

Elektronenholographie mittels Elektronenrückstreubeugung

Arbeit zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften

Matthias Gianfelice

geboren in Hagen

2022

Arbeitsgruppe Westphal

Fakultät Physik, Technische Universität Dortmund

1. Gutachter: Prof. Dr. Carsten Westphal
 2. Gutachter: Prof. Dr. Dr. Wolfgang Rhode
- Abgabedatum: 9. September 2022

Version: 9FB399D/222/29.11.2022

Zusammenfassung

Die von Dennis Gábor 1948 vorgestellte Holographie sollte zur besseren Auflösung des Elektronenmikroskops beitragen. Die Verwendung mit Elektronenwellen ließ aber noch auf sich warten: Erste experimentelle Umsetzung fand die Lichtholographie, in deren Rahmen die holographische Idee substantiell weiterentwickelt werden konnte. Durch Erzeugung eines Phasenkontrasts wird in einer zweidimensionalen Aufnahme neben der Amplituden- auch eine Phaseninformation gespeichert. Damit ist die Rekonstruktion der ursprünglichen Objektwelle und von jedem Betrachtungsort aus eine dreidimensionale Bildwiedergabe unter anderer Perspektive möglich.

Im Sinne der ersten Idee Gábors verfolgt die Elektronenholographie das Ziel, atomare Strukturen aufzulösen. Mittels der Computerholographie kann die Rekonstruktion der Kristalle sichtbar gemacht werden. In dieser Arbeit sollen Simulationen der Elektronenholographie mittels Electron Backscatter Diffraction (EBSD) auf planem Schirm zeigen, dass die erfolgreiche Rekonstruktion eines Wolframeinkristalls mit bereits heute vorhandenen Apparaturen möglich ist. Dazu wird zunächst ein theoretisches Fundament vorgestellt. Anhand der Schirmgröße und -auflösung wird ermittelt, unter welchen Voraussetzungen eine holographische Rekonstruktion möglich ist. Ermittelte Grenzwerte werden quantitativ erläutert. Durch Definition eines Bewertungskriteriums ist es möglich, die Rekonstruktionen automatisiert auszuwerten. Mit der Einführung der Additionsmethode werden die ursprünglichen Limitierungen in Teilen aufgehoben und die Rekonstruktion deutlich stabilisiert. Abschließend wird die experimentelle Umsetzbarkeit diskutiert.

Abstract

Holography, introduced by Dennis Gábor in 1948, was intended to contribute to a better resolution of the electron microscope. However, its use with electron waves still required further research at the time. The first experimental implementation was light holography, in the context of which the holographic idea could be substantially further developed. By generating a phase contrast, the phase information is stored in a two-dimensional image in addition to the amplitude information. This allows for the reconstruction of the original object wave and a three-dimensional image reproduction under a different perspective from any viewing point.

In accordance with Gábor's first idea, electron holography pursues the goal of resolving atomic structures. By using of computer holography the reconstruction of crystals can be made visible. In this work, simulations of electron holography using Electron Backscatter Diffraction (EBSD) on a planar screen will demonstrate that successful reconstruction of a tungsten single crystal is achievable with equipment currently available. For this purpose, a theoretical foundation is firstly prepared. Based on the screen size and resolution, it is determined under which conditions a holographic reconstruction is possible. The computed limit values are explained quantitatively. By defining a fitting criterion it is possible to evaluate the reconstructions automatically. With the introduction of the addition method the original limitations are improved and the reconstruction is significantly stabilized. Finally, the experimental feasibility is discussed.

*Was Dir widerfuhr, es mag vergehen.
Was Du daraus geformt, das soll bestehen.*
— Albertus Magnus, †1280

Michele und Tobias.
Sigrid und Walter.
Katharina.
Philip und Aphrodite.

Inhaltsverzeichnis

Veröffentlichungen und Konferenzteilnahmen	xi
1 Einleitung	1
2 Lichtholographie	5
2.1 Historie	5
2.2 Wellenbeschreibung	7
2.2.1 Ebene Welle	9
2.2.2 Kugelwelle	9
2.3 Herleitung	10
2.3.1 Rekonstruktion	12
2.3.2 Holographischer Prozess	14
2.3.3 Kohärenzbedingungen	16
2.4 Volumenhologramme	17
3 Elektronenrückstreubeugung	19
3.1 Elektronenbeugung	19
3.1.1 Wechselwirkungstiefe	19
3.1.2 Wellenform	21
3.2 EBSD	21
3.2.1 Kikuchi-Linien	22
3.2.2 Musterauswertung	24
3.2.3 Auflösungsvermögen	24
3.3 EBSD-Holographie	25
3.3.1 Differenzbildmethode	27
4 Computerholographie	31
4.1 Fourier-Synthese	31
4.1.1 Fourier-Reihen	32
4.1.2 Fourier-Transformation	33
4.2 Diskretisierung	34
4.2.1 Abtasttheorem	36
4.2.2 Aliasing	39
4.2.3 Diskrete Transformation	41
4.3 Holographie	42
4.3.1 Objekte	42
4.3.2 Wellenpropagation	43
4.3.3 Schirm	45
4.3.4 Rekonstruktion	46
4.3.5 Oversampling Ratio	46

5	Algorithmen	49
5.1	Punktbasiert	49
5.1.1	Faltungsmethode	51
5.1.2	Lookup-Tabellen	52
5.1.3	Wavefront Recording Plane	55
5.2	Weitere Methoden	56
5.3	Hardware	57
5.4	Eigenentwicklung	59
5.4.1	Wellenfelder	59
5.4.2	Koordinatentransformation	60
5.4.3	Kristallsampling	62
5.4.4	Rekonstruktion	62
6	Untersuchungen	65
6.1	Probensystem	65
6.1.1	Kristallstruktur	66
6.2	Aufbau	67
6.2.1	Lokale Umgebung	68
6.2.2	Rekonstruktionsvolumen	69
6.3	Bewertungskriterien	70
6.3.1	Klassifizierung	71
6.3.2	Wahrheitsmatrix	72
6.3.3	Gütekriterien	73
6.3.4	Atomkugeln	74
7	Ergebnisse	77
7.1	Standardexperiment	77
7.1.1	Visualisierung	80
7.2	F_1 -Score	82
7.3	Massenauswertung	85
7.3.1	Oversampling Ratio	87
7.3.2	Abtasttheorem	88
7.4	Musteraddition	91
7.4.1	$\zeta_{\mathcal{N}}$ -Score	93
8	Zusammenfassung und Ausblick	97
A	Holographischer Prozess	103
A.1	Intensität mehrerer Streuer	103
A.2	Allgemeine Rekonstruktion	104
A.3	Rückpropagation	105
B	Implementierung	107
B.1	WaveObject	108
B.1.1	Refwave	108
B.1.2	Wavefield	108
B.1.3	ConicWavefield	109

B.2	accessors	110
B.3	CyclePlot	110
B.4	crystals	111
B.4.1	crystals.Crystal	111
B.4.2	crystals.SuperCell	111
B.5	Orientation	112
B.6	utils	112
B.7	typing	112
C	Klassifizierungsgrößen	113
	Symbolverzeichnis	115
	Akronymverzeichnis	119
	Abbildungsverzeichnis	121
	Literaturverzeichnis	123
	Danksagungen	141

Veröffentlichungen und Konferenzteilnahmen

Veröffentlichungen

Teile dieser Arbeit wurden im folgenden Rahmen veröffentlicht:

- M. Gianfelice und C. Westphal
Electron holography by planar electron backscattered diffraction patterns
In: *Journal of the Optical Society of America A* 39.2 (2022), A1–A6
<https://doi.org/10.1364/JOSAA.442927>
- M. Gianfelice und C. Westphal
Simulated Imaging of Atomic Surface-Structures via Holographic Approaches by Planar Electron Backscattered Diffraction Patterns
In: *OSA Technical Digest JTU5.A.22* (2021)
<https://doi.org/10.1364/AIS.2021.JTu5A.22>

Konferenzteilnahmen

Teile dieser Arbeit wurden im folgenden Rahmen vorgestellt:

- M. Gianfelice und C. Westphal
Simulated Imaging of Atomic Surface-Structures via Holographic Approaches by Planar Electron Backscattered Diffraction Patterns
OSA Imaging and Applied Optics Congress, virtuell ausgetragen (2021)
- M. Gianfelice, C. Kohlmann und C. Westphal
Photoelectron holography—an experimentalist's approach to computer calculated reconstructions
DPG-Frühjahrstagung der Sektion Kondensierte Materie (SKM), Regensburg (2019)

1 Einleitung

Der vielleicht intuitivste Zugang zum Verständnis eines Systems erfolgt durch das unmittelbare Sehen. Seit jeher ist der Mensch darum bemüht, immer entferntere liegende Dinge sichtbar zu machen und immer kleinere Gegenstände aufzulösen.

Die Geschichte der Optik als Lehre des (sichtbaren) Lichts und seiner Ausbreitung beginnt dabei keineswegs erst mit dem Auftreten der ersten optischen Sehhilfen etwa um das 13. Jahrhundert [1, Kap. 1.2]. Vielmehr beschäftigt das Licht und seine Nutzbarmachung die Menschen bereits länger, als es schriftliche Aufzeichnungen gibt. So belegen Ausgrabungen in Ninive, dass bereits die Babylonier über Linsen und Metallspiegel verfügten, auch wenn deren Nutzung heute nicht klar ist und keine Kenntnisse über optische Gesetzmäßigkeiten voraussetzt [2, Kap. 1].

Die griechischen Philosophen beschäftigten sich ebenfalls mit dem Versuch einer theoretischen Beschreibung des Lichts. Maßgeblich ist dabei die Beschreibung der *Camera Obscura* durch Aristoteles [3, Kap. 5]. Er konnte eine partielle Sonnenfinsternis ohne die Gefahr vor Blendung dadurch beobachten, dass ein Abbild durch die Löcher eines Siebs auf den Boden projiziert wurde. Vergrößerte er den Abstand zwischen kleinem Projektionsloch und Abbildungsebene, vergrößerte sich auch das Abbild [4, Kap. 1]. Der arabische Gelehrte Al-Haitham stellte neben einer Strahltheorie zum menschlichen Sehen [5] auch fest, dass ein kleineres Loch eine schärfere Abbildung liefert [6, Kap. 1].

Giovanni Battista della Porta schlug vermutlich als erster vor, dieses Prinzip der Lochbildkamera zu nutzen, um Künstler zu unterstützen. Auf der Rückseite eines dem Loch gegenüberliegenden Schirms ließ sich ein Motiv abmalen [6, Kap. 1]. Durch Linsen und Spiegel war es dem Prämonstratenser Johann

Zahn möglich, die erste bekannte, portable Lochbildkamera zum Abmalen zu konstruieren. Dem Prinzip nach ist das bereits sehr mit der modernen Spiegelreflexkamera verwandt [3, Kap. 5].

Mit der von Joseph Nicéphore Niépce entwickelten *Heliographie* war es dann möglich, erstmals eine fotografische Abbildung anzufertigen [3, Kap. 19]. Eine mit Asphalt beschichtete Glasplatte konnte er nach der Belichtung mithilfe von Lavendelöl so aufbereiten, dass er ein lichtbeständiges Bild erhielt. Es handelte sich also um eine zweidimensionale Abbildung einer Szene, so wie sie auch das Auge sehen würde. Allerdings zunächst in einer negativen Darstellung, bevor er später ebenfalls Direktpositive aufnehmen konnte [6, Kap. 5].

Auch die vielen Weiterentwicklungen der letzten Jahre bis hin zur Farbfotografie [7] oder der Aufnahme mittels Halbleitersensoren in modernen Kameras [8] können einen Mangel aber nicht kompensieren: Durch die Projektion und deren Aufnahme geht viel Information verloren, nämlich die einer getreuen, dreidimensionalen Abbildung. Hierbei bedeutet getreu und dreidimensional die Erhaltung des Raumeindrucks. Die Fotografie zeigt auf einer Ebene immer die gleiche Perspektive, unabhängig der Betrachtungsrichtung.

Ein gewisser Tiefeneindruck kann durch die *Stereoskopie* dauerhaft eingefangen werden. Sie basiert auf der Grundidee, dass die menschliche Tiefenwahrnehmung durch zwei Augen realisiert wird. Das Gehirn ist in der Lage, aus zwei Bildern ein einziges zusammenzusetzen und die relative Entfernung einzelner Objekte zueinander abzuschätzen. Dazu wird eine Szene aus zwei Betrachtungswinkeln – nämlich an Stelle der Augen – aufgezeichnet [3, Kap. 6]. Wird nun jedem Auge durch eine Apparatur getrennt jeweils eines

der Bilder vorgelegt, entsteht auch nachträglich ein dreidimensionaler Raumeindruck [9, Kap. 3]. Heutzutage wird dieser Effekt mithilfe von Brillen mit Polarisations- oder Farbfiltren beispielsweise bei der Betrachtung entsprechend aufgezeichneter Kinofilme verwendet [10].

Allerdings wird auch hier kein vollständiges Objekt aufgezeichnet. Lediglich für eine feste Perspektive, die bei der Aufnahme festgelegt wurde, besteht nun eine persistente Abbildung. In erheblichem Unterschied dazu ist die *Holographie* in der Lage, eine echte Aufnahme des vollständigen Objekts zu machen und damit eine tatsächlich dreidimensionale Abbildung aufzuzeichnen [11, Kap. 1.1]. Statt lediglich die Intensität der vom abzubildenden Objekt ausgehenden Welle aufzunehmen, wird hier auch die Phaseninformation aufgezeichnet [12]. Dies gelingt durch Überlagerung mit einem kohärenten Hintergrund [13].

Ursprünglich für die Auflösung atomarer Strukturen mithilfe des Elektronenmikroskops entworfen, entwickelte Dennis Gábor 1948 die Holographie [14] und erhielt dafür 1971 den Nobelpreis [15]. Allerdings sollte noch etwas Zeit bis zur Umsetzung in atomarer Größenordnung vergehen. In der Zwischenzeit wurde die Lichtholographie – entsprechend der Benennung mit Wellenlängen sichtbaren Lichts – erfolgreich umgesetzt und substanziell weiterentwickelt. So war es möglich, echt dreidimensionale Objektwellen von einem zweidimensionalen Abbildungsmedium wiederzugeben.

Inzwischen haben Anwendungen der Holographie im Alltag an Bedeutung weitestgehend verloren, nachdem auch in Geschäften immer seltener Hologramme zu sehen sind. Lediglich Sicherheitsmerkmale, etwa auf Kreditkarten, sind noch bekannt [11]. Dabei werden durch Aufprägung Weißlicht-Reflexionshologramme realisiert, die auch unter diffusem Streulicht sichtbar sind [16, Kap. 3.5.6]. Auf Geldscheinen finden sich dagegen häufig mit dem Hologramm verwandte Kippbilder, sog. *Kinegramme*, bei denen es sich aber nicht um Hologramme im eigentlichen Sinne handelt. Ihre Abbildung ändert sich sprunghaft in Abhängigkeit vom Betrachtungswinkel [16, Kap. 3.5.7].

In der Forschung jedoch hat die Holographie signifikant an Bedeutung gewonnen und ist Gegenstand umfassender und hochaktueller Forschungsfragen geworden. Besonders im Rahmen der Computerholographie wird an echt dreidimensionalen Anzeigesystemen geforscht, die ohne die oben erwähnten stereoskopischen Brillen auskommen [17]. Nahe an Gábors ursprüngliche Idee heran kommt die Elektronenholographie, die wie von ihm vorgeschlagen mittels Elektronenwellen atomare Strukturen sichtbar macht [18].

Im Rahmen dieser Arbeit soll die Holographie mittels Elektronenrückstreubeugung vorgestellt werden. Dabei handelt es sich um ein etabliertes Verfahren in der Materialanalyse [19]. Wir werden zeigen, dass holographische Aufnahmen mit heute bereits zur Verfügung stehenden Apparaturen möglich sind und gleichzeitig eine Methode zur Verbesserung der Abbildung vorschlagen [20]. Ganz konkret ist damit also die Abbildung periodischer Strukturen in Größenordnungen weniger Ångström nach unseren Erkenntnissen möglich [21]. Über die Rekonstruktion mittels Computeralgorithmen werden Kristalle sicht- und damit unmittelbar begreifbar.

Zu Beginn soll daher erst in Kapitel 2 die Lichtholographie vorgestellt werden. Neben der Verfolgung des erwähnten historischen Ablaufs sind Lichtwellen intuitiver nachvollziehbar und optische Aufbauten dem Leser vermutlich vertrauter. Wir werden die Entwicklungsgeschichte ebenso wie eine mathematische Herleitung aus der Wellengleichung beleuchten.

Im Anschluss soll im Rahmen von Kapitel 3 die Elektronenrückstreubeugung als primärer Experimentalaufbau skizziert und erläutert werden. Ferner ist die Wellenbeschreibung von Elektronen nicht trivial [22]. Zusammen

mit der begrenzten Wechselwirkungstiefe bildet sie die Grundlage für die späteren Ausführungen. Die EBSD genannte Methode und ihre Messdaten werden erläutert, bevor gezeigt wird, wie diese Muster zur Holographie genutzt werden können. Insbesondere muss erläutert werden, wie sich die aus der Lichtholographie bekannten Bestandteile (Objekt, Referenz und Schirm) hier wiederfinden.

Es werden weiterhin Simulationen realer Apparaturen vorgestellt. Damit diese gelingen können, soll in Kapitel 4 erst auf Besonderheiten eingegangen werden, die bei der Portierung der holographischen Theorie auf diskrete, numerische Systeme auftreten. Verschiedene Effekte sind für die weitere Aufarbeitung relevant und werden beleuchtet. Insbesondere das Abtasttheorem [23, 24] und Aliasing sind ursächlich für physikalische Grenzen [11, Kap. 3.2.2 ff.].

Eine weitere Herausforderung stellt die Beschreibung der physikalischen Welt am Computer dar [25, Kap. 1.1]. Deren Modellierung soll in Abschnitt 4.3 erläutert werden, daraus folgt mit dem Oversampling Ratio σ ein weiterer, relevanter Parameter für das Experimentdesign [20]. Wir werden hier ferner erkennen, dass sich die Wellenausbreitung in einen Propagationsoperator formalisieren und die holographische Rekonstruktion mit genau demselben Prozess durchführen lässt [11, Kap. 5].

In der Computerholographie sind bereits viele Algorithmen zur Berechnung eines Hologramms etabliert [26]. Sie implementieren den Propagationsoperator und erlauben damit sowohl Holographie als auch Rekonstruktion. Es ist instruktiv, diese Algorithmen in Kapitel 5 kennenzulernen, um eine möglichst

effiziente Berechnungsgrundlage zu schaffen. Ferner soll auf spezialisierte Hardware und die im Rahmen dieser Arbeit entstandene Eigenentwicklung eingegangen werden.

In Kapitel 6 wird das simulierte Experiment genau beschrieben. Insbesondere soll der Wolframeinkristall näher erläutert werden, der als Probensystem aufgrund seiner einatomigen Basis gewählt wurde. Ferner werden vorbereitende Erläuterungen für die Einführung des F_1 -Score [27] als ein objektives Maß für die Übereinstimmung der Rekonstruktion mit der simulierten Realität getroffen.

Anhand dieses Maßes F_1 ist es dann in Kapitel 7 möglich, eine teilautomatisierte Auswertung vieler verschiedener Experimentalparameter durchzuführen. Im Rahmen dieser Arbeit wurden zwei Parameter gewählt, die miteinander korrelieren. Die Auflösung N^2 und die Kantenlänge d des quadratischen Schirms werden hier weiträumig überprüft. Daraus werden wir physikalische Grenzen in der Aufnahme aufzeigen, die sich in den Messdaten wiederfinden lassen. Durch Einführung der *Additionsmethode* können wir diese Ergebnisse noch weiter verbessern.

Abschließend soll nach einer Zusammenfassung in Kapitel 8 zunächst eingeordnet werden, inwieweit die hier erhobenen Erkenntnisse tatsächlich auch experimentell umsetzbar sind. Im Anschluss folgt ein Ausblick darauf, was nun für die weitere Erforschung zu tun bleibt. Im Anhang finden sich ferner weiterführende Erläuterungen zur holographischen Theorie mehrerer Streuer, der für die Simulationen entwickelten Bibliothek und mit dem F_1 -Score zusammenhängenden Klassifizierungsgrößen.

2 Lichtholographie

Im Kern zeichnet sich ein Hologramm gegenüber einer Fotografie durch die zusätzlich gespeicherte Phaseninformation aus. Dadurch ist eine echt dreidimensionale Wiedergabe des aufgenommenen Objekts möglich. Die Technik der Holographie wurde dabei ursprünglich für die Nutzung mit Elektronenwellen vorgestellt. Ihre erste rein praktische Anwendung hat sie jedoch in der Anwendung mit sichtbarem Licht gefunden.

Das vorliegende Kapitel soll in die Holographie einführen und anhand ihrer ersten Anwendung als Lichtholographie die Kernkonzepte erläutern, bevor sie in weiteren Kapiteln auf die für diese Arbeit relevante, atomare Größenordnung übertragen wird. Im Besonderen bietet sich dieser Zugang an, da sichtbares Licht alltäglich ist und viele optische Phänomene damit intuitiv nachvollziehbar sind.

Dazu soll zunächst die historische Entwicklung der Holographie dargelegt werden. Es schließt sich die mathematische Beschreibung von Wellen an, mit deren Hilfe sich dann auch der eigentliche holographische Prozess herleiten lässt. Abschließend soll mit Kohärenzbedingungen und Volumenhologrammen auf tiefergehende Themen der Lichtholographie eingegangen werden, die für die später vorgestellte Elektronenholographie zwar nicht unabdingbar, dennoch von großem Forschungsinteresse sind.

2.1 Historie

Anhand ihrer zeitlichen Entwicklungsgeschichte lassen sich einige Konzepte der Holographie nachvollziehen, die im weiteren Verlauf relevant werden. An dieser Stelle soll daher mithilfe der Historie schrittweise auf diese Grundlagen eingegangen werden. Denn die holographische Idee, die Dennis Gábor 1948

vorstellte und für die er 1971 den Nobelpreis der Physik erhielt, basiert auf vielen weiteren Arbeiten, die hier kurz vorgestellt werden sollen.

Bereits 1843 und damit 75 Jahre vor Gábors Holographie publizierte Ernst Abbe seine „Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung“. Dort beschrieb er insbesondere die Bildentstehung im Mikroskop als zweigeteilten Prozess [28], wie er in Abbildung 2.1 dargestellt ist.

Von dem abzubildenden Objekt gehen seiner Vorstellung nach Strahlen aus, die in der Brennebene des Objektivs interferieren und so ein Beugungsbild erzeugen – das *primäre Bild*. Dieses Bild ist rein virtuell, denn es wäre ja erst realisiert, wenn ein Detektor eingeschoben wird. Stattdessen kann nun dieses Bild als Ausgang weiterer Strahlen angenommen werden, die durch das Objektiv vergrößert werden und in einer weiteren Ebene zum *sekundären Bild* interferieren [28, Abschn. 6].

Durch diese Aufspaltung des Abbildungsprozesses an der objektnahen Brennebene des Objektivs lassen sich die beteiligten Prozesse jeweils durch die gleiche Transformation – die in Abschnitt 4.1.2 genauer erläuterte Fourier-Transformation – beschreiben. Damit war die Fourier-Optik geboren [30].

Dabei ist diese Aufteilung in zwei Beugungsvorgänge keineswegs nur eine theoretische Spielerei zur einfacheren Erklärung im Rahmen der geometrischen Optik. Mieczysław Wolfke schlug 1920 die experimentelle Umsetzung durch ein „Röntgenfoto“ vor [31]. Röntgenstrahlen haben deutlich kleinere Wellenlängen als sichtbares Licht und können damit kleinere Strukturen auflösen, sind aber nur sehr begrenzt für das menschliche Auge direkt sichtbar [32]. Wolfkes Idee nach sollte das mit Röntgenwellen entstandene Primärbild aufgenommen und mit sichtbarem Licht

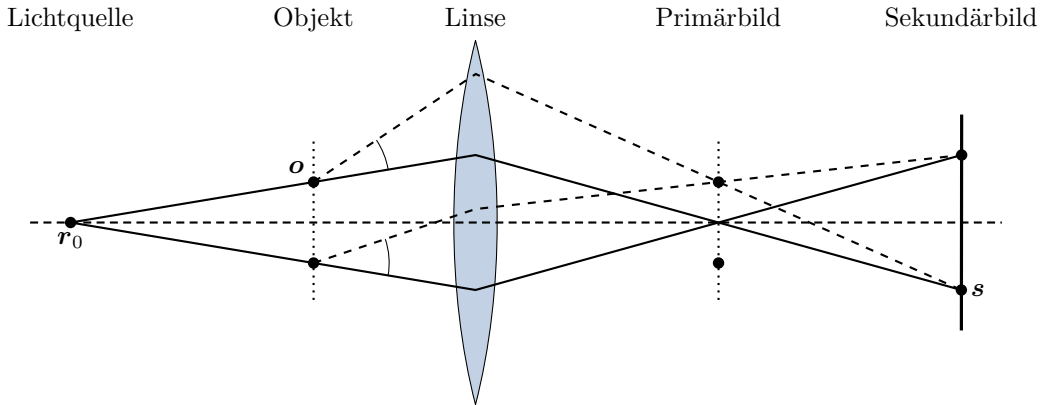


Abbildung 2.1: Abbildung im Mikroskop nach Abbe'scher Theorie zur Bildentstehung. Es wird zwischen einem primären Bild in der Bildebene der Lichtquelle und dem sekundären Bild des Gegenstands unterschieden. Darstellung nach Ref. [29, Abb. 1].

wiedergegeben werden. Er schränkte bereits hier ein, dass diese Fotografie nur eine Intensitätsverteilung aufnehmen kann, bei der die Phaseninformation der Objektwelle verloren geht. Neben monochromer, paralleler und senkrechter Beleuchtung ist damit auch ein symmetrisches Abbildungsobjekt nötig, denn erst dann ist das Beugungsbild des Beugungsbilds identisch mit dem Abbild des Objekts. Die große Ähnlichkeit zum holographischen Prozess ist kein Zufall: Das primäre Bild kann bereits als ein theoretischer Vorläufer des Hologramms verstanden werden [33, Kap. 1.2.1].

Hans Boersch gelang diese zweischrittige Abbildung durch ein anderes Verfahren. Das primäre Bild imitierte er durch Blenden, die er dann mit sichtbarem Licht bestrahlte [34]. Der Technik nach ist das bereits die erste überhaupt durchgeführte Holographie [35].

Sir William Lawrence Bragg entwickelte nach diesem Prinzip das *Röntgenstrahl-Mikroskop*. Die wesentlichen Reflexe einer Gitterbeugung bildete er durch Bohrungen in einer Messingplatte nach. Die Löcher stellten damit ebenfalls ein reziprokes Gitter dar, das wie bei Boersch nun auch mit sichtbarem Licht beleuchtet wurde [36]. Die Reflexe dieses Git-

ters konnten damit jedoch keine Phasenverschiebung abbilden. Durch das spätere Einsetzen von Glimmerplättchen in die Bohrlöcher wurde dies sehr begrenzt möglich. Das zusätzliche Material veränderte die Phase lokal, sodass die Darstellungen der Kristalle sichtbar besser mit der Kristallstruktur übereinstimmte [37–39].

Eine fundamentale Grundidee der Holographie formulierte erstmals Frederik Zernike mit der Nutzung eines „kohärenten Hintergrunds“ im Rahmen seiner *Phasenkontrastmethode* [13]. Durch die Interferenz der Objektwelle und eines kohärenten Hintergrunds entsteht ein Phasenkontrast, der sich in eine Intensitätsverteilung überführen lässt, wo bisher nur ein Amplitudenbild möglich war [12].

Bei seinen Versuchen, die Auflösung des von ihm genutzten Elektronenmikroskops zu verbessern, verband Gábor 1948 dann die Ideen Wolfkes und Zernikes zur holographischen Methode [14]. Dabei wählte er den Namen *Hologramm* für das aufgenommene Primärbild aus dem Griechischen $\delta\lambda\omicron\varsigma$ (holos, deutsch „ganz“), $\gamma\rho\acute{\alpha}\varphi\omega$ (grapho, deutsch „schreiben“) [40].

Bei der Erzeugung der kohärenten Strahlung machte sich Gábor die Kohärenzbedingung nach Verdet zunutze. Bei Verwendung einer hinreichend kleinen Lochblende verhält sich das durch sie dringende Licht wie das von einer Punktquelle am Ort der Blende ausgehende [41]. Bei seinen Experimenten mit sichtbarem Licht stellte Gábor aber fest: Die von ihm verwendete Geradeausholographie überlagert im Sekundärbild mit einem konjugiertem Bild, dessen Ursprung weiter unten erläutert werden soll.

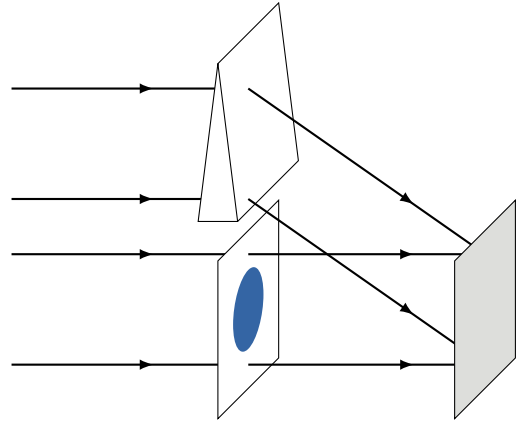
Emmett Leith und Juris Upatnieks schafften es 1962, das konjugierte Bild aus dem rekonstruierten zu legen, indem sie einen Winkel zwischen Referenz- und Objektwelle einführten [44]. Ermöglicht wird das durch die Zerteilung des Referenzstrahls, der entsprechend sehr kohärent sein muss. Die von ihnen verwendeten und vorgeschlagenen Aufbauten sind schematisch in Abbildung 2.2 dargestellt. Durch den 1960 in Ref. [45] neu vorgestellten Laser war das experimentell umsetzbar, während Gábor diese Möglichkeit noch nicht zur Verfügung stand [33, Kap. 1.2.2].

2.2 Wellenbeschreibung

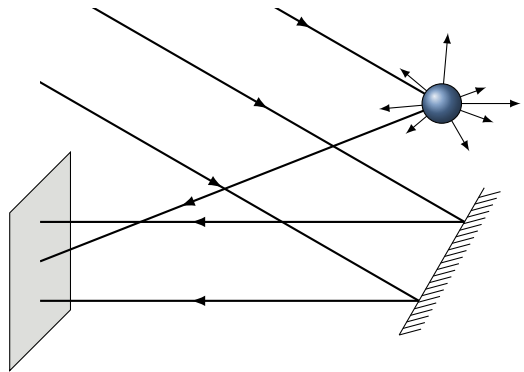
Für die weitere Behandlung ist zunächst ein grundlegendes mathematisches Verständnis von Licht nötig, das sich als elektromagnetische Welle durch eine Differentialgleichung beschreiben lässt. D'Alembert hat eine solche *Wellengleichung* erstmals 1747 für die Beschreibung schwingender Saiten eingeführt [1, Kap. 2.1.1].

In diesem Abschnitt soll nun allgemein eine Herleitung der Wellengleichung für elektromagnetische Wellen dargelegt werden, die im Wesentlichen Ref. [46, Kap. 4.3.1] folgt. Dazu sei zunächst ein Vakuum angenommen. Wir gehen also von einem elektromagnetischen Feld in einem ungeladenen Isolator mit Ladungsdichte

$$\rho = 0 \quad (2.1)$$



(a) Aufbau für die *Durchlichtholographie* nach Ref. [42]. Die Referenzwelle wird mithilfe eines Prismas schräg mit der durch das Objekt gedämpften Referenzwelle überlagert. Das wäre des Weiteren auch mit einem Spiegelpaar möglich.



(b) Darstellung der *Auflichtholographie* nach Ref. [43]. Ein Objekt wird mithilfe der Referenzwelle beleuchtet. Sie wird ferner mithilfe eines Spiegels am Hologramm anschließend mit der Objektwelle überlagert.

Abbildung 2.2: Stilisierte Darstellung zweier typischer Arten von Experimenten zur Aufnahme von Lichthologrammen wie sie Leith und Upatnieks [42, 43] vorschlugen.

sowie der vektoriellen elektrischen Stromdichte

$$\mathbf{j} = \mathbf{0} \quad (2.2)$$

aus.

Ferner nehmen wir zur Vereinfachung ein lineares und homogenes Medium an. Die elektrische und die magnetische Flussdichte lauten dann respektive

$$\begin{aligned} \mathbf{D} &= \varepsilon_r \varepsilon_0 \mathbf{E} \quad \text{und} \\ \mathbf{B} &= \mu_r \mu_0 \mathbf{H} \end{aligned} \quad (2.3a)$$

in Abhängigkeit der elektrischen und magnetischen Feldstärken \mathbf{E} und \mathbf{H} .

Die *Maxwell-Gleichungen* beschreiben elektromagnetische Felder und den Zusammenhang zwischen der elektrischen und magnetischen Komponente. Sie lauten im vorgestellten Fall

$$\nabla \cdot \mathbf{E} = 0, \quad (2.4a)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (2.4b)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad \text{und} \quad (2.4c)$$

$$\nabla \times \mathbf{B} = \varepsilon_r \varepsilon_0 \mu_r \mu_0 \frac{\partial \mathbf{E}}{\partial t} \quad (2.4d)$$

mit dem *Nabla-Operator* ∇ .

An dieser Stelle sei in Erinnerung gerufen, dass die doppelte Anwendung einer Rotation auf ein beliebiges Vektorfeld \mathbf{V} zu dem Ausdruck

$$\begin{aligned} \nabla \times (\nabla \times \mathbf{V}) &= \nabla(\nabla \cdot \mathbf{V}) - \nabla \cdot (\nabla \mathbf{V}) \\ &= \nabla(\nabla \cdot \mathbf{V}) - \nabla^2 \mathbf{V} \end{aligned} \quad (2.5)$$

mit dem verkürzenden *Laplace-Operator*

$$\nabla^2 \mathbf{V} \equiv \nabla \cdot (\nabla \mathbf{V}) \quad (2.6)$$

führt.

Ferner sei die Definition der Vakuumlichtgeschwindigkeit

$$c_0 = 1/\sqrt{\varepsilon_0 \mu_0} \quad (2.7)$$

wiederholt. In einem Medium erfährt sie eine Dämpfung um den sog. *Brechungsindex* n und lautet

$$c = \frac{1}{\sqrt{\varepsilon_0 \varepsilon_r \mu_0 \mu_r}} \stackrel{(2.7)}{=} \frac{c_0}{\sqrt{\varepsilon_r \mu_r}} =: \frac{c_0}{n}. \quad (2.8)$$

Mit den zuvor erwähnten Gleichungen erhalten wir aus der Anwendung der Rotation auf die jeweils linke und rechte Seite der Gleichung (2.4c) und unter Berücksichtigung des Satzes von Schwarz über die Vertauschbarkeit partieller Differentiationen respektive

$$\begin{aligned} \nabla \times (\nabla \times \mathbf{E}) \\ \stackrel{(2.5)}{=} \underbrace{\nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E}}_{=0 \text{ (2.4a)}} = -\nabla^2 \mathbf{E}, \end{aligned} \quad (2.9a)$$

$$\begin{aligned} \nabla \times \left(-\frac{\partial}{\partial t} \mathbf{B} \right) &= -\frac{\partial}{\partial t} (\nabla \times \mathbf{B}) \\ \stackrel{(2.4d)}{=} -\varepsilon_r \varepsilon_0 \mu_r \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} &= -\frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2}. \end{aligned} \quad (2.9b)$$

Beide Gleichungen (2.9) lassen sich nach Gleichung (2.4c) weiterhin gleichsetzen und ergeben dann

$$\nabla^2 \mathbf{E} - \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0. \quad (2.10)$$

Analog lässt sich auch mit Gleichung (2.4d) verfahren. Nach Bildung der Rotation auf beiden Seiten ergibt sich so

$$\nabla^2 \mathbf{B} - \frac{1}{c^2} \frac{\partial^2 \mathbf{B}}{\partial t^2} = 0. \quad (2.11)$$

Bei den Gleichungen (2.10) und (2.11) handelt es sich bereits um die Wellengleichung, die von beiden Komponenten des elektromagnetischen Feldes erfüllt wird. Sie lässt sich in die homogene Wellengleichung

$$\square \psi(\mathbf{r}, t) = 0. \quad (2.12)$$

verallgemeinern.

In der Notation (2.12) wird eine abkürzende Schreibweise in Gestalt des *D'Alembert-Operators*

$$\square\psi \equiv \nabla^2\psi - \frac{1}{c^2} \frac{\partial^2}{\partial t^2}\psi \quad (2.13)$$

verwendet.

2.2.1 Ebene Welle

Diese Wellengleichung (2.12) wird beispielsweise durch eine sog. *ebene Welle*

$$\psi(\mathbf{r}, t) = A e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)} \quad (2.14)$$

mit konstanter *Amplitude* A gelöst [1, Kap. 2.7]. Das Argument der Exponentialfunktion ohne imaginäre Einheit i wird als *Phase* der Welle bezeichnet.

Bei einer Momentaufnahme zu einer beliebigen aber festen Zeit $t = t_0$ bilden sich Flächen gleicher Phase aus. Diese erfüllen die Gleichung

$$\mathbf{k} \cdot \mathbf{r} = \text{const}, \quad (2.15)$$

sind also Ebenen, die senkrecht auf \mathbf{k} stehen. In Richtung des Vektors \mathbf{k} treten diese *Wellenfronten* genannten Ebenen gleicher Phase äquidistant immer wieder in einem Abstand von

$$\Delta\mathbf{r}_n \cdot \mathbf{k} = 2\pi n, \quad n \in \mathbb{Z} \quad (2.16)$$

zueinander auf. Zwei benachbarte Wellenfronten gleicher Phase, die für die ebene Welle namensgebend sind, werden dabei durch eine *Wellenlänge*

$$\lambda = 2\pi/k \quad (2.17)$$

voneinander getrennt. Das zugehörige \mathbf{k} wird daher auch *Wellenvektor* genannt [46, Kap. 4.3.2].

Der in der Lichthologie häufig genutzte Laser kann gut durch eine räumlich begrenzte ebene Welle (2.14) angenähert werden [47, Kap. 12.1]. Das Laserlicht verfügt

ebenfalls über eine dedizierte Ausbreitungsrichtung mit Wellenfronten konstanter Amplitude. Jedoch fehlt ihnen die unendliche Ausbreitung der ebenen Welle. Es ist also nötig, sie in ihrer Beschreibung räumlich zu begrenzen – etwa durch eine Amplitudenfunktion in Form einer Blende. In der Realität gilt es noch zu berücksichtigen, dass der Strahl mit zunehmendem Abstand zur Quelle aufweitet und bei großen Distanzen inkohärent wird.

Es sei bemerkt, dass die ebene Welle in der vorgestellten Form (2.14) zunächst eine rein mathematische Lösung der Wellengleichung (2.12) ist und ansonsten kein real existentes Pendant hat. Die mathematische ebene Welle dehnt sich unendlich weit aus, was in der Realität so nie zutreffen kann. Dennoch ist sie in ihrer Anwendung sehr nützlich, da sich auch komplexe Wellenfronten als lokal stark eingeschränkte ebene Wellen approximieren lassen [11, Kap. 3.2.3].

2.2.2 Kugelwelle

Da später für die Behandlung der Elektronen-holographie im Rahmen dieser Arbeit eine andere Form der Welle von besonderer Bedeutung ist, sollen alle weiteren Phänomene anhand einer anderen Lösung der Wellengleichung (2.12) erläutert werden: der *Kugelwelle*

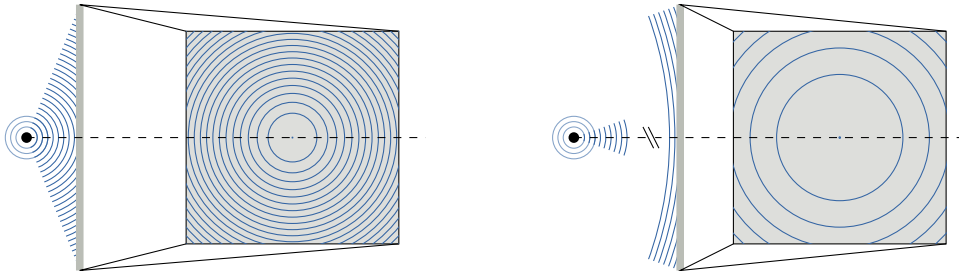
$$\psi(r, t) = \frac{A}{kr} e^{-i(kr - \omega t)}. \quad (2.18)$$

Sie wird hauptsächlich durch den Abstand r von ihrem Ausgangszentrum beschrieben [46, Kap. 4.3.5].

Liegt der Mittelpunkt \mathbf{r}_0 der Kugelwelle nicht im Nullpunkt, ist also $\mathbf{r}_0 \neq \mathbf{0}$, gilt im Allgemeinen

$$r(\mathbf{x}) = |\mathbf{x} - \mathbf{r}_0|. \quad (2.19)$$

Die Wellenfronten gleicher Phase bilden sich dabei als Kugelflächen mit Radius r um das Ausgangszentrum aus. Die Überlegungen



(a) Die Quelle befindet sich nah vor dem Schirm: Die Wellenfronten liegen eng beieinander.

(b) Die Quelle hat einen deutlich größeren Abstand zum Schirm: Der Abstand zwischen den Wellenfronten erhöht sich.

Abbildung 2.3: Eine Kugelwelle befindet sich zentral vor einem Schirm in unterschiedlichem Abstand. Ausgezeichnet sind die Wellenfronten, also die Flächen gleicher Phase. Die Darstellung folgt schematisch der Illustrationen in Ref. [48, Abb. 1 f.].

zu Wellenlänge und Wellenvektor gelten hier analog zur ebenen Welle, insbesondere gilt auch Gleichung (2.17).

Die Intensität muss aus Gründen der Energieerhaltung bei der Ausbreitung der Wellenfronten konstant bleiben [49, Kap. 34.4]. Im Falle der Kugelwelle beträgt die Fläche einer Wellenfront

$$O = 4\pi r^2. \quad (2.20)$$

Damit muss die Intensität im isotropen Fall invers dazu, also

$$I \stackrel{!}{\propto} \frac{1}{O} \propto \frac{1}{r^2} \quad (2.21)$$

sein. Sie berechnet sich als Betragsquadrat der Welle und ist daher proportional zur quadrierten Amplitude. Deren $1/r$ -Abhängigkeit führt zur geforderten Energieerhaltung

$$I = |\psi|^2 = \psi \cdot \psi^* \stackrel{(2.18)}{=} \left(\frac{A}{kr}\right)^2 \propto \frac{1}{r^2}. \quad (2.22)$$

Die Normierung auf die Länge des Wellenvektors k macht die Amplitude einheitenlos. Ferner erlaubt sie die Skalierung von Wellenlänge λ und Abständen r zur selben Zeit, ohne dass sich die zugrundeliegende Physik ändert: Skalierungseffekte heben sich in beiden Größen auf. Insbesondere im folgenden Kapitel 3

für die Behandlung atomarer Größenordnungen wird dies von besonderer Bedeutung sein. Bei Verkleinerung der Wellenlänge λ und damit Vergrößerung des Wellenvektorbetrags k werden die relevanten Abstände r in der Regel im gleichen Maße kleiner. Die Gesamtamplitude der Wellenfunktion bleibt also konstant.

Ferner ist die Wellengleichung (2.12) invariant unter Phasenverschiebungen $\Delta\varphi$. Da

$$e^{i\Delta\varphi} = \text{const} \quad (2.23)$$

gilt, ist im Allgemeinen

$$\square(\psi e^{i\Delta\varphi}) = e^{i\Delta\varphi} \underbrace{\square\psi}_{=0 \text{ (2.12)}} = 0. \quad (2.24)$$

Damit gilt für alle vorgestellten Wellen, dass sie auch unter einer konstanten Phasenverschiebung die Wellengleichung (2.12) erfüllen.

2.3 Herleitung

Gábor hatte 1948 wie beschrieben eine Methode vorstellen wollen, die das Auflösungsvermögen eines Elektronenmikroskops erhöht, um noch kleinere Strukturen sichtbar

zu machen. Da die Lichtholographie aber viel zugänglicher ist – immerhin ist sie unmittelbar für das Auge sicht- und nachvollziehbar – hat auch er mit lichtholographischen Experimenten begonnen [14]. Sie unterscheiden sich von denen, die auf die später erläuterten Elektronenholographie zurückgreifen, primär lediglich in einem Parameter: der Wellenlänge λ . Hier soll daher zunächst die Herleitung der (Licht-)Holographie im Allgemeinen beschrieben werden, bevor wir deren Prinzipien in Kapitel 3 auf Elektronenwellen übertragen.

Die in Gleichung (2.18) angenommene Kugelwelle soll für die weitere Betrachtung des holographischen Prozesses herangezogen werden. Da alle Prozesse zu einem bestimmten Zeitpunkt beobachtet werden und die zeitliche Entwicklung vernachlässigt werden kann, wird auf die Berücksichtigung der Zeitabhängigkeit verzichtet. Der Mittelpunkt der Kugelwelle liege an einem Ort \mathbf{r}_0 . Die Wellenfunktion lautet damit dann

$$\psi(\mathbf{r}) = \frac{A}{k|\mathbf{r} - \mathbf{r}_0|} e^{-i(k|\mathbf{r} - \mathbf{r}_0| - \varphi)}. \quad (2.25)$$

Gelingt es nun durch Abbildung der Phaseninformation, die Wellenfronten sichtbar zu machen, wird die Holographie auch anschaulich verständlich. Abbildung 2.3 stellt dazu zwei Szenarien gegenüber. Liegt ein Kugelwellenursprung nah am Aufnahmeschirm, so liegen die Wellenfronten kreisförmig um den Lotpunkt unter dem Kugelwellenursprung in geringem Abstand. Mit zunehmendem Abstand zum Schirm nimmt auch der Abstand der Kreisinge zueinander zu, sodass sich geometrisch die Lage und der Abstand des Quellenpunkts zur Aufnahmeebene bestimmen ließe. Seine Grenzen erfährt diese einfache Konstruktion dann, wenn der Ursprung nicht über der Aufnahmeebene und weit entfernt liegt [48].

Zur Herleitung der relevanten Wellenfunktionen sei zunächst ein abzubildendes Objekt angenommen, das lediglich aus einem Punkt be-

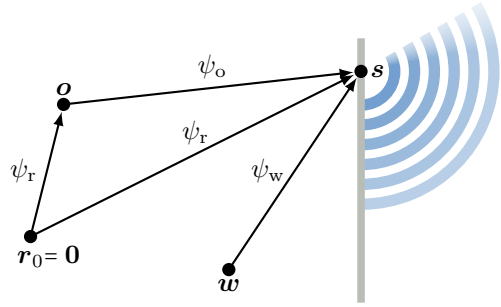


Abbildung 2.4: Skizze zur Herleitung der auf dem Schirm entstehenden Intensität und deren Wiedergabe: Eine Referenzwelle ψ_r beleuchtet ein Objekt. Die dort entstehende Objektwelle ψ_o überlagert mit der Referenzwelle am Schirm. Durch Beleuchtung mit einer Wiedergabewelle ψ_w entsteht eine neue Welle. Im Falle der holographischen Rekonstruktion entspricht sie im Wesentlichen der ursprünglichen Objektwelle.

steht. Diese Vereinfachung entspricht selbstverständlich nicht der Realität mit deutlich komplexeren Objekten. Das Ergebnis bleibt aber auch bei der Summe vieler überlagerter Wellen das gleiche, wie in Anhang A gezeigt wird.

Jede beliebige Wellenfront lässt sich nach dem Huygens'schen Prinzip als Ausgangspunkt vieler weiterer sogenannter Partikularwellen interpretieren [50, Kap. 8.2]. Entsprechend lassen sich auch annähernd beliebige Objekte und ihre Objektwellen holographieren, indem sie virtuell in viele Teilwellen mit jeweils punktförmigen Ursprung aufgeteilt werden.

Das Objekt befinde sich, wie in Abbildung 2.4 gezeigt, am Ort \mathbf{o} . Die zugehörige Wellenfunktion lautet dann

$$\psi_o(\mathbf{r}) = \frac{A_o}{k|\mathbf{r} - \mathbf{o}|} e^{-i(k|\mathbf{r} - \mathbf{o}| - \varphi_o)} \quad (2.26)$$

mit Objektamplitude A_o und nicht näher spezifizierter, konstanter Phasenverschiebung φ_o .

Die Intensität beträgt

$$I(\mathbf{r}) = |\psi(\mathbf{r})|^2. \quad (2.27)$$

Im Falle der Objektwelle ψ_o wird sie am Schirm \mathbf{s} gemessen – also wie in Abbildung 2.5a dargestellt aussehen. Wir erhalten

$$I(\mathbf{s}) = |\psi_o|^2 = \psi_o \cdot \psi_o^* = \left(\frac{A_o}{k|\mathbf{s} - \mathbf{o}|} \right)^2. \quad (2.28)$$

Es ist also lediglich die reine Amplitudeninformation in Gestalt von A_o erhalten.

Um die Phase sichtbar zu machen, ist nach Zernike die Überlagerung mit einer zweiten, kohärenten Welle nötig [12]. Sie wird mit gleicher Frequenz und damit einem Wellenvektor gleichen Betrags k gewählt. Diese wird als *Referenzwelle* ψ_r bezeichnet und habe ohne Beschränkung der Allgemeinheit eine Phasenverschiebung $\varphi = 0$ und ihren Ursprung im Nullpunkt $\mathbf{0}$. Damit ergibt sich ihre Wellenfunktion zu

$$\psi_r(\mathbf{r}) = \frac{A_r}{kr} e^{-ikr}. \quad (2.29)$$

Die Gleichheit zwischen Objekt- und Referenzwellenlänge ist dann in jedem Falle sichergestellt, wenn die Referenzwelle auch zur Beleuchtung des Objekts genutzt wird und damit ursächlich für die Auslösung der Objektwelle ist. In diesem Fall ergibt sich eine Objektwellenamplitude von

$$A_o \propto |\psi_r(\mathbf{o})| \quad (2.30)$$

und eine Phasenverschiebung von

$$\varphi_o = \text{phase}[\psi_r(\mathbf{o})] + \varphi'. \quad (2.31)$$

Aus der Überlagerung beider Wellen und

$$\cos x = \frac{1}{2}(e^{+ix} + e^{-ix}) \quad (2.32)$$

erhalten wir die Intensität der Überlagerung zu

$$I(\mathbf{r}) = |\psi_o(\mathbf{r}) + \psi_r(\mathbf{r})|^2 = \bar{A}(\mathbf{r}) + 2 \frac{A_o A_r}{k^2 |\mathbf{r} - \mathbf{o}| r} \cos(k(|\mathbf{r} - \mathbf{o}| - r) - \varphi_o) \quad (2.33)$$

mit einem allgemeinen Amplitudenoffset

$$\bar{A}(\mathbf{r}) := \frac{A_o^2}{k^2 |\mathbf{r} - \mathbf{o}|^2} + \frac{A_r^2}{k^2 r^2}. \quad (2.34)$$

Diese Intensität (2.33) wird nun auf dem Schirm in irgendeiner Form aufgezeichnet. In früheren Experimenten geschah dies etwa durch eine Fotoplatte, die dann eine Schwärzung T in Abhängigkeit der Intensität aufweist. Deren Intensitätsabhängigkeit ist durchaus relevant für die Rekonstruktion – allerdings stark abhängig von der konkret verwendeten Methodik der Aufnahme. Im Folgenden soll die Abhängigkeit von der Intensität als linear approximiert werden. Dann entsteht ein konstanter, vernachlässigbarer Faktor.

Konkret sei eine virtuelle Schwärzung

$$T(I) = \frac{I}{I_{\max}} \propto I \quad (2.35)$$

angenommen, die der normierten Intensität entspricht.

Ein Hologramm im Folgenden enthält also ein „Schwärmungsmuster“ T , das direkt proportional zur Intensität ist. Zu jedem Kugelwellenursprung entsteht darin ein Kreisring-system. Beim Übergang auf mehrere Streuzentren in der Objektwelle überlagern sich deren Systeme im Muster, wie es beispielhaft in Abbildung 2.5b–d dargestellt ist.

2.3.1 Rekonstruktion

Die geschwärmte Platte soll nun mit einer Wiedergabewelle

$$\psi_w(\mathbf{r}) = \frac{A_w}{\kappa |\mathbf{r} - \mathbf{w}|} e^{-i(\kappa |\mathbf{r} - \mathbf{w}| - \varphi_w)} \quad (2.36)$$

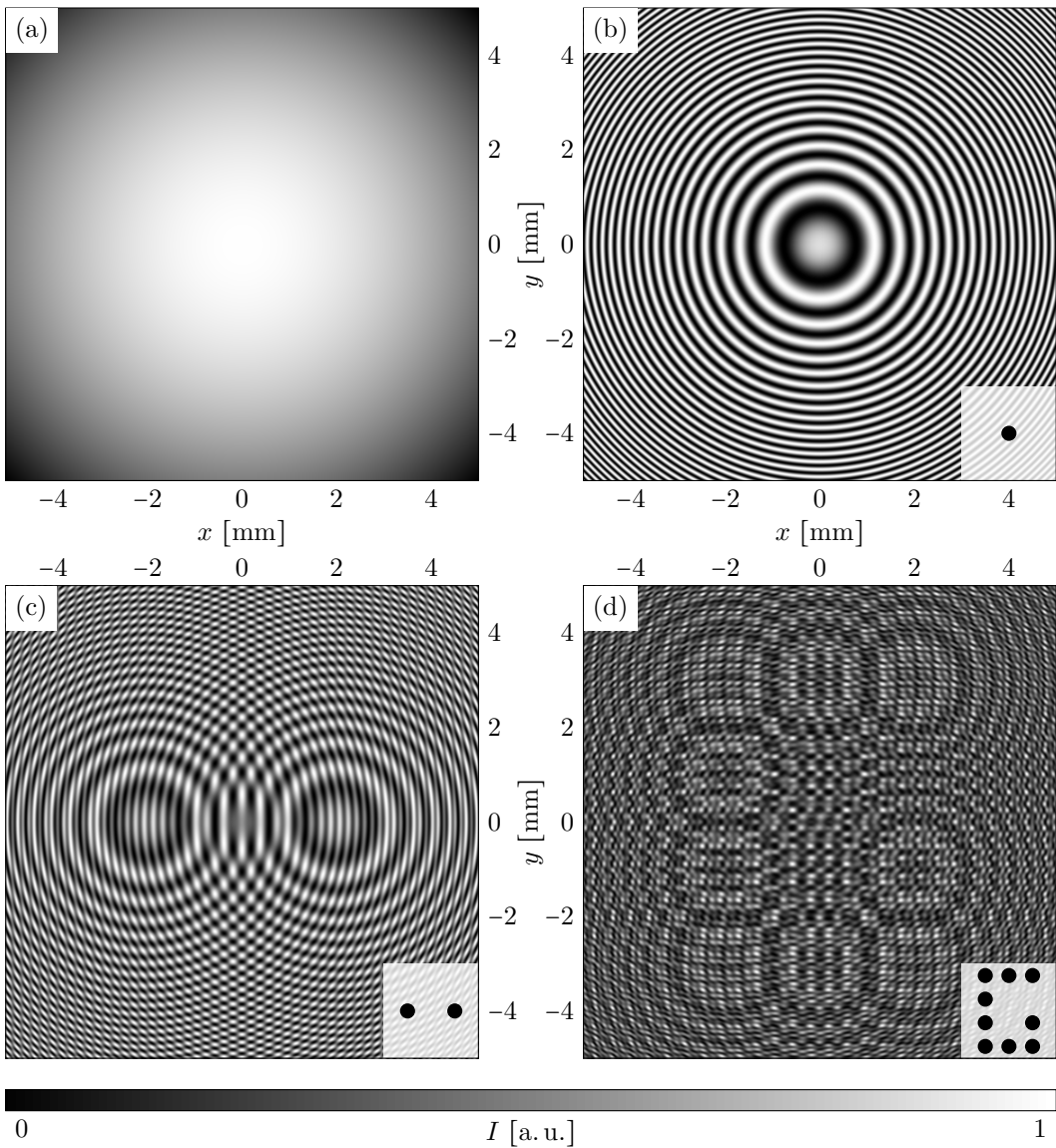


Abbildung 2.5: Darstellungen von Simulationsergebnissen holographischer Rechnungen auf einem Schirm nach Normierung der Intensität. (a) Intensitätsmuster der einfachen Kugelwelle, ohne jede Überlagerung. (b)–(d) Überlagerung der dargestellten Bilder aus ein, zwei und neun Punktstreuer mit einer Kugelwelle als Referenz. Jeder Punktstreuer erzeugt ein eigenes Kreisringssystem, die hier miteinander interferieren. Mit zunehmender Zahl an Punktstreuern wird es – wie im Falle des stilisierten Buchstaben G in (d) – schwieriger, das ursprüngliche Bild zu erraten.

mit Wellenvektor $\boldsymbol{\kappa}$, Ursprung \boldsymbol{w} und konstanter Phasenverschiebung φ_w beleuchtet werden. Das Muster der Schwärzung dämpft die Wiedergabewelle ab und erzeugt eine weitere Welle mit entsprechend abgeschwächter Amplitude. Es ergibt sich damit am Schirm mit Gleichung (2.32) eine Überlagerung

$$\begin{aligned} T[I(\boldsymbol{s})] \cdot \psi_w(\boldsymbol{s}) &\propto \bar{A}(\boldsymbol{s}) \cdot \psi_w(\boldsymbol{s}) \\ &+ \hat{A}(\boldsymbol{s}) e^{i[k(|\boldsymbol{s}-\boldsymbol{o}|-s)-\kappa|\boldsymbol{s}-\boldsymbol{w}|- \varphi_o - \varphi_w]} \\ &+ \hat{A}(\boldsymbol{s}) e^{-i[k(|\boldsymbol{s}-\boldsymbol{o}|-s)+\kappa|\boldsymbol{s}-\boldsymbol{w}|- \varphi_o + \varphi_w]} \end{aligned} \quad (2.37)$$

mit der Amplitude

$$\hat{A}(\boldsymbol{s}) = \frac{A_o A_r}{k^2 |\boldsymbol{s}-\boldsymbol{o}|s} \cdot \frac{A_w}{\kappa |\boldsymbol{s}-\boldsymbol{w}|}. \quad (2.38)$$

Hier soll der Spezialfall gleicher Wellenlänge von Referenz- und Wiedergabewelle und damit gleichem Wellenvektorbetrag betrachtet werden. Gleichzeitig gehe sie ebenfalls vom Ursprung der Referenzwelle aus. Es gilt also

$$k = \kappa \quad \text{und} \quad \boldsymbol{w} = \mathbf{0}. \quad (2.39)$$

Damit entsprechen sich beide Wellen bis auf eine konstante Phasenverschiebung. Wir erhalten am Ort des Schirms komplexe Wellenamplituden von

$$\begin{aligned} T[I(\boldsymbol{s})] \cdot \psi_w(\boldsymbol{s}) &\propto \bar{A}(\boldsymbol{s}) \cdot \psi_w(\boldsymbol{s}) \\ &+ \hat{A}(\boldsymbol{s}) e^{i(k(|\boldsymbol{s}-\boldsymbol{o}|-2s)-\varphi_o - \varphi_w)} \\ &+ \hat{A}(\boldsymbol{s}) e^{-i(k|\boldsymbol{s}-\boldsymbol{o}|- \varphi_o + \varphi_w)}. \end{aligned} \quad (2.40)$$

Das Ergebnis in Gleichung (2.40) setzt sich aus drei Termen zusammen. Beim ersten Term handelt es sich um eine einfache Dämpfung der Wiedergabewelle. Der dritte ist von besonderem Interesse, denn er besteht – von einem Amplitudenfaktor und einer konstanten Phasenverschiebung φ_w abgesehen – aus der ursprünglichen Objektwelle ψ_o . Im zweiten Term findet sich deren konjugierte Form ψ_o^* mit einer zusätzlichen Phasenverschiebung

$$\Delta\varphi = 2ks + \varphi_w. \quad (2.41)$$

Bei dieser neuen Welle handelt es sich um das konjugierte Bild, das Gábor durch seine Überlagerung mit dem eigentlichen Bild in der Rekonstruktion Schwierigkeiten bereite [11, Kap. 7.6]. In der Literatur wird es auch häufig als *Zwillingsbild* bezeichnet [51].

Dass in der rekonstruierten Welle im Wesentlichen die Objektwelle ψ_o enthalten ist, hat eine weitreichende Konsequenz. Denn für einen Beobachter ist nun nicht mehr unterscheidbar, ob diese Objektwelle tatsächlich von dem angestrahlten, ursprünglichen Objekt stammt oder eine entsprechende „Simulation“ aus einem Hologramm ist. Damit ist das durch das Hologramm erzeugte Bild nach Abbildung auf einen Detektor – wie etwa dem menschlichen Auge – nicht mehr von einem echten Objekt zu unterscheiden. Das echte Objekt wird dem Beobachter damit aufgrund der Äquivalenz der ursprünglichen und der rekonstruierten Objektwelle „vorgetäuscht“.

Im wesentlichen Unterschied zu einer Fotografie ist aber insbesondere möglich, das Objekt dreidimensional zu erfahren. Durch die Imitation der Objektwelle an jeder Stelle im Raum ist dadurch insbesondere auch die Veränderung der Lage des Detektors bei gleichzeitiger Veränderung der wahrgenommenen Perspektive und Parallaxe möglich.

In letzter Konsequenz bedeutet das auch, dass diese holographische Objektwelle ihrerseits selbst wie die eines Objekts fotografiert werden kann. Dabei lässt sich sogar der Kamerafokus beliebig einstellen, als wäre das Objekt real vorhanden. Es lassen sich damit insbesondere verschiedene Stellen des aufgenommenen Objekts fokussieren.

2.3.2 Holographischer Prozess

Der gesamte Prozess der Holographie ist in Abbildung 2.6 schematisch dargestellt und soll aufgrund seiner Bedeutung hier erneut kurz zusammengefasst werden. Das exemplarische Objekt besteht aus zwölf Streuzentren.

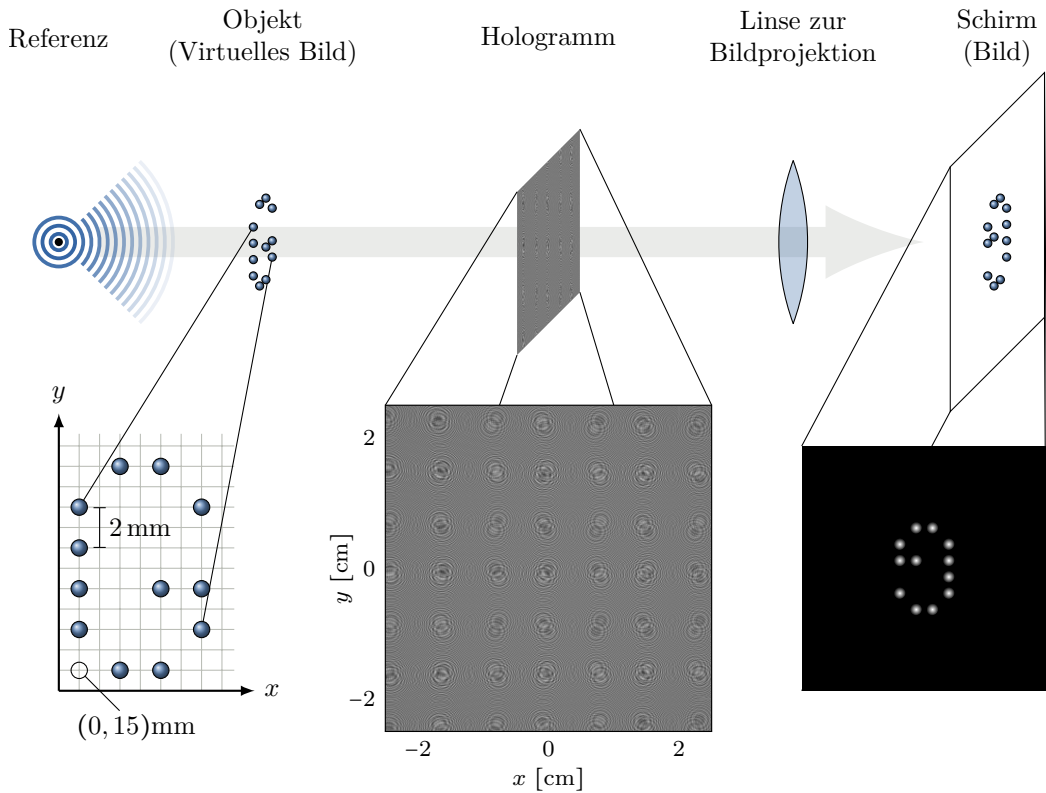


Abbildung 2.6: Schematischer, nicht maßstäblicher Aufbau des vollständigen holographischen Prozesses mit Rekonstruktion: Das Objekt besteht hier aus wenigen Streupunkten und ergibt einen stilisierten Buchstaben G. Es wird durch die Referenzwelle beleuchtet und erzeugt das gezeigte Hologramm. Entfernt man nun das Objekt und beleuchtet das Hologramm mit der gleichen Referenzwelle, wird die Objektwelle derart initiiert, dass durch Verwendung einer Linse das Bild auf einen Schirm projiziert werden kann – ganz, wie es etwa auch im Auge geschieht. Es wird ein virtuelles Bild an Stelle des Objekts wahrgenommen, das auf dem Schirm ausgemacht werden kann. Durch die geometrische Abbildung ist das Bild hier hinsichtlich beider Achsen gespiegelt. Vergleiche auch Ref. [52, Abb. 1.8].

Die Referenzwelle beleuchtet diese und überlagert mit der summierten Objektwelle als Hologramm.

Wird nun das Objekt aus dem Strahlengang entfernt und das Hologramm mit derselben Referenzwelle beleuchtet, wird die Objektwelle wie oben beschrieben imitiert.

Mit einer Linse lässt sich das Bild auf einen Schirm projizieren. Es entsteht ein aus optischen Gründen jeweils an der x - und y -Achse gespiegeltes Bild auf dem Schirm. Dem Betrachter scheint das ursprüngliche Objekt nun als virtuelles Bild. Es ist ihm nicht möglich, zwischen dem echten Objekt und dem durch das Hologramm simulierten virtuellen Bild als Herkunft der Objektinformation zu unterscheiden.

Im vorliegenden Beispiel wurde das Objekt um 15 mm in x -Richtung von der optischen Achse entfernt. Das erleichtert die Rekonstruktion, da die auf der optischen Achse dominante Referenzwelle dort weniger stark auftritt und sich somit leichter unterdrücken lässt. Ferner entsteht das imaginäre Bild nicht im Abbildungsvolumen und kann somit nicht mit dem gewünschten Bild überlagern. Für die Darstellung des Bilds auf dem Schirm wurde diese Verschiebung kompensiert.

Es ist also bereits in diesem einfachen Beispiel gelungen, ein zweidimensional gespeichertes Bild getreu zu rekonstruieren. Dabei ist inzwischen klar, dass das Hologramm nicht vergleichbar zu einer Fotografie ist. Stattdessen ist nicht unterscheidbar, ob es überhaupt je eine Aufnahme gegeben hat. Im Falle der Lichthologie rekonstruiert hier also ein echt räumliches Bild der aufgezeichneten Gegenstände.

Lediglich die Referenzwelle ist dabei möglichst exakt zu reproduzieren. Daher bietet es sich grundsätzlich bei holographischen Experimenten an, eine möglichst einfache Geometrie zu wählen, die im Anschluss zur Rekonstruktion leicht nachzustellen ist.

2.3.3 Kohärenzbedingungen

Bisher ist von idealen Wellen und Rahmenbedingungen ausgegangen worden. Insbesondere wurden Zeitabhängigkeiten vernachlässigt. Inwieweit die Holographie auch außerhalb der simulierten Idealwelt funktioniert, soll hier kurz dargestellt werden.

Insbesondere sind die Anforderungen an die Kohärenz der beteiligten Wellen groß. Das bedeutet, dass die Wellen zu jedem beliebigen Zeitpunkt t und an jedem Ort \mathbf{r} eine konstante Phasenverschiebung zueinander aufweisen müssen. Quantisieren lässt sich die Abhängigkeit des Rekonstruktionserfolgs von der Kohärenz durch die *Theorie der partiellen Kohärenz*.

In den theoretischen Ableitungen oben wurde von einer monochromen, unendlich kohärenten, perfekt punktförmigen Lichtquelle ausgegangen. In der Realität weist eine Strahlungsquelle aber eine gewisse Ausdehnung auf. Deren Störung lässt sich als Fourier-Reihe aus unendlich vielen monochromatischen Wellen darstellen [53]. Im Falle der Punktquelle wird dabei nur ein Term behandelt. Ein weiterer Störeinfluss stammt von Amplitude und Phase einer realen Lichtquelle, die natürlicherweise Fluktuationen aufweisen [50, Kap. 10].

Um diese Einflüsse quantisieren zu können, definieren wir den komplexen Kohärenzgrad

$$\gamma = \frac{\langle V_o V_r^* \rangle}{\sqrt{\langle V_o V_o^* \rangle \langle V_r V_r^* \rangle}}. \quad (2.42)$$

Sein Betrag gibt an, inwieweit Objekt- und Referenzwelle interferieren. Bei V_o und V_r handelt es sich gemäß der Definition aus Ref. [50, Kap. 10.2] um Beträge der sog. *analytischen Signale*, die aus den Erregerwellen ψ_o und ψ_r hervorgehen.

Im Rahmen der Theorie der partiellen Kohärenz ist die Intensität als zeitliches Mittel über diese Signale definiert mittels

$$I = \langle V V^* \rangle = \frac{1}{2t_0} \int_{-t_0}^{t_0} V(t) V^*(t) dt. \quad (2.43)$$

Ein Kohärenzgradbetrag von $|\gamma| = 0$ bedeutet also, dass die beiden Wellen zueinander inkohärent sind, die Intensitäten also einfach überlagert sind. Ein Betrag $|\gamma| = 1$ hingegen gibt an, dass sie voll miteinander interferieren.

Der Kontrast Λ des Hologramms lässt sich definieren als

$$\Lambda = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}}. \quad (2.44)$$

Der maximale Kontrast Λ_{\max} ist begrenzt durch den Kohärenzgrad, es gilt nach Ref. [33, Kap. 4.2.3]

$$\Lambda_{\max} = |\gamma|. \quad (2.45)$$

Es lässt sich etwa für eine Aufnahme über das Zeitintervall einer Sekunde abschätzen, dass eine Abweichung der Frequenz von bereits $\Delta\omega = 0,5 \text{ Hz}$ ausreicht, um die Rekonstruktion vollständig unmöglich zu gestalten. Mit Abnahme der Aufnahmezeit nimmt diese Frequenztoleranz antiproportional zu [33, Kap. 4.2.4]. Die Verwendung verschiedener Lichtquellen ist damit nahezu ausgeschlossen. Insbesondere können Bewegungen des aufgenommenen Objekts Probleme verursachen, wenn sich beispielsweise durch Dopplerverschiebung die Wellenlänge ändert [54].

In lightholographischen Experimenten gilt es daher insbesondere auch, Vibrationen zu minimieren. Schwingungsamplituden der lichtumlenkenden Bauteile dürfen dabei eine Größenordnung von ungefähr $\lambda/20$ nicht überschreiten. Dies gelingt durch aufwändige und schwere Aufbauten zur Erzeugung weitgehend vibrationsfreier Experimente [33, Kap. 5.2.2]. Beispielsweise sind schwere Granitplatten für Experimentiertische im Gegensatz zu Stahl deutlich besser in der Lage, mechanische Schwingungen zu dämpfen. Eine Gummilagerung dieser Tischplatten reduziert den Eintrag der Gebäudeschwingungen in den Experimentalaufbau. Ferner werden diese Experimentiertische präferiert unmittelbar auf dem Fundament platziert.

2.4 Volumenhologramme

Bisher wurde von einer zweidimensionalen Fläche als Aufnahmemedium ausgegangen. Diese Art der Intensitätsspeicher werden daher *Flächenhologramme* genannt. Tatsächlich verfügen insbesondere Fotoplatten aber über eine gewisse räumliche Ausdehnung senkrecht zur Fläche, die ebenfalls geschwärzt wird. Es handelt sich um dreidimensional zu behandelnde *Volumenhologramme*. Sie sind für die weitere Arbeit nicht von besonderer Relevanz, sollen aber aufgrund eines besonderen Einsatzgebiets erwähnt werden: der *Weißlichtholographie*.

Zur Erstellung von Hologrammen, die verschiedene Farben gleichzeitig rekonstruieren, schlugen Leith und Upatnieks vor, mehrere Flächenhologramme aufeinander zu schichten [43]. Für eine Auswahl an Grundfarben wird je ein Hologramm auf eine Platte aufgezeichnet. Nun werden diese Platten übereinander gelegt und unter Verwendung der gleichen Farben rekonstruiert.

Ein Problem wird dabei sofort erkennbar: Die Einzelbilder werden auch mit nicht der jeweiligen Referenzwelle entsprechenden Farbkomponenten rekonstruiert – es entstehen unerwünschte Fehlerbilder. Dies ließe sich in gewissem Umfang vermeiden, indem – ähnlich wie bei der schräg eingestrahelten Referenzwelle zur Verschiebung des konjugierten Bilds – hier die verschiedenen Farbreferenzwellen unter unterschiedlichen Winkeln eingestrahlt werden. Freilich muss dies bei der Rekonstruktion aber in exakt gleicher Art und Weise geschehen, was einen sehr hohen experimentellen Aufwand bedeutet.

Im Volumenhologramm ist die Rekonstruktion unter einem gleichen Winkel aber denkbar. Denn mit größerem Wellenlängenabstand $\Delta\lambda$ zwischen Referenz- und Wiedergabewellenlänge wird das unerwünschte Bild immer schwächer. Weißlichtholographie ist mit Volumenhologrammen also grundsätzlich einfacher realisierbar [33, Kap. 3.4.2].

3 Elektronenrückstreuung

Nachdem zuvor die Holographie mit sichtbarem Licht behandelt wurde, skalieren wir in diesem Kapitel die untersuchten Objekte auf die Größenordnung der atomaren Abstände in Kristallen. Dafür ist zunächst die Erweiterung der Wellentheorie auf Elektronen nötig. Zu Beginn dieses Kapitels sollen daher zunächst einige relevante Welleneigenschaften beschrieben werden.

Im Anschluss soll mit der Elektronenrückstreuung die Untersuchungsmethode vorgestellt werden, auf der die hier behandelte Art der Elektronenholographie basiert und bereits Anwendung bei der Untersuchung atomarer Kristalle findet. Durch die Einführung der abschließend erläuterten Differenzbildmethode können mit dieser Art der Beugung auch holographische Untersuchungen durchgeführt werden. Dort sollen die Feinheiten der experimentellen Anwendung beschrieben werden, bevor sie in späteren Kapiteln algorithmisch umgesetzt werden.

3.1 Elektronenbeugung

De Broglie postulierte 1924 im Rahmen seiner Dissertation, dass der bisher für Photonen angenommene Welle-Teilchen-Dualismus für jede Materie Anwendung finden kann [55]. Auch Elektronen lassen sich so Welleneigenschaften zuordnen, die mit einer energieabhängigen, sog. *de-Broglie-Wellenlänge*

$$\lambda_B = \frac{h}{\sqrt{2m_e E_{\text{kin}}}} \quad (3.1)$$

beschrieben werden.

Damit lassen sich auch mit Materiewellen Beugungsexperimente durchführen. Bekannte Experimente mit sichtbarem Licht – eben auch die in Kapitel 2 für Licht behandelte Holographie – können auf diese deutlich kleineren Wellenlängen λ_B übertragen werden. Da

diese in oder unterhalb der Größenordnung atomarer Abstände liegen, lassen sie sich insbesondere zur Untersuchung von Kristallen nutzen.

In der Festkörperphysik wird die Wahl einer Probe typischerweise auf diese Kristalle eingeschränkt. Dabei handelt es sich um Festkörper, deren intrinsische Struktur durch Bravais-Gitter mit zugehörigen Atombasen beschrieben werden kann. In der Folge zeigen alle Atome eine langreichweitige Ordnung mit diskreter Translationsinvarianz [56, Kap. 1.1].

Allerdings kann nicht der gesamte Kristall analysiert werden. Aufgrund einer begrenzten Reichweite der Elektronen haben lediglich nah an der Oberfläche liegende Schichten Einfluss auf die Messergebnisse.

3.1.1 Wechselwirkungstiefe

Die eingestrahlenen Elektronen können nur bis zu einer bestimmten Tiefe in den Kristall eindringen. Diese Wechselwirkungstiefe ist durch die *mittlere freie Weglänge* – auch Inelastic Mean Free Path (IMFP) genannt – λ_{IMFP} gegeben. Er wird im weiteren Verlauf dieser Arbeit noch relevant sein, weshalb er an dieser Stelle bereits eingeführt werden soll.

Primär ist der IMFP von der kinetischen Energie, deutlich weniger dominant vom Material abhängig [57]. Für die phänomenologische Betrachtung kann die verwandte Mean Escape Depth (MED) hinzugezogen werden, die die maximale Tiefe detektierter Elektronen bei Spektroskopie-Messungen angibt und ähnlich zur IMFP skaliert [58]. Die typische Bindungsenergie eines Valenzelektrons liegt im Festkörper in einer Größenordnung von etwa $E_B \approx 10 \text{ eV}$. Für Energien weit oberhalb dieser Bindungsenergie $E_{\text{kin}} \gg E_B$ können wir die Valenzelektronen mit nun vernachlässigbaren

Bindungen also als freies Elektronengas annähern. Daher wird die MED auch erst für höhere Energien angegeben. Die Elektronenabsorption wird damit primär durch eine Abhängigkeit von der mittleren Elektron-Elektron-Distanz beschrieben, die für alle Materialien annähernd gleich ist [59, Kap. 1.2].

Bei der Berechnung von Wechselwirkungstiefen wird hier die von Shinotsuka et al. vorgeschlagene Approximation aus Ref. [60] verwendet, die diese bereits in Einheiten von nm angibt. Dabei soll an dieser Stelle zum Zwecke der Vergleichbarkeit die Konvention übernommen werden, Fitparameter ohne Einheit anzugeben. Die Ergebnisse der Berechnungen sind hier in den Einheiten zu lesen, die nachstehend geklammert angegeben sind. Die mittlere freie Weglänge lautet approximiert dann

$$\lambda_{\text{IMFP}}(E_{\text{kin}}) = \frac{\hat{E}}{E_{\text{p}}^2 \left[\beta \ln(\gamma \hat{E}) - \frac{C}{E_{\text{kin}}} + \frac{D}{E_{\text{kin}}^2} \right]} \quad [\text{nm}]. \quad (3.2)$$

Dabei geht Gleichung (3.2) mit dem Energieparameter

$$\hat{E} \equiv E_{\text{kin}} \quad (3.3)$$

zurück auf eine nicht-relativistische Form, die aus der Bethe-Bloch-Gleichung abgeleitet wird [61]. In ihr finden sich bereits alle anderen Parameter in äquivalenter Form wieder. Insbesondere sind das die kinetische Elektronenenergie E_{kin} und die Plasmonenenergie

$$E_{\text{p}} = 28,816 \cdot \sqrt{\frac{N_{\text{v}} \cdot \rho}{M}} \quad [\text{eV}], \quad (3.4)$$

also die Energie der quantisierten Elektronenoszillationen.

Beide Energien sind – zusammen mit der später verwendeten Bandlückenenergie E_{g} – jeweils bereits in eV angegeben. N_{v} beschreibt

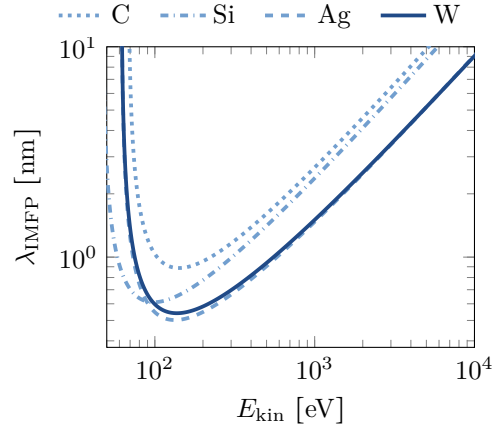


Abbildung 3.1: Die gezeigte Kurve nach Refs. [60, 61] wird aufgrund der qualitativen Ähnlichkeit für viele Festkörper als „Universelle Kurve“ bezeichnet.

die Zahl der Valenzelektronen je Atom oder Molekül. ρ ist die Dichte des Materials in Einheiten von g cm^{-3} , M seine atomare Masse.

Bei den übrigen Parametern handelt es sich dabei um Werte, die sich empirisch aus den Messdaten ermitteln lassen. Diese Parameter lauten im Einzelnen

$$\beta = -1,0 + \frac{9,44}{\sqrt{E_{\text{p}}^2 + E_{\text{g}}^2}} + 0,69\rho^{0,1} \quad [\text{eV}^{-1} \text{ nm}^{-1}], \quad (3.5a)$$

$$\gamma = 0,191\rho^{-0,5} \quad [\text{eV}^{-1}], \quad (3.5b)$$

$$C = 19,7 - 9,1 \cdot U \quad [\text{nm}^{-1}], \quad (3.5c)$$

$$D = 534 - 208 \cdot U \quad [\text{eV nm}^{-1}], \quad (3.5d)$$

$$U = \frac{N_{\text{v}} \cdot \rho}{M} \stackrel{(3.4)}{=} \left(\frac{E_{\text{p}}}{28,816} \right)^2, \quad (3.5e)$$

wobei sie nur unwesentlich von denjenigen früherer Publikationen mit nicht-relativistischer Form abweichen [60, 61]. Mit dem relativistischen Korrekturfaktor

$$\alpha(E_{\text{kin}}) = \left(1 + \frac{E_{\text{kin}}}{(2m_{\text{e}}c_0^2)^2} \right) \left(1 + \frac{E_{\text{kin}}}{m_{\text{e}}c_0^2} \right)^{-2} \quad (3.6)$$

ist die in Gleichung (3.2) angegebene Schätzung mit Energie-Parameter

$$\hat{E} = \alpha(E_{\text{kin}}) \cdot E_{\text{kin}} \quad (3.7)$$

bis zu einer kinetischen Energie von $E_{\text{kin}} = 200 \text{ keV}$ und somit für die in dieser Arbeit relevanten Energiebereiche gültig [60].

Diese in Abbildung 3.1 gezeigte Funktion wird aufgrund ihrer Ähnlichkeit für verschiedene Festkörper auch häufig *Universelle Kurve* genannt – obwohl diese Benennung in die Irre führt [62, Kap. 2.1.2]. Sehr wohl sind nämlich materialabhängige Effekte abseits des qualitativ stets ähnlichen Verlaufs erkennbar [63].

3.1.2 Wellenform

Für die in Abschnitt 2.3 beschriebene Herleitung der Lichthographie wurde eine isotrope Wellenausbreitung angenommen. Die Beschreibung der Wellenfunktion von Elektronen gestaltet sich dabei aber deutlich schwieriger, da sie stark anisotrop ist [22]. Insbesondere ist daher die Annahme reiner Kugelwellen nicht haltbar. Um holographische Experimente durchführen zu können, bedarf es also zunächst einer anderen mathematischen Beschreibung dieser Wellenform.

Bei hohen kinetischen Energien der Elektronen bildet sich ein Kegel in Ausbreitungsrichtung der Wellen aus. Die Streuamplitude außerhalb dieses Kegels, der durch den Öffnungswinkel Φ bestimmt wird, ist schnell vernachlässigbar [64]. Bei den im Rahmen dieser Arbeit untersuchten kinetischen Energien von $E_{\text{kin}} \geq 10 \text{ keV}$ verschwindet die Amplitude außerhalb eines Öffnungswinkels von $\Phi \geq 20^\circ$ [65]. Innerhalb dieses Kegels kann die Welle als isotrop approximiert und damit gleich der Form einer Kugelwelle angenommen werden [66]. Diese konische Wellenform, wie sie in Abbildung 3.2 qualitativ dargestellt ist, soll in allen weiteren Berechnungen angenommen werden.

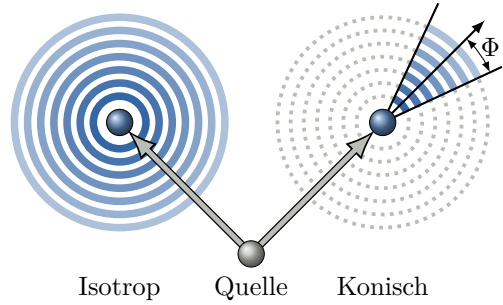


Abbildung 3.2: Gegenüberstellung der relevanten Wellenformen: Die isotrope Welle links kann für Elektronen nicht angenommen werden, stattdessen kann für große Energien die konische Form rechts angenähert werden. Innerhalb eines Kegels mit Öffnungswinkel Φ ist die Wellenausbreitung approximiert gleich der isotropen, außerhalb des Kegels Null.

3.2 EBSD

Vier Jahre nach de Broglies Postulat über die Materiewellen untersuchten Nishikawa und Kikuchi erstmals Beugungsbilder von Elektronen, die sie mit einer kinetischen Energie von $E_{\text{kin}} = 50 \text{ keV}$ unter streifendem Einfall eines Winkels von $\vartheta = 6^\circ$ auf eine Spaltebene eines Calcitkristalls (CaCO_3) lenkten. Deren de-Broglie-Wellenlänge beträgt nach Gleichung (3.1)

$$\lambda_B \approx 5,5 \text{ pm} = 0,055 \text{ \AA} \quad (3.8)$$

und liegt damit weit unter den atomaren Abständen innerhalb des Kristalls. Dort streuen sie daher am Kristallgitter.

Aufgrund des in Abschnitt 3.1.1 beschriebenen IMFP sind für die Holographie lediglich inelastische Interaktionen an einem oberflächennahen Atom relevant. Im Anschluss können die Elektronen nun ein- oder mehrfach an weiteren Atomen elastisch beugen oder unmittelbar aus der Probe austreten. Auf einem Schirm, der im vorgestellten Experiment in

einem Abstand von $D = 6,4$ cm zur Probe positioniert und analog zu lithographischen Experimenten durch eine Fotoplatte realisiert ist, kann dann ein sog. *Kikuchi-Muster* aufgezeichnet werden. Die verwendete Technik wurde dabei aufgrund der Rückstreuung der Elektronen zunächst Backscatter Kikuchi Diffraction (BKD) genannt, bevor sie später in Electron Backscatter Diffraction (EBSD) umbenannt wurde [67].

3.2.1 Kikuchi-Linien

Im aufgenommenen Muster sind Streifen – die *Kikuchi-Linien* – erkennbar. Vergleiche dazu auch Abbildung 3.6 weiter unten. Diese Linien treten dabei auch häufig parallel in Paaren mit geringem Abstand auf. Eine dieser Linien wird mit einer hohen Intensität aufgenommen, erscheint im Muster also hell, die andere regulär mit geringer Intensität, erscheint also dunkel. Sie grenzen einen Bereich mit zum Untergrund verschiedener Intensität ein. Dieses und vergleichbare Systeme werden daher als *Kikuchi-Bänder* bezeichnet [68].

Die Bildung der Kikuchi-Linien ist dabei zunächst qualitativ analog zu den bei der Röntgenstrahlung als *Kossel-Linien* bezeichneten Beugungserscheinungen zu erklären [69, 70]. Diese lassen sich aus der Bragg-Gleichung herleiten. Geometrisch wird diese anhand von Abbildung 3.3 leicht nachvollziehbar. Eintreffende Strahlen fallen unter dem Winkel ϑ auf den aus mehreren, im Abstand a aufeinander angeordneten Gitterebenen bestehenden Kristall. Die reflektierten Strahlen haben damit einen direkt ablesbaren summierten Gangunterschied von

$$\Delta = 2a \sin \vartheta. \quad (3.9)$$

Die Strahlen interferieren nun genau dann konstruktiv, wenn dieser Gangunterschied Δ einem ganzzahligen Vielfachen $n \in \mathbb{N}$ der Wellenlänge λ entspricht. Es ergibt sich die *Bragg-Bedingung* [56, Kap. 2.2.1]

$$\Delta = 2a \sin \vartheta = n\lambda. \quad (3.10)$$

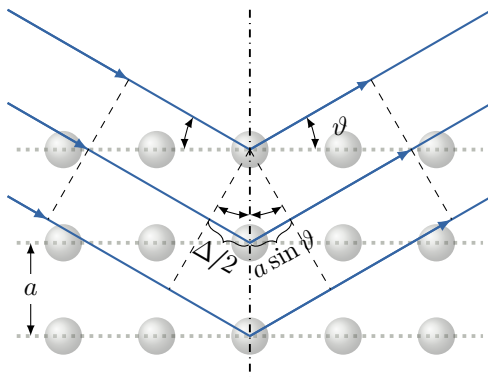


Abbildung 3.3: Geometrische Darstellung zur Ableitung der Bragg-Bedingung.

Die Kossel-Linien entstehen, wenn die Atome im Kristall zur isotropen Emission von Röntgenstrahlung angeregt werden. Ein Teil dieser Strahlung reflektiert an den Netzebenen und interferiert dann mit der unreflektierten Strahlung. Da Wellenlänge λ und Gitterkonstante a für das System gegeben sind, ist auch der Winkel konstant. Die Intensitätsmaxima breiten sich also in der Form eines Kegels – dem Kossel-Kegel – aus. Auf dem Detektor wird der Schnitt von Kegel und Schirmebene als Kossel-Linie nachgewiesen [71, Kap. 3.9]. Aus deren Lage lässt sich die Orientierung der für die Beugung verantwortlichen Netzebenen und damit die des Kristalls ablesen [70].

Im Falle der Kikuchi-Linien kann deren Form durch ähnliche geometrische Überlegungen aus zwei Konstruktionen des reziproken Gitters abgeleitet werden. Zum einen ist das die Brillouin-Zone, die primitive Zelle des reziproken Gitters – vgl. Ref. [56, Kap. 2.1.4], zum anderen die Ewaldkugel, die die von-Laue-Bedingung visualisiert [56, Kap. 2.2.2]. Die Bragg-Bedingung (3.10) ist erfüllt für Schnitte zwischen den Brillouin-Zonengrenzen und einer Ewaldkugel-Konstruktion um eine punktförmige Quelle innerhalb des Kristalls. Diese Schnitte lassen sich dann auf den Schirm projizieren [72]. Es handelt sich also um eine Zentralprojektion der

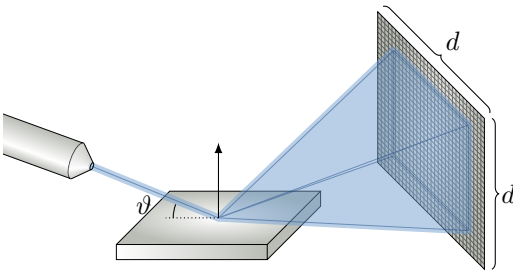


Abbildung 3.4: Schematische Darstellung des experimentellen EBSD-Aufbaus. Der eintretende Elektronenstrahl ist dabei zur Probenoberfläche verkippt [20, Abb. 1a].

Schnitte auf der Ewaldkugel aus der Quelle auf den Schirm: eine sog. *gnomonische* Projektion.

Die quantitative Erklärung ist dabei komplexer und die hier erbrachte qualitative Beschreibung über Kossel-Kegel für das vollständige Verständnis der Kikuchi-Bänder sicher nicht ausreichend. Ein Grund dafür liegt darin, dass die Atome die Elektronen inkohärent zur einfallenden Welle streuen – die klassische Theorie aber nur für die kohärente Beugung anwendbar ist [73]. Bisher wurden auch weiter auftretende Effekte vernachlässigt, die hier an Bedeutung gewinnen [74, 75]. So verringert auftretende Absorption die Intensität, während Mehrfachstreuung zusätzliche Interferenzen erzeugt. Erst durch von Laues relativistische Theorie zur Beschreibung von „Materiewellen und ihre(n) Interferenzen“ ließen sich die auftretenden Muster auch quantitativ erläutern [71, Kap. 5.25]. Inzwischen ist durch ein umfangreicheres Verständnis und die gesteigerten Kapazitäten moderner Rechner auch die Simulation dieser Kikuchi-Bänder sehr genau möglich, für weitere Erläuterungen sei auf Ref. [76] verwiesen.

Bei üblichen EBSD-Maschinen wird der Elektronenstrahl heute, wie in Abbildung 3.4 dargestellt, typischerweise in einem Winkel von $\vartheta = 20^\circ\text{--}30^\circ$ zur Probenoberfläche ein-

gestrahlt [67]. Auch Nishikawa und Kikuchi haben ihre Probe wie oben beschrieben unter streifendem Einfall bestrahlt. Denn ohne einen Winkel ϑ fern ab des Lots ist der Kontrast nicht ausreichend, um die Kikuchi-Bänder überhaupt ablesen zu können. So sind beispielsweise die Muster von Lithiumfluorid (LiF) für Winkel von $\vartheta \geq 35^\circ$ überhaupt nicht mehr sichtbar [77].

Das wird anschaulich bei Betrachtung der Rutherford-Streuung von geladenen Teilchen klar [77]. Im zugehörigen Erstexperiment wurden dazu α -Teilchen an einer Goldfolie gestreut. Die Anzahl der gestreuten Teilchen ließ sich demnach durch die *Rutherford'sche Streuformel* [78, Kap. 1.3.3] beschreiben:

$$dN_\alpha = N_\alpha n \Delta x \frac{Z^2 e^4}{(4\pi\epsilon_0)^2 m_\alpha^2 v_\infty^4 \sin^4\left(\frac{\varphi}{2}\right)} d\Omega. \quad (3.11)$$

Wenn N_α Teilchen eine Materie der Dicke Δx mit Kerndichte n und Kernladungszahl Z durchdringen, gibt dN_α die entsprechende Zahl der abgelenkten Teilchen an. Insbesondere ist der differentielle Streuquerschnitt mit

$$\frac{dN_\alpha}{d\Omega} \propto \sin^{-4}\left(\frac{\varphi}{2}\right) \quad (3.12)$$

stark winkelabhängig. Mit zunehmendem Winkel nimmt er sehr schnell ab.

Durch einen streifenden Einfall erhöht sich demnach die Intensität der inelastisch gestreuten Elektronen. Ein noch flacherer Winkel erhöht diese Intensität zusätzlich und wird für oberflächensensitive Messungen, wie etwa bei der Nutzung von Reflection High-Energy Electron Diffraction (RHEED), genutzt. Auch hier sind holographische Untersuchungen – zumindest in der Theorie – möglich. So lässt sich beispielsweise mithilfe von Raytracing prüfen, in welchem Rahmen RHEED-Holographie einsetzbar ist [79]. Allerdings leidet die räumliche Auflösung unter

dem flachen Winkel. Ferner wird die Bildauswertung signifikant durch zunehmende Verzerrung des Strahlflecks und des Detektorbilds erschwert [67]. RHEED wird aufgrund der hohen Oberflächensensitivität primär etwa bei der Epitaxie eingesetzt. Auf Grund der großen Verwandtschaft dieses Systems mit der in der Arbeit simulierten Apparatur lassen sich alle nachfolgend beschriebenen Erkenntnisse qualitativ dennoch weitgehend übertragen.

Unabhängig vom gewählten Einfallswinkel treffen die Elektronenwellen nach der Beugung an der Probe auf einen Schirm. Beide sind in heutigen Experimenten üblicherweise $D = 2\text{ cm}$ voneinander entfernt. Der Schirm hat eine Fläche der Größenordnung $d^2 = 5 \times 5\text{ cm}^2$, auf deren Fläche sich etwa $N^2 = 1000 \times 1000$ diskrete Pixel befinden. Inzwischen wird der Schirm nicht mehr durch eine Fotoplatte, sondern normalerweise durch eine CCD-Kamera realisiert [19]. Die Abkürzung CCD steht dabei für die verwendete Technik, es handelt sich um ein „Charge Coupled Device“. Auf dem Detektor lässt sich nun das eigentliche Muster messen.

3.2.2 Musterauswertung

Für eine vollständige Charakterisierung ist die Auswertung vieler Muster nötig. Erst die automatisierte Erkennung und Klassifizierung der Kikuchi-Bänder ermöglichte die umfangreiche Anwendung der EBSD [80]. Zunächst gelang die computergestützte Bestimmung der Linienparameter durch verhältnismäßig einfache Algorithmen, wie die zeilenweise Suche von Intensitätsmaxima [81]. Es folgte eine Kantenerkennung mithilfe eines Burns-Algorithmus, bei dem Pixel anhand einer Gradientenrichtung – nicht anhand der Intensität des Gradienten wie intuitiv anzunehmen wäre – gruppiert werden [82, 83]. Durchgesetzt haben sich aber Methoden der Analyse mithilfe einer Hough- [84, 85] und später der Radon-Transformation [86, 87].

Die Hough-Transformation wurde ursprünglich definiert, um als Verarbeitungsverfahren Linien in Bildern zu erkennen und in Geradenparameter zu überführen [88]. Sie lässt sich allerdings für verschiedene zu erkennende Formen in Bildern generalisieren [89] und wird in Bilderkennungsalgorithmen für sog. *Template Matching* angewendet [90, 91]. Die ältere Radon-Transformation geht zurück auf die Projektion einer Punkt- auf eine Geradenfunktion [92]. Auch sie lässt sich verallgemeinert für beliebige Formen anwenden. Beide Transformationen sind dabei eng miteinander verwandt [93].

Im Folgenden möchten wir die Probe genauer beleuchten, an der die eigentliche EBSD-Messung stattfindet. Denn in einem bisher als unendlich ausgedehnt und regelmäßig angenommenen Kristall existieren Schadstellen. Etwa verschiedene sog. *Kristalliten*, die sich in ihrer Ausrichtung unterscheiden und damit voneinander zu unterscheidende Gitter darstellen. Die sie trennende Grenzfläche ist eine Diskontinuität und damit ein zweidimensionaler Kristalldefekt, der *Korngrenze* genannt wird. Die Struktur der Kristalliten in ihrer Gesamtheit wird als *Gefüge* bezeichnet.

Inzwischen lässt sich EBSD bereits für viele Arten der Untersuchung dieses Gefüges eines Werkstoffs nutzen. Durch Verschieben der Probe innerhalb der Apparatur lässt sich diese Kartographieren. Korngrenzen können charakterisiert werden, indem die Kristallitenorientierungen wie oben beschrieben für einzelne Bereiche bestimmt werden. Ferner lassen sich Kristallspannungen etwa aus der Unschärfe des Musters ablesen. Weitere Untersuchungen beschäftigen sich mit der Textur, dem Verformungsverhalten und vielen weiteren Aspekten der Probe [19, 76].

3.2.3 Auflösungsvermögen

Die räumliche Auflösung eines EBSD-Experiments – also der Größe des Interaktionsvolumens – ist von vielen Parametern neben

der Wechselwirkungstiefe abhängig, wie etwa der Strahlbreite [94]. Insbesondere aber auch von der gewählten kinetischen Energie E_{kin} der einfallenden Elektronen, was nicht ausschließlich auf den IMFP zurückzuführen ist. Durch Monte-Carlo-Simulationen lässt sich eine untere Grenze einer Auflösung im Bereich weniger nm abschätzen. Mit zunehmender Energie steigt auch sie stark an [95].

Es gibt verschiedene Ansätze, diese Auflösung zu verbessern, wie etwa den der Energieselektion. Nur 5% der am Schirm ankommenden Elektronen sind tatsächlich elastisch gestreut und tragen die Strukturinformation. Die restlichen, inelastisch gestreuten Elektronen tragen nur zum Rauschen bei [96]. Ihre Energie ist allerdings geringer. Vergleichsweise junge Entwicklungen aus den letzten 15 Jahren sind daher Energiefilter, die inelastisch gestreute Elektronen vom Muster entfernen und damit den Kontrast deutlich erhöhen können [97]. Da der Energieverlust des Streuprozesses materialabhängig ist, machen sie die Messungen chemisch sensitiv [98] und sind damit auch für die hier beschriebene Holographie hochinteressant, da sie die Anwendungsmöglichkeiten erweitern.

Keller und Geiss haben eine transmittierende Art der Apparatur vorgeschlagen, die eine signifikante Auflösungsverbesserung bringt, da keine Rückstreuung mehr gemessen wird. Richtigerweise wäre sie aber der Transmission Kikuchi Diffraction (TKD) zuzuordnen [99]. Diese Methode lässt sich durch dünne Substrate deutlich verbessern [100]. Auch bei der Untersuchung von auf Substraten aufgetragenen Partikeln durch Rückstreuung profitiert die Auflösung von der Verringerung der Substratdicke [101].

3.3 EBSD-Holographie

Eine Vielzahl an Untersuchungsmethoden mittels EBSD wurden bereits erwähnt. Im Rahmen dieser Arbeit soll die holographische

Methode nun auf EBSD-Messungen übertragen werden. Verschiedene andere Arten der Elektronenholographie wurden dabei bereits für andere Beugungsmethoden von niedrigen kinetischen Energien ab $E_{\text{kin}} = 20 \text{ eV}$ [102] bis hin zu hohen Energien von $E_{\text{kin}} = 100 \text{ keV}$ [18, 103] vorgestellt und angewendet. Allerdings stellen in diesen Experimenten häufig Rauschen oder Artefakte große Probleme für eine erfolgreiche Rekonstruktion dar.

Im Falle von Niederenergie-Elektronenbeugung, der sog. Low-Energy Electron Diffraction (LEED), sind beispielsweise erfolgreiche Rekonstruktionen möglich – diese leiden allerdings unter einer stark eingeschränkten räumlichen Reichweite. Im Rahmen von Ref. [104] gelang die LEED-Holographie beispielsweise nur durch voriges Mitteln über mehrere Hologramme, die mit verschiedenen kinetischen Energien aufgenommen wurden.

Für holographische Untersuchungen mittels EBSD müssen wir zunächst analog zur Lichtholographie eine Referenzwelle definieren. Diese entspricht aber nicht der naheliegenden einfallenden Elektronenwelle. Gegenüber den austretenden Elektronen unterscheidet sie sich in der kinetischen Energie um einige eV [105–107]. Ein- und austretende Wellen sind, wie bereits beschrieben, inkohärent zueinander [73] und beeinflussen sich damit im zeitlichen Mittel nicht. Die bereits oben beschriebenen Energiefilter ließen sich hier im Übrigen auch zu einer späteren Verbesserung des Kontrasts einsetzen, indem sie die höherenergetische einfallende Elektronenwelle für das Hologramm eliminieren. Die so extern eingebrachten Elektronen können also vollständig ignoriert werden.

Stattdessen definieren wir den ersten Streuprozess an einem Atom innerhalb der Probe als Ursprung der Referenzwelle. Diese Welle streut nun an weiteren Atomen, sodass sich eine Überlagerung mehrerer gestreuter austretender Wellen und der ungestreuten Referenzwelle ergibt. Wie Abbildung 3.5 zeigt,

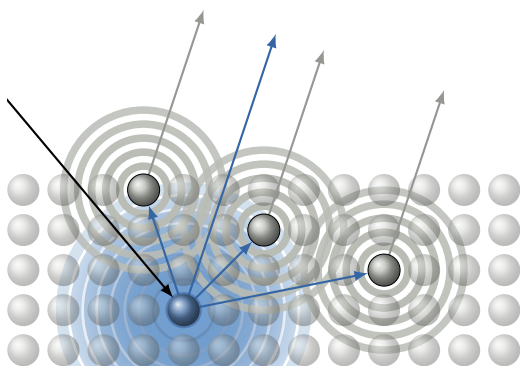


Abbildung 3.5: Die einfallende Elektronenwelle streut an einem Atom, das als Referenzwellen-Quelle dient. Von dort aus werden sekundäre Streuwellen ausgelöst, die in Summe dann das Hologramm ergeben [20, Abb. 1b].

liegt die Referenzwellen-Quelle damit effektiv innerhalb der Probe [108]. Die Abstände zwischen den einzelnen Streuern liegen dabei in der Größenordnung der Elektronenwellenlänge. Daher können wir das Intensitätsmuster aus der Überlagerung dieser Wellen – das EBSD-Beugungsmuster – als ein Hologramm, also eine Überlagerung einer Referenzwelle mit der zugehörigen Objektwelle, interpretieren [109].

Bei Beachtung der endlichen Breite des einfallenden Strahls auf den Kristall stellt sich die Frage, ob die Annahme eines einzelnen „ersten“ Streuprozesses zur Erzeugung der Referenzwelle überhaupt valide ist. Würden nämlich mehrere Referenzwellen-Quellen im Kristall zeitgleich entstehen, kann die Referenzwelle nicht mehr als Kugelwelle mit individuellem punktförmigen Zentrum angenommen werden – die Simulationen werden beliebig kompliziert. Zur qualitativen Abschätzung der Anzahl dieser Referenzstreuprozesse nehmen wir einen Strahldurchmesser von

$$d_{\emptyset} = 100 \mu\text{m} = 10^6 \text{ \AA} \quad (3.13)$$

an. Wie oben beschrieben, gibt die mittlere freie Weglänge die Tiefe des „Wechselwirkungszyllinders“ an, dessen Volumen wir für die Schätzung bestimmen wollen. Auf konkrete Energien und Kristalle soll erst später eingegangen werden, an dieser Stelle sei die mittlere freie Weglänge mit $\lambda_{\text{IMFP}} = 150 \text{ \AA}$ und eine typische Gitterkonstante mit $a = 3 \text{ \AA}$ angegeben. Wir erhalten damit also etwa

$$N_{\text{refs}} = \frac{\pi(d_{\emptyset}/2)^2 \cdot \lambda_{\text{IMFP}}}{a} \approx 4 \cdot 10^{13} \quad (3.14)$$

potentielle Referenzwellen-Quellen. Allerdings ist diese große Zahl kein Grund zur Beunruhigung, denn mehrere Gründe sprechen dafür, dass wir weiterhin an der beschriebenen Theorie festhalten können.

Erstens ist die Wechselwirkungsreichweite viel kleiner als der Strahldurchmesser, also

$$\lambda_{\text{IMFP}} \ll d_{\emptyset}. \quad (3.15)$$

Dass zwei Streuprozesse in der gleichen örtlichen Umgebung auftreten ist folglich deutlich unwahrscheinlicher.

Sollten zweitens innerhalb der Reichweite des IMFP mehrere Streuprozesse gleicher Phase auftreten, so ist davon auszugehen, dass ein ausreichend zeitlicher Abstand zwischen ihnen besteht, sodass die gegenseitige Beeinflussung absolut vernachlässigbar ist [52, Kap. 2.4.2].

Drittens besitzt ein Elektronenstrahl nur eine endliche Kohärenz, die von der verwendeten Elektronenquelle abhängig ist. Nur innerhalb eines sich auf der Probe ausbildenden Kohärenzbereichs stammende Wellenamplituden interferieren überhaupt miteinander. Andere Wellen sind unabhängig voneinander – also phasenunkorreliert. Es tritt am Schirm nur eine Addition ihrer Intensitäten auf [62]. Im zeitlichen Mittel besteht damit auch keine gegenseitige Beeinflussung der „ersten“ Streuprozesse und ihrer Folgewellen mit denen anderer Referenzwellen [20].

Gerade diese Intensitätsaddition hat noch eine weitreichende Folgerung für die Simulation von Hologrammen. Gleiche lokale Umgebungen erzeugen die selben Hologramme. Es reicht damit also völlig aus, in einem simulierten Kristall alle sich unterscheidenden lokalen Umgebungen zu identifizieren und deren Hologrammintensitäten zu addieren. Im realen Experiment ließen sich Hologramme verschiedener lokaler Umgebungen dann durch die weiter oben beschriebene chemisch sensitive Filterung wie etwa nach Refs. [97, 98] voneinander separieren und einzeln rekonstruieren. Allerdings ist daher auch keine direkte Abbildung einzelner Streuer möglich. Aufgrund der effektiven Mittlung über verschiedene lokale Umgebungen verschwinden vereinzelte Fehlstellen in der Rekonstruktion.

Die in Abschnitt 3.1.2 beschriebene kegelförmige Wellenform wird genau hier relevant: Entlang der Verbindungsachse der Referenzwellen-Quelle über das Streuatom bildet sich der Ausbreitungskegel aus. Dies geschieht wie ausgeführt jedoch nur für hohe Energien. Es sei daher erwähnt, dass durchaus auch Ansätze zur Lösung des Problems der anisotropen Elektronenwellenfunktion im Rahmen der Elektronen holographie bei niedrigeren als den hier untersuchten Energien E_{kin} existieren. Die Behandlung wird aber bedeutend komplexer und führt an dieser im Detail Stelle zu weit. Durch andere als die unten beschriebene Differenzbildmethode geartete Anpassung des Rekonstruktionsintegrals ist in Einzelfällen so etwa eine Rekonstruktion möglich [110, 111]. Ferner gelingt die Messung in einigen Fällen durch einen sehr flachen Winkel bei der Messung mittels niederenergetischer Elektronen [112]. Solche RHEED-Experimente wurden bereits in Abschnitt 3.2.1 thematisiert. Wie beschrieben, steht der gesteigerten Intensität durch flacheren Winkel eine ungenauere Bildauswertung gegenüber [67, 79].

Bei der niederenergetischen Elektronenbeugung tritt das in Abschnitt 2.3.1 beschriebene

Zwillingsbild ebenfalls auf und führt zu den für holographische Experimente bereits bekannten Problemen. Es wird ebenfalls rekonstruiert, obwohl es das aufgenommene Objekt nicht real darstellt. Im niederenergetischen Fall ließ es sich lediglich durch mathematische Tricks unterdrücken [113]. Auch im Falle der Röntgenholographie lassen sich entsprechende Algorithmen zur Verbesserung des rekonstruierten Signals entwerfen [114].

Da das rekonstruierte Objekt im Fall der hier vorgeschlagenen Elektronen holographie mittels EBSD aber zwischen Emitteratom und Schirm liegt, tritt das Zwillingsbild jenseits des Emitteratoms tiefer im Kristall auf [51]. Aufgrund der Orientierung der Wellenkegel in Richtung Schirm kann das Zwillingsbild also per Definition nicht im Rekonstruktionsvolumen erscheinen [52], was einen großen Vorteil der in Abschnitt 5.4.4 noch näher beschriebenen Rekonstruktion ins Kristallvolumen darstellt.

3.3.1 Differenzbildmethode

Ein mit heutigen Methoden aufgenommenes Muster ist in der oberen Hälfte der Abbildung 3.6 dargestellt. Dabei enthält ein Beugungsmuster insbesondere die nullte Beugungsordnung, die keine holographische, sondern nur geometrische Information in sich trägt, das Muster aber stark dominiert [108]. Dies schließt insbesondere die beschriebenen Kikuchi-Bänder ein. Es bedarf also eines Verfahrens, um diese nullte Ordnung weitestgehend zu eliminieren und damit die holographische Strukturinformation höherer Ordnung sichtbar zu machen.

Zur Motivation eines solchen Verfahrens sei zunächst an die Beugung am Einfachspalt erinnert. Dort tritt das nullte Beugungsmaximum immer an gleicher Stelle auf – völlig unabhängig von der Wellenlänge des Lichts. Mit veränderter Wellenlänge verschieben sich aber die Maxima und Minima höherer Ordnung [115, Kap. 6.4.1.6]. Analog verhält es

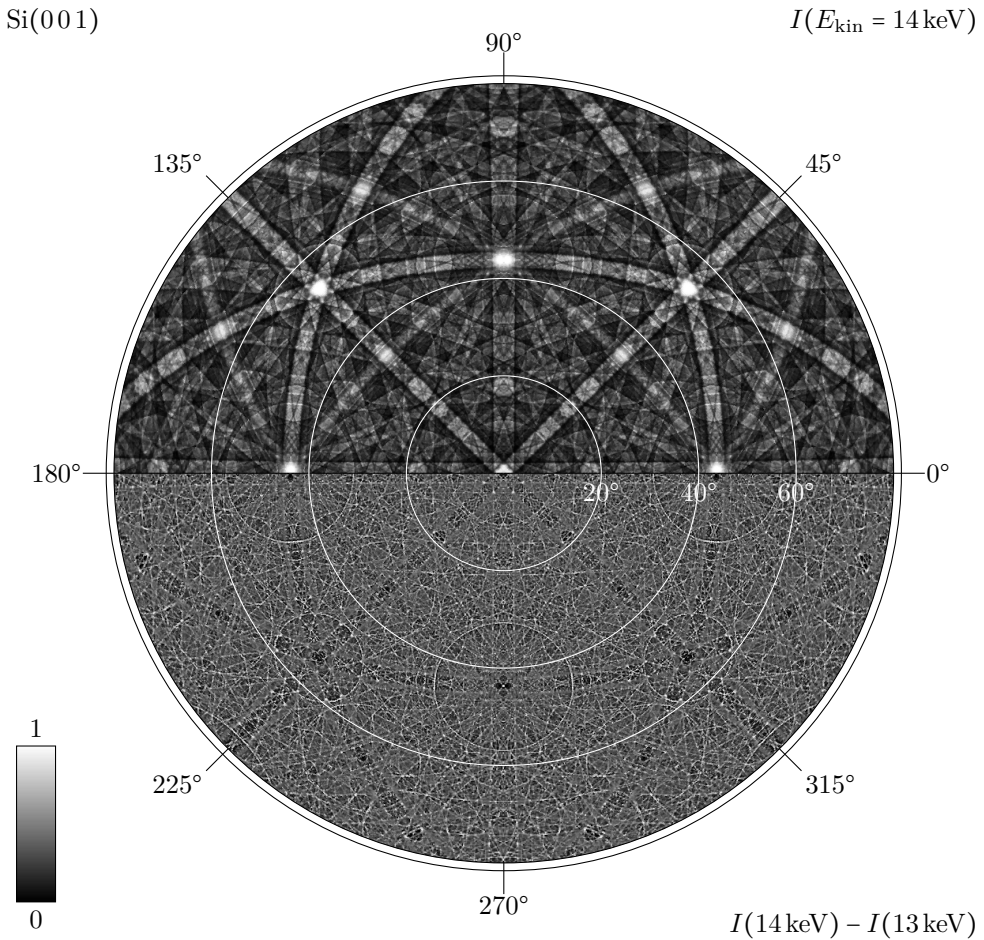


Abbildung 3.6: Aufnahme eines experimentellen EBSD-Musters eines Silizium(001)-Kristalls oben und das Differenzbild nach Abschnitt 3.3.1 zweier dieser Muster unten, die bei verschiedenen kinetischen Energien aufgenommen wurden. Vergleich nach Daten aus Ref. [52], eigene grafische Darstellung.

sich bei den hier betrachteten Beugungsmustern nach EBSD. Die geometrische Information in der nullten Beugungsordnung tritt immer an derselben Stelle auf. Und das völlig unabhängig von der Wellenlänge, also der kinetischen Energie der Elektronen, während dies auf die höheren Ordnungen nicht zutrifft.

Omori et al. machen sich diesen Effekt bei der Differential Photoelectron Holography (DPH) zunutze [116]. Statt einer einzelnen Intensität im Rekonstruktionsintegral bedienen sie sich einer partiellen Ableitung $\partial/\partial\mathbf{k}$ nach dem Wellenvektor \mathbf{k} . In einer einfacheren numerischen Ausführung bedeutet dies, dass zwei aufgenommene Muster voneinander subtrahiert werden können. Werden nämlich zwei Bilder bei unterschiedlicher kinetischer Energie aufgenommen, finden sich die geometrischen Informationen in gleicher Ausprägung und können durch die Subtraktion eliminiert werden, während die höheren Ordnungen erhalten bleiben. Abbildung 3.6 zeigt in der unteren Hälfte ein solches Muster nach Differenzbildung. Und tatsächlich ähnelt das Differenzbildmuster optisch stark den Hologrammen, die wir bereits aus der Lichtholographie kennen.

Diese einfache Ausprägung einer einzelnen Subtraktion im Rahmen der DPH, die im Folgenden *Differenzbildmethode* genannt werden

soll, ist hinreichend beschrieben. Der Schritt zur Überführung von simulierten „vollwertigen“ EBSD-Mustern durch Differenzbildung in reine Hologramme soll daher im weiteren Verlauf der Arbeit ausgelassen werden. Die vollständige Simulation dieser Prozesse erhöht hier die Komplexität stark, ohne den Erkenntnisgewinn signifikant zu steigern. Ist im Folgenden von einem EBSD-Hologramm die Rede, ist damit eigentlich die Differenz aus zwei Mustern gemeint, die zusammen wie beschrieben ein Hologramm ohne dominierende geometrische Information bilden.

Wenn im Laufe dieser Arbeit die Aufnahme mehrerer Hologramme zu einer Probe beschrieben ist, bedeutet das aber nicht automatisch, dass diese Zahl in Gedanken für das konkrete Experiment zu verdoppeln ist. Denn ein erforderliches Differenzbild lässt sich auch durch Kombination anderer Aufnahmen finden. Ab drei Mustern lassen sich so durch Permutation drei Paare finden, die dann drei Hologramme ergeben. Analog gilt dies für mehr Ausgangsmuster, sodass die Zahl benötigter Messungen ab drei Mustern unverändert bleibt. Alternativ ließe sich unterstützend zu allen Messungen ein einziges zusätzliches Subtraktionsmuster aufnehmen, das von allen anderen Mustern abgezogen wird. Die Anzahl benötigter Muster erhöht sich dann lediglich konstant um eins.

4 Computerholographie

Die mittels Elektronenholographie untersuchten Objekte haben atomare Größenordnung. Damit ist ihre Rekonstruktion dem menschlichen Auge nicht unmittelbar zugänglich. Stattdessen muss sie also berechnet und dann visualisiert werden. Dazu bietet sich die Verwendung eines Computers an.

Die digitale Berechnung von Hologrammen ist dabei bereits etabliert. Großes Interesse wird etwa der Entwicklung von 3D-Displays entgegengebracht, die durch die Beleuchtung berechneter Hologramme echt dreidimensionale Darstellungen erzeugen, ohne dabei auf etwaige Hilfsmittel wie stereoskopische Brillen zurückgreifen zu müssen [17]. Miniaturisiert lassen sich solche Displays dann beispielsweise in Augmented-Reality-Brillen einsetzen, wo sie mitten im Blickfeld virtuelle Objekte einbetten können [117].

Da die EBSD-Holographie in dieser Arbeit im Rahmen computergestützter Simulationen durchgeführt wird, ist die Behandlung der Computerholographie nötig. Ihre vornehmlich im Zusammenhang mit der simulierten Lichtholographie gewonnenen Erkenntnisse sollen dabei im Laufe der weiteren Arbeit selektiv auf die Verwendung mit EBSD angewandt werden.

Zu Beginn dieses Kapitels soll mit der Fourier-Transformation ein wichtiges Hilfsmittel in vielen Bereichen der Physik eingeführt werden. Im Speziellen ist sie auch für die hier relevante (Fourier-)Optik und damit für die folgenden Abschnitte von großer Bedeutung.

Im Anschluss werden die bei der Berechnung am Computer zwangsläufig auftretende Diskretisierung genauer behandelt. Aus der bisher weitgehend analytischen Betrachtung des holographischen Prozesses wird eine numerische, wenn diskrete Systeme behandelt werden. Aus kontinuierlichen Integralen werden

diskrete Summen. Dieser Übergang in nicht-kontinuierliche Systeme findet bereits im Experiment genau dann statt, wenn das Aufnahme-medium nicht mehr kontinuierlich ist – also beispielsweise, wie in Kapitel 3 beschrieben, durch einen CCD-Sensor realisiert ist. In Abschnitt 4.2 sollen die hierbei auftretenden Probleme und Folgen dargelegt werden.

Nebenbei treten auch *numerische Fehler*, wie etwa Rundungsfehler, aufgrund der verwendeten Datentypen oder Ungenauigkeiten in Algorithmen auf, die etwa aus Erwägungen zur Senkung der Laufzeit folgern können. Derartige Fehlertypen sollen kurz angeschnitten, aber aufgrund ihrer geringeren Signifikanz nicht weiter vertieft werden.

Abschließend sollen die Teilschritte der Holographie am Computer in Abschnitt 4.3 theoretisch erläutert werden, bevor im Folgekapitel konkrete Implementierungen behandelt werden.

4.1 Fourier-Synthese

Besonders nützlich für die Untersuchung der Effekte am Computer ist die Theorie der Signalverarbeitung. Im Einklang mit ihr werden im Folgenden Funktionen als zeitkontinuierliche *Signale* $s(t)$ angenommen. Ein Signal wird hierbei genau dann als zeitkontinuierlich beschrieben, wenn es für beliebige t in einem zusammenhängenden Intervall reeller Zahlen ohne Unterbrechung definiert ist.

Die *Fourier-Transformation* ist eines der elementarsten Werkzeuge der Signalverarbeitung. Da sie auch für viele weitere Felder der Holographie, wie etwa die Fourier-Optik [30, Kap. 1.4], interessant ist, soll sie in diesem Abschnitt anhand der genannten Signale beschrieben werden. Die Darlegung orientiert sich dabei an Ref. [46, Kap. 4.3.6] und Ref. [118, Kap. 3].

4.1.1 Fourier-Reihen

Die Fourier-Analyse lehrt, dass sich jede quadratintegrale Funktion $s(t)$ innerhalb des Intervalls $[-t_0, +t_0]$ in eine *Fourier-Reihe*

$$s(t) = s_0 + \sum_{n=1}^{\infty} \left[A_n \cos\left(\frac{n\pi}{t_0}t\right) + B_n \sin\left(\frac{n\pi}{t_0}t\right) \right] \quad (4.1)$$

entwickeln lässt [53, 119]. Quadratintegabel ist eine Funktion dabei genau dann, wenn das Integral ihres Betragsquadrats konvergiert [120, Kap. 7.2.4].

A_n und B_n stellen in Gleichung (4.1) Gewichte für die periodischen Funktionen mit Frequenz $n/2t_0$ und $n \in \mathbb{N}$ dar. Eine derartige Dekomposition ist möglich, da die Funktionen

$$\frac{1}{\sqrt{2t_0}}, \quad \frac{1}{\sqrt{t_0}} \cos\left(\frac{n\pi}{t_0}t\right), \quad \frac{1}{\sqrt{t_0}} \sin\left(\frac{n\pi}{t_0}t\right) \quad (4.2)$$

im Intervall $[-t_0, +t_0]$ ein vollständiges Orthonormalsystem bilden [46, Kap. 2.3.5].

Für die Koeffizienten dieser Fourier-Reihe (4.1) gilt

$$s_0 = \frac{1}{2t_0} \int_{-t_0}^{+t_0} s(t) dt, \quad (4.3a)$$

$$A_n = \frac{1}{t_0} \int_{-t_0}^{+t_0} s(t) \cos\left(\frac{n\pi}{t_0}t\right) dt, \quad (4.3b)$$

$$B_n = \frac{1}{t_0} \int_{-t_0}^{+t_0} s(t) \sin\left(\frac{n\pi}{t_0}t\right) dt. \quad (4.3c)$$

Die parameterunabhängige Konstante s_0 stellt dabei einen Gleichanteil der Funktion dar, deren Veränderung durch eine Superposition der periodischen trigonometrischen Sinus- und Kosinusfunktionen abgebildet werden. Anschaulich geben die *Fourier-Koeffizienten* genannten Gewichte A_n und B_n damit

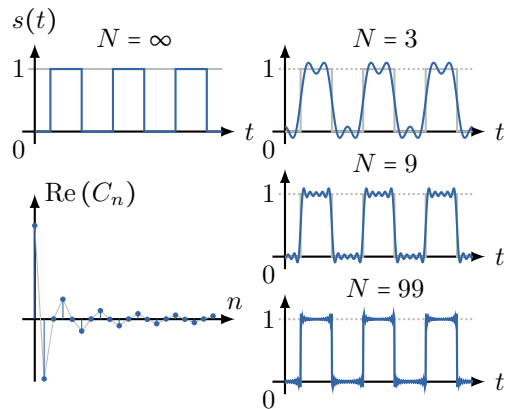


Abbildung 4.1: Die periodische Rechteckfunktion lässt sich äquivalent als Fourier-Reihe (4.8) mit unendlicher Summe darstellen. Die Fourier-Koeffizienten C_n können komplexe Zahlen sein, ihre Imaginärteile sind in diesem Falle aber 0. Es wird also lediglich ihr Realteil dargestellt. Mit zunehmendem N zur Darstellung der Reihe wird auch die Annäherung an die ursprüngliche Rechteckfunktion immer genauer. Darstellung adaptiert nach Ref. [118, Abb. 3.2 f.].

an, welche Frequenzanteile sich im entwickelten Signal befinden.

Am Beispiel einer Rechteckfunktion sind in Abbildung 4.1 Fourier-Koeffizienten dargestellt. Auch wird ebendort erkennbar, dass durch vorzeitiges Abbrechen der Fourier-Reihe bei $n = N$ eine Annäherung der Funktion möglich ist. Bei starken Funktionswertänderungen kommt es dabei zu Überschwingern, die als *Gibbs'sches Phänomen* bezeichnet werden [121]. Sie folgen aus starken Wertänderungen. Im Extremfall der Unstetigkeit, wie in der gezeigten Rechteckfunktion, wären unendlich viele Frequenzen zur Abbildung nötig – deren Aufnahme ist in der Realität aber nicht leistbar. Je nach physikalischer Anwendung stellen diese Gibbs-Schwingungen ein tatsächlich zu beachtendes Problem dar. Beispielsweise in der Simulation realer

Systeme im Bereich stark korrelierter Materie [122, Abschn. 3]. Dort lassen sich diese Probleme durch einen Faltungskern, Kernel genannt, umgehen. Neben dem oben gezeigten Fourier'schen existieren etwa auch der Fejér'sche [123, Gl. 3.38b] oder der Jackson-Kernel [124].

Zur Weiterentwicklung der Fourier-Reihe (4.1) definieren wir ein weiteres vollständiges Orthonormalsystem aus den Funktionen

$$v_n(t) := \frac{1}{\sqrt{2t_0}} \exp\left(i \frac{n\pi}{t_0} t\right), \quad n \in \mathbb{Z}. \quad (4.4)$$

Mit der *Euler'schen Formel*

$$e^{i\varphi} \equiv \cos \varphi + i \sin \varphi \quad (4.5)$$

lassen sich die trigonometrischen Funktionen ausdrücken als

$$\cos\left(\frac{n\pi}{t_0} t\right) = \sqrt{\frac{t_0}{2}} (v_n(t) + v_{-n}(t)) \quad \text{und} \quad (4.6a)$$

$$\sin\left(\frac{n\pi}{t_0} t\right) = -i \sqrt{\frac{t_0}{2}} (v_n(t) - v_{-n}(t)). \quad (4.6b)$$

Ist das Signal $s(t)$ zusätzlich noch periodisch mit einer Periodenlänge von $2t_0$, also

$$s(t + 2t_0) \equiv s(t) \quad \forall t, \quad (4.7)$$

gilt Gleichung (4.1) für alle t auch außerhalb des gegebenen Intervalls und die Reihe (4.1) lässt sich ferner zu

$$s(t) = \sum_{n=-\infty}^{\infty} C_n \exp\left(i \frac{n\pi}{t_0} t\right) \quad (4.8a)$$

mit den Amplituden

$$C_n := \frac{1}{2t_0} \int_{-t_0}^{+t_0} s(t) \exp\left(-i \frac{n\pi}{t_0} t\right) dt \quad (4.8b)$$

umschreiben. Diese als *Fourier-Synthese* bezeichnete Entwicklung einer Funktion lässt

sich aufgrund der Periodizität insbesondere über jeden zusammenhängenden Bereich der Länge $2t_0$ vornehmen. Es ist lediglich zu beachten, dass sich beide Integrationsgrenzen gleich verschieben, um eine vollständige Periode abzubilden.

4.1.2 Fourier-Transformation

Wir definieren nun die Parameter

$$d_\omega := \frac{\pi}{t_0}, \quad \omega_n := nd_\omega \quad \text{und} \quad \tilde{s}_n := C_n \cdot t_0 \sqrt{\frac{2}{\pi}}. \quad (4.9)$$

Der Parameter d_ω stellt also den Abstand zwischen benachbarten ω_n dar. Aus Gleichung (4.8) erhalten wir damit

$$s(t) = \frac{1}{\sqrt{2\pi}} \sum_{n=-\infty}^{+\infty} \tilde{s}_n e^{i\omega_n t} d_\omega, \quad (4.10a)$$

$$\tilde{s}_n = \frac{1}{\sqrt{2\pi}} \int_{-t_0}^{+t_0} s(t) e^{-i\omega_n t} dt. \quad (4.10b)$$

Heben wir nun die Beschränkung $2t_0$ -periodischer Funktionen auf, anders formuliert: wird die Periodizität beliebig vergrößert durch

$$\left[-t_0, +t_0\right]_{t_0 \rightarrow \infty}, \quad (4.11)$$

geht zeitgleich der ω_n -Abstand gegen Null, also

$$d_\omega \rightarrow 0. \quad (4.12)$$

Aus der diskreten Summe (4.10a) wird ein kontinuierliches Integral und aus den Koeffizienten (4.10b) eine Funktion, sodass

$$s(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \tilde{s}(\omega) e^{i\omega t} d\omega, \quad (4.13a)$$

$$\tilde{s}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} s(t) e^{-i\omega t} dt \quad (4.13b)$$

gilt.

Dabei bezeichnet man (4.13b) als die *Fourier-Transformierte* und notiert

$$\mathcal{F}\{s(t)\}(\omega) := \tilde{s}(\omega) \quad (4.14a)$$

und für ihre Umkehrung

$$\mathcal{F}^{-1}\{\tilde{s}(\omega)\}(t) := s(t) \quad (4.14b)$$

oder in einer symbolischen Notation nach Ref. [123, Kap. 2.1]

$$s(t) \circ \text{---} \tilde{s}(\omega) \quad \text{und} \quad (4.15a)$$

$$\tilde{s}(\omega) \text{---} \circ s(t). \quad (4.15b)$$

Der konkrete Vorfaktor der Fourier-Transformation (4.13) ist dabei abhängig von der gewählten Konvention und nicht weiter relevant, wenn für Hin- und Rücktransformation jeweils die richtige zugehörige äquivalente Darstellung gewählt wird [125, Kap. 15.3.1.1]. Im Folgenden treten äquivalente Darstellungen konsistent paarweise auf.

Die bisher als Parameter eingeführte Variable ω ist dabei auch physikalisch deutbar. Es handelt sich um eine Frequenz, bei der Fourier-Transformierten $\tilde{s}(\omega)$ damit um das Frequenzspektrum des Signals $s(t)$ [125, Kap. 15.3.1.3].

Besonders nützlich ist die Fourier-Transformation bei der Berechnung von Faltungen

$$s_1(t) * s_2(t) := \int_{-\infty}^{+\infty} s_1(\tau) s_2(t - \tau) d\tau, \quad (4.16)$$

da sich durch Transformation das *Faltungstheorem*

$$\mathcal{F}\{s_1(t) * s_2(t)\}(\omega) = \tilde{s}_1(\omega) \cdot \tilde{s}_2(\omega) \quad (4.17)$$

ergibt [125, Kap. 15.3.1.3]. Es gilt analog auch – je nach Konvention mit einem variierenden Vorfaktor und hier im Besonderen aufgrund der Art der Notation in Gleichung (4.13) auch ohne Vorfaktor – umgekehrt [126, Kap. 2.3]

$$\mathcal{F}\{s_1(t) \cdot s_2(t)\}(\omega) = \tilde{s}_1(\omega) * \tilde{s}_2(\omega). \quad (4.18)$$

Die Fourier-Transformation hat viele weitere Eigenschaften, die bei Bedarf lediglich referenziert werden. Es sei für eine vollständigere Übersicht stattdessen auf Ref. [125, Kap. 15.3.1] verwiesen.

Da die aufgezeichneten Hologramme zweidimensional sind, soll die Transformation an dieser Stelle explizit auf mehrere Dimensionen erweitert werden. Es seien dazu die d -dimensionalen komplexwertigen Vektoren

$$\mathbf{t}, \boldsymbol{\omega} \in \mathbb{C}^d \quad (4.19)$$

definiert, die jeweils die Funktionsparameter in \mathbf{t} und Frequenzkomponenten in $\boldsymbol{\omega}$ enthalten. Für die Transformation gilt dann verallgemeinert nach [127, Kap. 2.1]

$$s(\mathbf{t}) = \left(\frac{1}{2\pi}\right)^{d/2} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \tilde{s}(\boldsymbol{\omega}) e^{i\boldsymbol{\omega}\mathbf{t}} d^d\boldsymbol{\omega}, \quad (4.20a)$$

$$\tilde{s}(\boldsymbol{\omega}) = \left(\frac{1}{2\pi}\right)^{d/2} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} s(\mathbf{t}) e^{-i\boldsymbol{\omega}\mathbf{t}} d^d\mathbf{t}. \quad (4.20b)$$

4.2 Diskretisierung

Werden Wellen am Computer simuliert, wird aus der kontinuierlichen Wellenfunktion völlig unabhängig von der konkret gewählten Wellenart eine Menge diskreter Messwerte. Mit der Vorarbeit zur Fourier-Transformation können wir nun die daraus folgenden Konsequenzen im Rahmen der Signalverarbeitung behandeln.

Im Zusammenhang mit der Darstellung und Klassifizierung solcher Signale am Computer gibt es hier zwei Effekte zu unterscheiden: *Abtastung* und *Quantisierung*. Deren Erläuterung folgt im Wesentlichen Ref. [118, Kap. 4].

Im Falle der Abtastung wird das kontinuierliche Signal zu verschiedenen Zeitpunkten gemessen. Diese Abtastung findet bezüglich aller Funktionsparameter statt, soll hier aber exemplarisch anhand der Zeit erfolgen. Alles

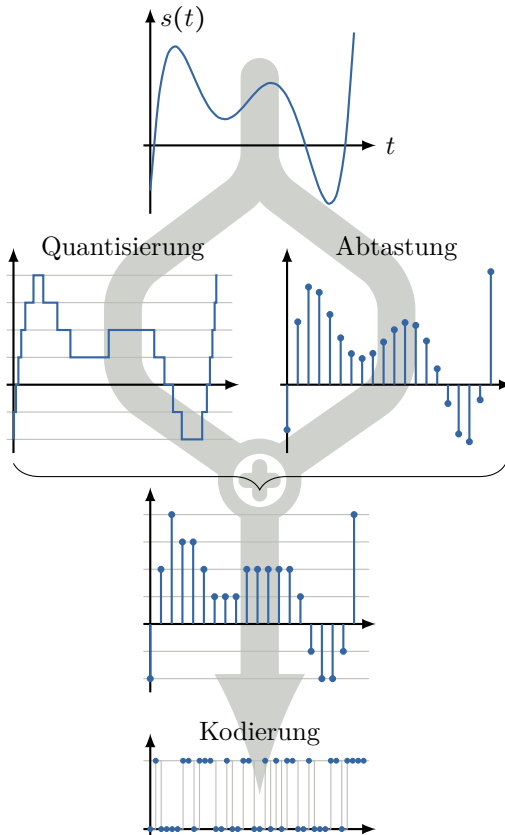


Abbildung 4.2: Bei der Aufnahme am Computer wird ein Signal sowohl quantisiert als auch abgetastet. Beide Prozesse sind zunächst getrennt dargestellt. Bei der Quantisierung wird das Signal wertdiskret und auf vorgegebene Quantisierungsstufen gerundet. Bei der Abtastung wird das Signal durch ein Messen zu bestimmten Zeitpunkten zeitdiskret. Beide Prozesse vereint ergeben das digitale Signal. Durch eine abschließende Kodierung wird es binär und kann durch den Computer verarbeitet und gespeichert werden. Darstellung adaptiert nach Ref. [118, Abb. 4.1].

Beschriebene lässt sich analog auf die später noch relevanten Ortsparameter übertragen. Damit ergibt sich ein zeitdiskretes Signal aus einer beliebigen Funktion $s(t)$ mittels

$$s_a(t) = s(t) \cdot \sum_n \delta(t - t_n). \quad (4.21)$$

Bei $\delta(x)$ handelt es sich um die Kronecker-Delta-Distribution, die gemäß

$$\delta(x) = \begin{cases} 1 & x = 0, \\ 0 & \text{sonst} \end{cases} \quad (4.22)$$

definiert ist [128, Kap. 5]. Die Aufnahme erfolgt in Gleichung (4.21) zu beliebigen Zeitpunkten t_n . Typischerweise sind diese äquidistant gewählt, sodass

$$s_a(t) = s(t) \cdot \sum_n \delta(t - n\Delta t) \quad (4.23)$$

mit konstantem Δt gilt. Wie in Abbildung 4.2 zu erkennen, ist das Signal damit nur zu bestimmten Zeiten überhaupt bekannt, weshalb das Signal nun auch *zeitdiskret* genannt wird.

In der Realität sind solche Abtastimpulse selbstverständlich nicht beliebig kurz, eine exaktere Abbildung dieses Umstands wäre mit einer Rechteckfunktion möglich. Darauf soll an dieser Stelle zu Gunsten der Übersichtlichkeit aber verzichtet werden.

Technisch bedingt ist es ferner nicht möglich, jede beliebige Zahl im Computer darzustellen. Daher passiert automatisch auch eine Quantisierung des Signalwerts. Häufig werden Messwerte als *Gleitkommazahlen* aufgenommen. Deren genaue Beschaffenheit ist stark abhängig von der konkreten Implementierung und hier daher nicht weiter relevant. Allerdings sei festgestellt: zwar sind diese Gleitkommazahlen in der Lage, sowohl sehr kleine Zahlen wie das Planck'sche Wirkungsquantum mit einem Zahlenwert von $h \sim 10^{-34}$, als auch sehr große Zahlen wie die Avogadro-Konstante mit einem Zahlenwert von $N_A \sim 10^{23}$ recht gut abzubilden. Dennoch gibt es innerhalb der Größenordnungen unterschiedliche Schrittweiten, in denen Zahlen abgebildet werden können [129, Kap. 4.2]. Beim

Aufnehmen des Signals muss ein Wert also zwangsläufig auf eine dieser Quantisierungsstufen mit einem nicht unbedingt uniformen Abstand Δs gerundet werden – was Messfehler einführt. Ein entsprechend quantisiertes Signal lautet

$$s_q(t) = \text{sgn}(t) \cdot \Delta s \cdot \left\lfloor \frac{|t|}{\Delta s} + \frac{1}{2} \right\rfloor. \quad (4.24)$$

Dabei bezeichnet $\text{sgn}(x)$ die Signumfunktion, die das Vorzeichen ihres Arguments angibt [128, Kap. 4]. Ferner findet die auf ganze Zahlen abrundende Gaußklammer $\lfloor x \rfloor$ Verwendung. Das Signal ist nach dieser Art der Quantisierung nun auch *wertdiskret*. Die dadurch eingeführten Fehler sollen an dieser Stelle übergangen werden, da sie hier zu weit führen und gegenüber den Effekten der Abtastung weit weniger stark wahrnehmbar sind.

Das wert- und zeitdiskrete Signal lässt sich nun – abhängig vom konkreten Datentyp – kodieren. Das bedeutet, dass eine Abbildung auf ein *zweiwertiges* Signal vorgenommen wird: Das am Computer verwendete Signal besteht lediglich aus den Binärzahlen 0 und 1. Effekte dieser Kodierung wie das *Überlaufen* stellen in der Implementierung durchaus Probleme dar, sollen an dieser Stelle aber ebenfalls ausgeblendet werden. Sie sind stark von der konkreten Plattform und Programmierart abhängig und können weitestgehend über die Wahl eines passenden Datentyps vermieden werden. Es sei für eine ausführliche Behandlung dieser Fehlerquellen auf die Einführung in die Numerik in Algorithmen gemäß Ref. [129] verwiesen.

4.2.1 Abtasttheorem

Bei Betrachtung eines Videofilms, der ein fahrendes Auto zeigt, fällt auf, dass die Räder sich ab gewissen Rotationsgeschwindigkeiten scheinbar rückwärts drehen. Dieser Effekt kommt durch eine zu langsame Aufnahme der Einzelbilder des Videos zustande: Bevor ein

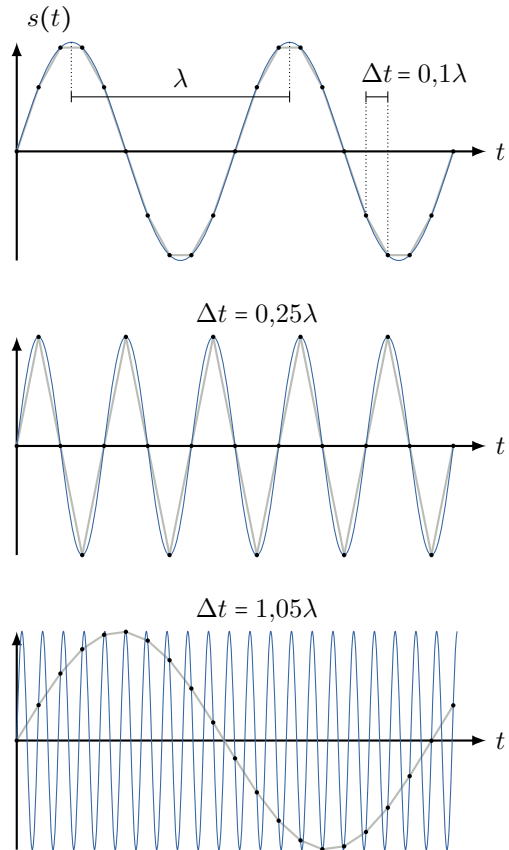


Abbildung 4.3: Bei festem Abtastabstand Δt kann mit zunehmender Frequenz und damit abnehmender Wellenlänge λ das Signal nicht mehr rekonstruiert werden. Liegen nur Informationen über die diskreten mit Punkten markierten Messwerte vor, könnte ein deutlich langwelligeres Signal interpoliert werden. Eigene Darstellung nach Ref. [130, Abb. 3.3] und Ref. [11, Abb. 3.4].

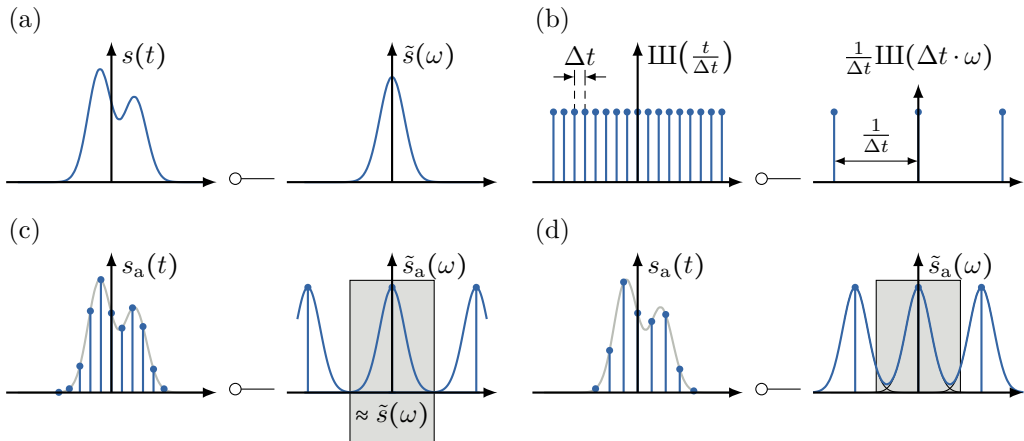


Abbildung 4.4: Darstellung eines (abgetasteten) Signals und seiner Fourier-Transformierten in Anlehnung an Ref. [11, Abb. 4.2 f.]. (a) Das Fourier-Spektrum des gezeigten Signals lässt sich kontinuierlich transformieren. (b) Die Fourier-Transformierte des Delta-Kamms ist ebenfalls ein Delta-Kamm mit invertiertem Abstand. Bei engerer Abtastung erhöht sich entsprechend der Abstand der Abtastung im Fourier-Raum, respektive umgekehrt. (c) Bei der „richtigen“ Abtastung ist es möglich, die Fourier-Transformierte mithilfe der Rechteckfunktion zu isolieren. Das ursprüngliche Signal lässt sich zweifelsfrei durch Rücktransformation rekonstruieren. (d) Bei der Unterabtastung hingegen schieben sich die Fourier-Transformierten ineinander: Die zweifelsfreie Rekonstruktion ist nicht möglich, da sich die Frequenzanteile überlappen.

neues Foto des Rads geschossen wurde, hat es bereits fast eine volle Umdrehung vollführt. Bei schneller Bildfolge dreht es sich scheinbar rückwärts. Dieser Effekt wird *Unterabtastung* genannt und soll im Folgenden quantifiziert werden, da er in ähnlicher Form auch bei der Berechnung von Hologrammen auftritt.

Die in Gleichung (4.23) eingeführte Abtastfrequenz

$$\Delta\omega := 1/\Delta t \quad (4.25)$$

kann keineswegs willkürlich gewählt werden. Bei der in Abbildung 4.3 gezeigten Aufzeichnung verschiedener Samplingintervalle fällt nämlich auf: Es lassen sich bei zu geringer Abtastrate lediglich langwelligere Funktionen durch die Punkte interpolieren. Eine untere Grenze für diese Abtastrate gibt das *Abtasttheorem* an. Bei dessen Herleitung orientieren wir uns an Ref. [11, Kap. 4.5].

Wir transformieren dazu zunächst die abgetastete Funktion (4.23) mittels des Faltungstheorems (4.17) zu

$$\begin{aligned} \tilde{s}_a(\omega) &\stackrel{(4.23)}{=} \mathcal{F}\left\{s(t) \cdot \sum_n \delta(t - n\Delta t)\right\}(\omega) \\ &\stackrel{(4.18)}{=} \tilde{s}(\omega) * \mathcal{F}\left\{\sum_n \delta(t - n\Delta t)\right\}(\omega). \end{aligned} \quad (4.26)$$

Die dabei auftretende Summe der Dirac-Distributionen lässt sich in die Einheits-Dirac-Impulsfolge

$$\text{III}(t) := \sum_{n=-\infty}^{\infty} \delta(t - n) \quad (4.27)$$

nach Ref. [128, Kap. 10] zusammenfassen. Ihre Fourier-Transformierte lautet

$$\mathcal{F}\{\text{III}(t)\}(\omega) = \sum_{m=-\infty}^{\infty} \delta(\omega - m) = \text{III}(\omega). \quad (4.28)$$

Damit sie wertgleich in Gleichung (4.26) eingesetzt werden kann, ist zunächst eine Dehnung durch Skalierung des Funktionsparameters erforderlich. Mit dem Ähnlichkeitstheorem gemäß [125, Kap. 15.3.1.3]

$$\mathcal{F}\left\{s\left(\frac{t}{a}\right)\right\}(\omega) = |a| \cdot \tilde{s}(a\omega) \quad (4.29)$$

lässt sich daraus auch die Fourier-Transformierte der gedehnten Dirac-Impulsfolge zu

$$\begin{aligned} \mathcal{F}\left\{\sum_{n=-\infty}^{\infty} \delta(t - n\Delta t)\right\}(\omega) \\ = \frac{1}{|\Delta t|} \sum_{m=-\infty}^{\infty} \delta\left(\omega - \frac{m}{\Delta t}\right). \end{aligned} \quad (4.30)$$

berechnen [118, Kap. 3.8.2].

Wir können Gleichung (4.26) für

$$\Delta t > 0 \quad (4.31)$$

damit weiterentwickeln zu

$$\begin{aligned} \tilde{s}_a(\omega) &\stackrel{(4.30)}{=} \tilde{s}(\omega) * \frac{1}{\Delta t} \sum_{m=-\infty}^{\infty} \delta\left(\omega - \frac{m}{\Delta t}\right) \\ &\stackrel{(4.25)}{=} \Delta\omega \sum_{m=-\infty}^{\infty} \tilde{s}(\omega - m\Delta\omega), \end{aligned} \quad (4.32)$$

wenn wir im letzten Schritt die Faltung gemäß Gleichung (4.16) durchführen. Das Fourier-Spektrum eines Signals, das mit einer Periodizität von Δt abgetastet wurde, hat folglich eine Periodizität von $\Delta\omega$. Anschaulich ist diese Beziehung in Abbildung 4.4a-c dargestellt.

Sofern nun ein Frequenzbestandteil größer ist als $\Delta\omega/2$, überlagert er, wie in Abbildung 4.4d abgebildet, im Frequenzspektrum mit der benachbarten Periode. Damit ist keine Rekonstruktion dieses Frequenzanteils mehr möglich. Es muss also der Zusammenhang

$$|\omega| < \frac{\Delta\omega}{2} = \frac{1}{2\Delta t} \quad (4.33)$$

gelten.

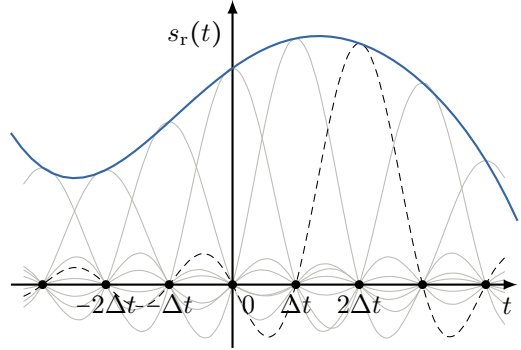


Abbildung 4.5: Nach Ref. [118, Abb. 4.10] adaptierte Darstellung des Signals als Summe der approximierenden sinc-Funktionen. Zur besseren Visualisierung ist eine dieser Funktion hervorgehoben.

Zur Rekonstruktion des Signals wird zunächst mithilfe der Rechteckfunktion $\Pi(t)$ nach Ref. [128, Kap. 4]

$$\Pi(t) = \begin{cases} 0 & |t| > 1/2 \\ 1/2 & |t| = 1/2 \\ 1 & |t| < 1/2 \end{cases} \quad (4.34)$$

der Kernteil des periodischen Spektrums aufgelöst. Dafür wird die Rechteckfunktion auf die Breite einer Frequenzspektrumsperiode $\Delta\omega$ angepasst und wir erhalten

$$\tilde{s}_r(\omega) = \Pi\left(\frac{\omega}{\Delta\omega}\right) \cdot \tilde{s}_a(\omega). \quad (4.35)$$

Mit der Spaltfunktion sinc als Fourier-Transformierte der Rechteckfunktion $\Pi(t)$ [128, Kap. 6] und dem Faltungstheorem ergibt sich

$$\begin{aligned} s_r(t) &= \frac{1}{\Delta t} \operatorname{sinc}\left(\frac{t}{\Delta t}\right) * s_a(t) \\ &= \frac{1}{\Delta t} \sum_{n=-\infty}^{\infty} s(n\Delta t) \operatorname{sinc}\left(\frac{t - n\Delta t}{\Delta t}\right). \end{aligned} \quad (4.36)$$

Abbildung 4.5 zeigt, wie die Annäherung des Signals durch die Summe verschiedener Spaltfunktionen möglich ist.

Gleichung (4.33) zeigt also: Die größte messbare Frequenz ist begrenzt. Sie wird *Nyquist-Frequenz* ω_{NY} genannt [131, Kap. 5.1.1]. Damit die im Signal maximal vorkommende Frequenz unterhalb der Nyquist-Frequenz liegt, muss die Abtastfrequenz $\Delta\omega$ nach Gleichung (4.33) also mindestens doppelt so groß sein. Diesen Zusammenhang nennt man *Nyquist-Shannon-Abtasttheorem* [23, 24, 132]

$$\Delta t^{-1} = \Delta\omega > 2\omega_{\text{max}}. \quad (4.37)$$

Nach Whittaker [133], Kotelnikov [134] und Shannon wird es auch *WKS-Abtasttheorem* genannt, gelegentlich aber auch *Someya* zugeschrieben [135, 136].

4.2.2 Aliasing

Wird entgegen der Forderung des Abtasttheorems (4.37) mit einer Nyquist-Frequenz unterhalb der Maximalfrequenz abgetastet, so kommt es zu den oben beschriebenen Fehlinterpretationen des Signals. Diese werden *Aliasing* genannt.

Der bereits in Abbildung 4.3 angedeutete Effekt lässt sich quantifizieren. Denn die falsch wahrgenommene Frequenz tritt als Spiegelung an der Abtastfrequenz auf. Anhand eines Beispiels aus der Tontechnik lässt sich dies leicht nachvollziehen: Ein Tonsignal mit einer eigentlich für das menschliche Gehör nicht wahrnehmbaren Komponente bei $\omega_{\text{max}} = 30 \text{ kHz}$ wird mit einer Abtastfrequenz von $\Delta\omega = 48 \text{ kHz}$ aufgenommen. Die Nyquist-Frequenz beträgt folglich $\omega_{\text{NY}} = 24 \text{ kHz}$ und liegt damit unterhalb der Störkomponentenfrequenz ω_{max} . Diese spiegelt an der Abtastfrequenz und ergibt eine Störkomponente bei $48 \text{ kHz} - 30 \text{ kHz} = 18 \text{ kHz}$, die plötzlich sehr wohl hörbar ist. In der Praxis lässt sich dieses Problem allerdings nicht unbegrenzt durch Erhöhung der Abtastfrequenz lösen. Stattdessen werden *Anti-Aliasing-Filter* eingesetzt, die den Effekt unterdrücken sollen. In diesem konkreten Beispiel ließe sich etwa ein Tiefpass

einsetzen, der nur Frequenzen unterhalb der Nyquist-Frequenz ω_{NY} passieren lässt, während er andere, höhere Frequenzen ausblendet. Damit werden die Frequenzen, die später als Störanteile im abgetasteten Signal auftauchen könnten, bereits vor der Abtastung abgeblockt [131, Kap. 5.1.2].

In der Realität existieren keine Tiefpässe mit unendlich steilen Flanken, weshalb eine Überabtastung nötig ist. In der Tontechnik gilt beispielsweise als im hörbaren Basisband enthalten jede Frequenz, die unterhalb von $\omega_{\text{max}} = 20 \text{ kHz}$ liegt. Dennoch wird etwa im Rahmen des Audio-Compact-Disc-Standards eine Abtastfrequenz von $\Delta\omega = 44,1 \text{ kHz}$ gewählt. Einerseits erlaubt die daraus folgende Nyquist-Frequenz von $\omega_{\text{NY}} = 22,05 \text{ kHz}$ eine gewisse Steigung in der Flanke des Tiefpasses, andererseits war sie mit damaligen Videostandards kompatibel. Nebenbei ergibt sich diese Abtastfrequenz aus der Quadratur der Multiplikation der vier kleinsten Primzahlen nach

$$\Delta\omega = (2 \cdot 3 \cdot 5 \cdot 7)^2 \text{ Hz} = 44\,100 \text{ Hz}, \quad (4.38)$$

was erlaubt, viele ganzzahlige Produkte abzubilden.

Eine zu starke Überabtastung gilt es aber zu vermeiden, da jeder Messwert mehr Speicherplatz benötigt und gesteigerten Rechenaufwand mit sich bringt. Auch für die Holographie ist ein entsprechend guter Kompromiss zu wählen.

Besonders deutlich wird die negative Auswirkung des Abtasttheorem beim Holographieren durch die Betrachtung der dreidimensionalen ebenen Welle in Abbildung 4.6. Der die Ausbreitungsrichtung angegebene Wellenvektor \mathbf{k} befindet sich in einem Winkel ϑ_y zu einer Samplefläche, auf der äquidistant Pixel angeordnet sind. Die Schirmkoordinaten lassen sich als

$$\mathbf{r} = \begin{pmatrix} x_n \\ y_m \\ z_0 \end{pmatrix} = \begin{pmatrix} n\Delta x \\ m\Delta y \\ z_0 \end{pmatrix} \quad (4.39)$$

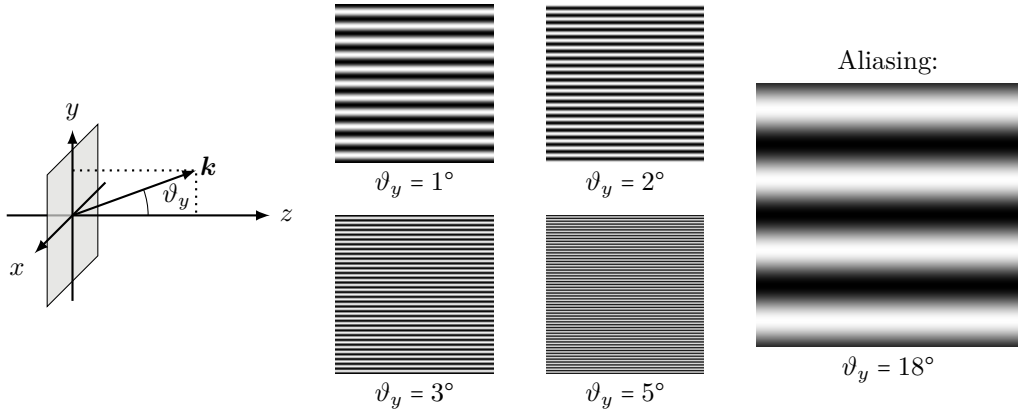


Abbildung 4.6: Beispiele einer Abtastung von ebenen Wellen in eigener Darstellung nach Ref. [11, Abb. 3.9]. Mit zunehmendem Winkel ϑ_y der Ausbreitungsrichtung nimmt die projizierte Wellenlänge immer weiter ab. Durch Alias-Effekte kommt es aber plötzlich zu einer scheinbaren Zunahme der Wellenlänge durch eine Unterabtastung. In den gezeigten Bildern wurde die Wellenlänge des roten HeNe-Lasers $\lambda = 633 \text{ nm}$ gewählt. 200 Pixel pro Dimension haben einen Abstand von jeweils $\Delta x = \Delta y = 1 \mu\text{m}$ zueinander.

beschreiben und zur Bestimmung der in Abbildung 4.6 gezeigten Muster in die ebene Welle (2.14) einsetzen. Es wird ersichtlich, dass die Auflösung in y -Richtung ab einem bestimmten Winkel nicht mehr ausreicht: Die eigentlich kontinuierlich zunehmende Frequenz wird aufgrund eines Aliasing-Fehlers plötzlich als abnehmend wahrgenommen.

Um diesen Effekt einzuschränken, kann ein maximaler Winkel $\vartheta_{y,\text{max}}$ bestimmt werden, unter dem eine fehlerfreie Abtastung bei gegebenem Messpunktabstand noch möglich ist. Die auf die y -Achse projizierte Welle in Abhängigkeit des Winkels ϑ_y hat die Wellenlänge

$$\lambda_y = \frac{\lambda}{\sin \vartheta_y}. \quad (4.40)$$

Damit muss der Abtastabstand nach Abtasttheorem (4.37)

$$\Delta y < \frac{\lambda}{2 \sin \vartheta_y} \quad (4.41)$$

sein. Da die Auflösung Δy zuvor in der Regel bereits feststeht, ergibt sich praktisch

eher umgekehrt ein maximal möglicher Abbildungswinkel von

$$\vartheta_{y,\text{max}} = \arcsin\left(\frac{\lambda}{2\Delta y}\right), \quad (4.42)$$

der für die x -Richtung völlig analog gilt [11, Kap. 3.2.5].

Bei zusätzlich frei wählbarem $z_0(z)$ wird der Zusammenhang deutlich komplexer, es gilt aber auch weiterhin, dass innerhalb der Grenzen (4.42) in jedem Falle eine korrekte Aufnahme des Signals möglich ist [11, Kap. 3.2.6].

Kugelwellen können lokal als ebene Wellen angenähert werden, sodass die hier angegebenen Maximalwinkel grundsätzlich in erster Näherung auch für sie angenommen werden können [11, Kap. 3.3.3.1]. Auf einem holographischen Schirm lässt sich damit ein Bereich gesichert frei von Aliasing feststellen, der sich im Idealfall über den gesamten Schirm erstreckt.

Tatsächlich lässt sich jede beliebige Lösung der Wellengleichung in ebene Wellen nach

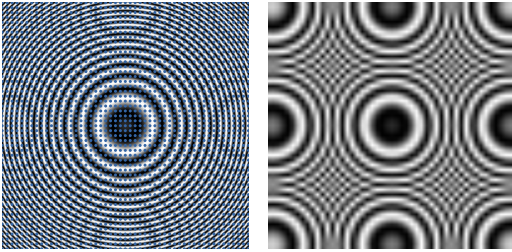


Abbildung 4.7: Eine Unterabtastung erzeugt eine völlig falsche Vorstellung vom Hologramm. Das eigentlich von einem Streuer erzeugte Muster (es entspricht genau einer Fresnel-Zonenplatte) wird links in einer Auflösung von 1001×1001 Pixeln dargestellt. Das eingezeichnete Punktraster deutet als optische Täuschung bereits an, welchen Effekt die Unterabtastung hat. Tatsächlich handelt es sich um das reine Kreisringssystem ähnlich zu Abbildung 2.5b. Durch Abtastung mit insgesamt nur 51×51 Pixeln im aufgelegten Raster ergibt sich ein rechts dargestelltes, periodisches Muster, das von mehreren Streuern stammen könnte. Diese Unterabtastung kann bereits bei der Darstellung am Computer erfolgen, da dieser Bilder quantisiert in Pixeln wiedergibt.

Gleichung (2.14) entwickeln. Das macht den in Abschnitt 2.2.1 als rein mathematische Lösung eingeführten Wellentyp noch nützlicher in der Behandlung und Annäherung vieler weiterer Wellenphänomene [46, Kap. 4.3.7].

Abschließend sei bemerkt, dass Aliasing-Fehler etwa beim Drucken oder der Darstellung am Bildschirm auftreten können – auch hier liegen diskrete Datenpunkte zugrunde, die eine Grafik abbilden. Daher können auch Interpretationsfehler bei der Betrachtung eines in der vorliegenden Arbeit dargestellten Hologramms auftreten. Besonders deutlich wird das bei der Betrachtung von Abbildung 4.7: Durch zu geringe Abtastung wirkt das Hologramm plötzlich grundverschieden.

4.2.3 Diskrete Transformation

Ein Signal wird realistisch nur mit einer bestimmten Anzahl N an Messpunkten abgetastet, die ferner auch nur in eine begrenzte Anzahl M von Frequenzwerten überführt werden können. Wir nehmen analog zu Ref. [11, Kap. 4.6] die Abtastabstände Δt und $\Delta \omega$ als konstant an, sodass die diskrete Fourier-Transformation (kurz DFT nach dem englischen „Discrete Fourier Transformation“) in Anlehnung an Gleichung (4.13b) nun

$$\tilde{s}(\omega) = \sum_{n=0}^{N-1} s(n\Delta t) e^{-2\pi i n \Delta t m \Delta \omega} \quad (4.43)$$

lautet. Da das Signal im abgetasteten Intervall wie das Spektrum als periodisch angenommen wird, ist $M = N$ zur eindeutigen Transformation ausreichend [118, Kap. 4.3.6]. Werden die Mess- und Frequenzwerte in Vektoren \mathbf{s} , $\tilde{\mathbf{s}}$ notiert, lässt sich die Summe (4.43) mithilfe einer Matrix

$$\tilde{\mathbf{s}}(\omega) = F \cdot \mathbf{s}(t), \quad F_{nm} = e^{-2\pi i n \Delta t m \Delta \omega} \quad (4.44)$$

umformulieren.

Zur Bestimmung dieser Transformationsmatrix F wäre also die Berechnung von N^2 Elementen nötig. Dieser Rechenaufwand wird in der Landau-Notation [137–139] mit $\mathcal{O}(N^2)$ notiert. Damit skaliert das Problem sehr ungünstig und bindet viele Ressourcen.

Die Fast Fourier Transformation (FFT) ist ein Algorithmus, der durch Zerlegung der einzelnen Berechnungen der Matrixelemente die Komplexität auf $\mathcal{O}(N \log N)$ senkt. Diese FFT bildet die Grundlage jeder heutigen Fourier-Transformation am Computer. Der Algorithmus soll aus Gründen der Übersichtlichkeit hier nicht näher behandelt werden, stattdessen sei auf Ref. [140, Kap. 10f.] verwiesen. Weitere Spezialfälle der diskreten Fourier-Transformation, insbesondere aber die reelle Fourier-Transformation, sollen hier aus denselben Gründen ebenfalls nicht betrachtet werden, wenngleich sie in der konkreten Implementierung holographischer Berechnungen Anwendung finden können.

4.3 Holographie

An dieser Stelle knüpfen wir an die Erkenntnisse aus der Lichtholographie in Kapitel 2 an und übernehmen sie implizit zusammen mit allem, was für die Elektronenholographie aus Abschnitt 3.3 notwendig ist, für die Simulation am Computer. Dabei wird begrifflich zwischen mehreren Komponenten der Holographie unterschieden. Wird ein erstelltes Hologramm mithilfe eines Sensors digital aufgezeichnet, so ist von Digital Holography (DH) die Rede. Das Hologramm lässt sich aber selbst bereits als Computer Generated Hologram (CGH) am Computer virtuell erzeugen. Dieses „virtuelle“ Hologramm lässt sich dann sogar „real“ mithilfe eines Folien-druckers ausdrucken und durch Beleuchtung optisch rekonstruieren. Selbstverständlich ist aber auch die Rekonstruktion am Computer möglich. Alle angesprochenen Teilkomponenten bilden zusammen die Computer Holography (CH) [11, Kap. 2.2].

In Abbildung 4.8 sind die einzelnen Bestandteile der CH als Flussdiagramm dargestellt. Da in dieser Arbeit sowohl die Holographie als auch die Rekonstruktion in einer Simulation geschehen, muss der gesamte Prozess vor diesem Hintergrund untersucht werden. Dazu sollen in diesem Abschnitt alle relevanten Teilschritte erläutert werden, bevor im nächsten Kapitel 5 die praktische Implementierung skizziert wird.

4.3.1 Objekte

Erste Herausforderung der CGH ist bereits die virtuelle Beschreibung eines zu holographierenden Objekts. Die Entwicklung immer detaillierterer Computerspielegrafik der letzten Jahre zeigt, dass die Darstellung komplizierter dreidimensionaler Objekte möglich ist. Jedoch gibt es einen entscheidenden Unterschied zwischen dieser Computer Graphics (CG) und der Computer Holography, der anhand von Abbildung 4.9 erklärbar ist.

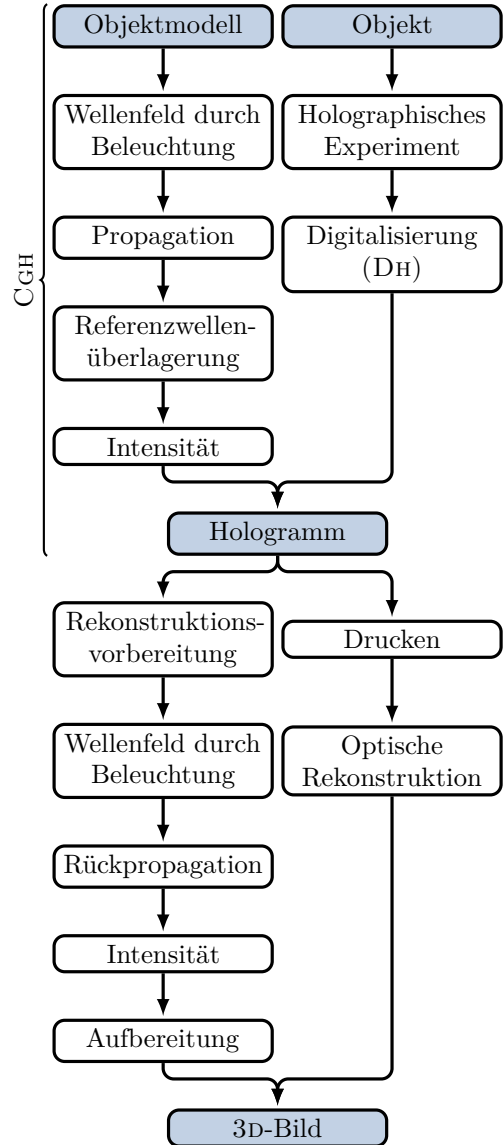


Abbildung 4.8: Flussdiagramm zum Ablauf des vollständigen holographischen Vorgangs am Computer links und in optischen Aufbauten rechts sowie deren Interaktionsmöglichkeit.

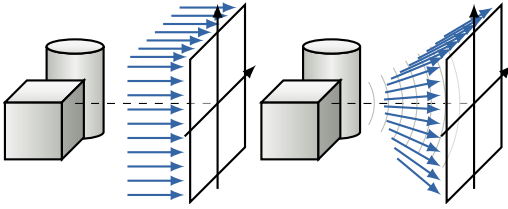


Abbildung 4.9: Schematische Unterscheidung nach Ref. [11, Abb. 2.6] des Renderings dreidimensionaler Objekte im Falle der Computer-Grafik durch Projektion (links) und der Computer-Holographie durch Wellenpropagation (rechts).

Im Rahmen von CG wird typischerweise sog. *Raytracing* verwendet, um zu entscheiden, was der Anwender „sieht“. Ziel ist, ein dreidimensionales Objekt auf eine zweidimensionale Ebene abzubilden. Der Betrachter meint eine Fotografie vor sich zu haben. In einer einfachen Implementierung wird dabei von jedem Beobachtungspunkt ein Strahlengang (*Ray*) durch die Bildebene hindurch zu einer möglichen Quelle verfolgt [141, Kap. 2]. Häufig findet diese Abbildung auf einen Bildschirm statt, die Strahlen können dann parallel konstruiert werden. Wenn diese Strahlen unterbrochen werden, ist eine weitere Berechnung nicht notwendig, das Objekt verbirgt sich hinter einem anderen Objekt. Im Falle des in Abbildung 4.9 dargestellten vorwärtsgerichteten Raytracings wird vom Objekt aus auf die Schirmpixel konstruiert, im rückwärtsgerichteten Fall entsprechend von den Pixeln auf die Objekte hin [142, Kap. 2.1 f.]. Nach dem Ausgangsobjekt werden beide Arten des Tracings respektive auch *Object-Order* oder *Image-Order-Rendering* genannt [143, Kap. 4].

Die CH benötigt deutlich mehr Berechnungen: Da die Perspektive nicht bereits feststeht, muss die Information des gesamten Objekts aufgenommen werden. Statt der wie oben beschriebenen *Projektion* bedarf es also einer echten Berechnung der *Propagation*,

wie sie qualitativ bereits in Abschnitt 2.2 angedeutet wurde. Es genügt nicht, aus jedem Punkt des Objekts eine Intensitätsinformation am Ort des Schirms zu berechnen, vielmehr muss der vollständige komplexe Wert der Wellenfunktion berechnet werden. Auf dem holographischen Schirm wird erst im zweiten Schritt eine Intensität aufgenommen. Zuvor erfolgt die Interferenz aller beteiligten Wellen mit der Referenzwelle.

Die Darstellung des Objekts innerhalb der Simulation richtet sich dabei stark nach den verwendeten Propagations-Algorithmen und wird im Folgekapitel 5 genauer thematisiert.

4.3.2 Wellenpropagation

Für die Beschreibung der Wellen wird immer die Wellengleichung (2.12) gelöst. Im Folgenden soll der Prozess der Berechnung der Welle an einem beliebigen Ort mit *Propagation* bezeichnet werden, deren Erläuterung folgt dabei im Wesentlichen Ref. [11, Kap. 5].

Im Allgemeinen sind für deren exakte Bestimmung aber Randbedingungen erforderlich. Eine mögliche Randbedingung im Rahmen der Lichtholographie ist etwa eine Blende, durch die die Welle tritt. Ist die Blende in Relation zur Wellenlänge hinreichend klein, tritt Streuung auf.

Beschrieben wird sie, wie beispielsweise in Abbildung 4.10a gezeigt, durch eine binäre Blendenfunktion

$$b(x, y): \mathbb{R}^2 \rightarrow \{0, 1\} \quad (4.45)$$

mit einem festen $z = z_0$, die innerhalb der Blende 1 und außerhalb 0 zurückgibt.

Wir hatten bereits in Abschnitt 2.3 die für Anwendungen in der Holographie irrelevante Zeitabhängigkeit aus der allgemeinen Wellenfunktion entfernt. Das bedeutet nun, dass wir keine Information über die konkrete „Zeitrichtung“ der Wellen haben. Deshalb sprechen wir per Konvention bei Ausbreitung in Richtung des Wellenvektors \mathbf{k} von *Vorwärts-*, im

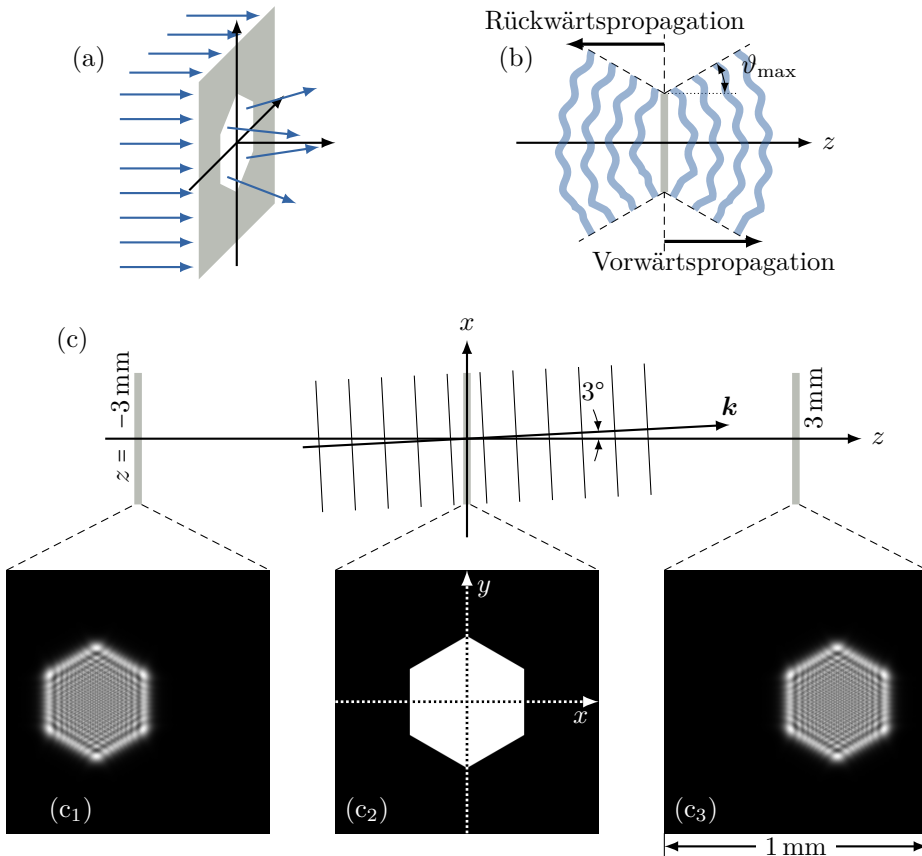


Abbildung 4.10: Darstellung der Propagation nach Ref. [11, Abb. 5.1 ff.]. (a) Ist eine Blende hinreichend klein gegen die Wellenlänge, tritt Streuung auf. (b) In der Simulation ist durch Ausblenden der Zeitkomponente eine Vorwärts- und Rückwärtspropagation entlang einer ausgezeichneten Richtung definierbar. Dabei lässt sich die Propagation nach Gleichung (4.42) auf einen Maximalwinkel ϑ_{\max} einschränken. (c) In einem konkreten Simulationsaufbau ergibt die Blende (c_2) in beide Richtungen durch schrägen \mathbf{k} -Vektor ein verschobenes Beugungsbild. Es wurde dabei das Licht eines HeNe-Lasers mit $\lambda = 633 \text{ nm}$ angenommen, die Pixel aller drei Grafiken haben einen Abtastabstand von $\Delta x = \Delta y = 2 \mu\text{m}$.

gegenläufigen Fall von *Rückwärtspropagation*, vergleiche dazu auch die Darstellung in Abbildung 4.10b und die Ausführungen in Anhang A.3. Im Beispiel der Abbildung 4.10 zeigt der Wellenvektor dabei primär in z -Richtung.

Grundsätzlich sind auch andere Randbedingungen denkbar. Bei der Rekonstruktion wird das Hologramm etwa neu beleuchtet und damit selbst zu einer Randbedingung

$$b(x, y): \mathbb{R}^2 \rightarrow [0, 1]. \quad (4.46)$$

Dieser Vorgang soll an dieser Stelle soweit formalisiert werden, dass er im Folgekapitel tatsächlich konkret umgesetzt werden kann. Dazu führen wir den Translationspropagationsoperator \mathcal{P}_d ein. Er bilde die Amplituden- und Phasenverteilung an einer Ebene auf eine parallele Ebene im Abstand d mittels

$$\mathcal{P}_d\{\psi(x, y; z_0)\} = \psi(x, y; z_1) \quad (4.47)$$

ab, wobei sich der Abstand zu

$$d = z_1 - z_0 \quad (4.48)$$

ergibt.

Die Rückpropagation ist daher durch \mathcal{P}_{-d} beschrieben. Entsprechend muss

$$\mathcal{P}_{-d}\{\mathcal{P}_d\{\psi(x, y; z_0)\}\} = \psi(x, y; z_0) \quad (4.49)$$

gelten.

Wenn statt einer festen Ebene $z = z_0$ nun aber zusätzlich die dritte Dimension als Parameter zugelassen wird, kann die Beschreibung beliebig komplex werden. Das ist nicht nur im Rahmen dreidimensionaler abzubildender Objekte oder der in Abschnitt 2.4 beschriebenen Volumenhologramme der Fall, sondern etwa auch bereits bei einem rotierten Schirm, wie er im Rahmen dieser Arbeit Anwendung findet.

Daher abstrahieren wir den Propagationsoperator, indem wir ihn mittels

$$\mathcal{P}_V\{\psi(\mathbf{r})\} = \psi(\mathbf{r}'), \quad \mathbf{r}' \in V \quad (4.50)$$

in ein beliebiges Volumen V abbilden lassen. Eine Rückpropagation sei durch

$$\mathcal{P}_{-V} \quad (4.51)$$

notiert.

Damit ist die Aufgabe eines Algorithmus, diesen Propagationsoperator zu implementieren. Es existieren dazu viele Ansätze, die in Kapitel 5 angerissen werden sollen. Die effiziente Berechnung dieser Propagation ist wesentliche Aufgabe der CH. Gelingt diese Berechnung in Echtzeit, ist etwa die reale Konstruktion holographischer Displays erst denkbar.

4.3.3 Schirm

Die in Abschnitt 2.3 eingeführte Fotoplatte, die in Kapitel 3 ihre Ablösung durch einen CCD-Sensor fand, soll hier als Schirm formalisiert werden. Er sei neben der eigenen Lage insbesondere durch seine Auflösung und Größe charakterisiert. Zur Vereinfachung seien nur quadratische, ebene Schirme angenommen. Sie sollen eine Kantenlänge von d und in jeder der beiden Richtungen jeweils N Pixel besitzen. Insgesamt werden also N^2 Pixel in einem quadratischen Raster auf einer Fläche d^2 liegen. Diese Pixel seien äquidistant positioniert und infinitesimal klein. Daher gilt für ihren Abstand untereinander

$$\Delta x = \Delta y = d/N. \quad (4.52)$$

Der durch die Simulation rein virtuelle Schirm ist grundsätzlich in der Lage, die Amplituden- und Welleninformationen des Wellenfelds an seinen Messpunkten anzugeben, da diese für jeden weiteren Schritt zunächst berechnet werden müssen. In der Realität wird aber nur die Feststellung der Intensität am Ort der Schirmpixel möglich sein, weshalb die reine Wellenfeldinformation bewusst nicht genutzt wird. Stattdessen wird das Betragquadrat zur Ermittlung dieser Intensität berechnet und als eigentliche Messgröße am

Schirm ausgegeben. Genau hier kann das Aliasing, wie in Abschnitt 4.2.2 beschrieben, ein großes Problem aufwerfen. Die richtige Wahl der Schirmparameter ist also unabdingbar.

Die gemessene Intensität wird nun angepasst. Wie in Abschnitt 2.3 beschrieben, besteht zwischen Transmissionsgrad für die Rekonstruktion und eigentlicher Intensität eine medienabhängige Beziehung, die hier zu Gleichung (2.35) angenommen wurde.

4.3.4 Rekonstruktion

Das aufgezeichnete Feld zeigt nun nicht das abzubildende Objekt, sondern trägt die Information über dessen „Aussehen“ in anders kodierter Form in sich. Es braucht eine Rekonstruktion.

Das Rendern einer solchen Rekonstruktion ist aufwendig, da etwa *Occlusion* zu beachten ist, also die Verdeckung von Objekten im Hintergrund durch Objekte, die im Vordergrund stehen. Diese ist abhängig von der Perspektive – ein am Computer rekonstruiertes Hologramm muss aber alle möglichen Perspektiven anbieten können [11, Kap. 11.1]. Für die Arbeit an der Elektronenholographie ist das nicht weiter relevant. Ziel ist es nicht, eine möglichst treffende dreidimensionale Darstellung für einen Beobachter zu ermöglichen, sondern einen „absoluten“ Blick auf das Rekonstruktionsvolumen zu werfen. Szenenparameter der Lichtholographie, wie etwa Perspektive und Fokus, verlieren hier also ihre Bedeutung.

Im Rahmen der Simulation wird die Rekonstruktion in ein konkretes Volumen durchgeführt. Auch dieses soll symmetrisch und damit würfelförmig sein. Insgesamt N_r^3 Pixel liegen dazu analog zu den Beschreibungen zum Schirm in einem Volumen d_r^3 .

Nach Gleichung (2.37) handelt es sich hierbei lediglich um eine Rückpropagation der normierten Transmission multipliziert mit der Referenzwelle in das Rekonstruktionsvolumen. Durch Intensitätsbildung ist dann

die Rekonstruktion abgeschlossen, gegebenenfalls wird das erhaltene Bild anschließend noch aufbereitet.

4.3.5 Oversampling Ratio

Wir haben bereits festgestellt, dass eine Unterabtastung zu Fehlern in der Holographie führt. An dieser Stelle soll die Überabtastung quantisiert werden. Denn sie bringt, wie in Abschnitt 4.2.2 beschrieben, mehr Daten mit sich, die gespeichert, verarbeitet und berechnet werden müssen.

Sind auf einem simulierten Schirm N_{scr} Pixel vorhanden, so liegen hier N_{scr} reelle Zahlen vor. Im Rekonstruktionsvolumen mit N_{rec} Pixeln gilt es aber damit N_{rec} komplexe Zahlen und damit $2N_{\text{rec}}$ Unbekannte zu bestimmen. Damit das daraus konstruierbare Gleichungssystem nicht unterbesetzt ist, muss

$$N_{\text{scr}} \geq 2N_{\text{rec}} \quad (4.53)$$

gelten. Durch diese Ungleichung motiviert, lässt sich das *Oversampling Ratio*

$$\sigma = \frac{N_{\text{scr}}}{N_{\text{rec}}} \geq 2 \quad (4.54)$$

definieren [144]. Im Falle unseres Problems gilt

$$N_{\text{scr}} = N^2 \quad \text{und} \quad N_{\text{rec}} = N_r^3. \quad (4.55)$$

Im rein d -dimensionalen Fall lässt sich das lineare Oversampling Ratio

$$\sigma' = N/N_r \quad (4.56)$$

definieren. Dabei seien N und N_r wieder Pixelzahlen in jeweils einer Dimension. Ganz allgemein ergibt sich für das lineare Oversampling Ratio damit aus Gleichung (4.54)

$$\sigma' \geq 2^{1/d} \quad (4.57)$$

als Voraussetzung für eine d -dimensionale Szene [145]. In dieser Arbeit kann es wohl nur als erste Näherung gelten, denn hier wird ein

zweidimensionales Objekt (der Schirm) mit einem dreidimensionalen Rekonstruktionsvolumen verknüpft.

In einem realen Experiment muss ferner immer davon ausgegangen werden, dass einzelne gemessene Intensitätswerte aufgrund von Fehlern fehlen oder maskiert werden müssen. Sei N_{mis} die Zahl fehlender Pixel, so ergibt sich daraus ein Verhältnis

$$\eta = N_{\text{mis}}/N_{\text{scr}} \quad (4.58)$$

fehlender zu allen Pixeln. Damit ist

$$N_{\text{val}} = N_{\text{scr}} - N_{\text{mis}} \stackrel{(4.58)}{=} (1 - \eta)N_{\text{scr}} \quad (4.59)$$

die Zahl der Pixel, die tatsächlich für die Rekonstruktion verwertbar sind [146]. Aus diesem Grund muss bei Zulassung von Fehlern statt Gleichung (4.54) genauer

$$\frac{N_{\text{val}}}{N_{\text{rec}}} \stackrel{(4.59)}{=} \frac{(1 - \eta)N_{\text{scr}}}{N_{\text{rec}}} \stackrel{(4.54)}{=} (1 - \eta)\sigma \geq 2 \quad (4.60)$$

gelten. Für den Anteil fehlender Pixel ergibt sich damit

$$\eta = 1 - \frac{2}{\sigma}. \quad (4.61)$$

Die doppelte Schirmgröße (genauer: Schirmfläche) bei gleichbleibenden Pixelabständen erlaubt also, dass die Hälfte der Pixel zur Bestimmung der Rekonstruktion ausreicht.

Damit wurden anhand der Computerholographie zwei wesentliche Parameter zur Bestimmung der Schirmdimension bestimmt. Eine Mindestanzahl an Pixeln wird einerseits durch das Abtasttheorem (4.37) in Abhängigkeit der abzubildenden Wellen vorgegeben, andererseits gibt die Anzahl der Pixel im Rekonstruktionsvolumen – die ihrerseits wieder anhand des Abtasttheorems festzulegen sind – durch das Oversampling Ratio eine Untergrenze an Pixeln an. Zusätzlich können weitere Pixel durch Erhöhung des Oversampling Ratio zur Stabilisierung aufgeschlagen werden. Dies geht allerdings auf Kosten der Laufzeit und des Speicherbedarfs.

5 Algorithmen

In Kapitel 4 haben wir uns mit der Computer Holography als aus theoretischen Überlegungen abgeleiteter Methode beschäftigt. Auf dem Weg zur Entwicklung beispielsweise der genannten realistischen Darstellung von echt dreidimensionalen Bildern ohne Hilfsmittel über Displays existieren bereits verschiedenste Programme, die die bisherigen theoretischen Überlegungen der Vorkapitel umsetzen. Sie behandeln typischerweise Experimente der in Kapitel 2 beschriebene Lichtholographie. Die dort gewonnenen Erkenntnisse der konkreten algorithmischen Umsetzung sollen hier erläutert und an die Besonderheiten der Elektronenholographie angepasst werden.

Erste theoretische Ansätze zur Berechnung von CGH lassen sich bis 1966 zurückdatieren [147]. Zunächst wurden in experimentellen Umsetzungen Binärhologramme berechnet, deren holographische Muster also noch keine Graustufen enthielten [148–150]. Auch hier trat analog zu den in Abschnitt 2.1 beschriebenen Experimenten ein störendes Zwillingsbild auf. Analog zur dargelegten Entwicklung der Lichtholographie gelang es durch einen hinreichend großen Winkel zwischen Objektebene und Referenzwelle, dieses Zwillingsbild in der Rekonstruktion vom eigentlichen Objekt zu trennen [151].

Bei der konkreten Implementierung ist es nötig, eine geeignete Beschreibung des abzubildenden, dreidimensionalen Objekts zu finden. Das gelingt etwa durch ein Array aus Koordinaten und Intensitätswerten im Objekt liegender Punkte. Ferner gilt es dann, aus dieser Ansammlung von Daten, Pixel genannt, ein typischerweise zweidimensionales Hologramm zu berechnen [152].

Erschwerend hierbei sind diese dreidimensionalen Szenen und Hologramme für reale Anwendungen sehr groß; die Pixelabstände hingegen in der Größenordnung der holographie-

renden Welle und damit sehr klein. Es werden also viele Punkte auf ebenfalls viele Pixel abgebildet, was wiederum zu einem sehr großen, exponentiell wachsenden Rechenaufwand führt. An ein typisches Video mit 25 Bildern je Sekunde ist damit scheinbar nicht zu denken.

Im Rahmen dieses Kapitels werden Zusammenhänge zwischen Experiment und Simulation beschrieben. Hierbei unterscheiden wir zwischen algorithmischen Aspekten und Hardware-Fragestellungen. Abschließen soll das Kapitel mit einer Übersicht über die vom Autor der vorliegenden Arbeit für die EBSD-Holographie geschriebene Bibliothek, die in den Folgekapiteln ihre Anwendung findet. Sie erlaubt die Berechnung der Propagation von verschiedenen Wellen, insbesondere aber auch von Elektronenwellen unter Berücksichtigung der in Abschnitt 3.1.2 beschriebenen Besonderheiten. Da sie auf etablierten Konzepten aufbaut, müssen diese zunächst vorgestellt werden.

5.1 Punktbasiert

Eine Gruppe von Algorithmen folgt bei der Berechnung der Propagation stark der Argumentation in Abschnitt 2.3. Diese *punktbasierten* Algorithmen bilden die Grundlage der Forschungen in dieser Arbeit und sollen im Folgenden anhand der Überlegungen aus Ref. [152, Abschn. 2 ff.] erläutert werden. Ihre Grundidee ist in Abbildung 5.1 skizziert.

Das Ausgangsobjekt sei eine dreidimensionale, mittels CG dargestellte Form. Im Rahmen der CGH liegen, wie in Abschnitt 4.3 beschrieben, die Daten in diskreter Form vor. Typischerweise bedeutet das im Falle der Ausgangsobjekte, dass sie als eine Ansammlung selbstleuchtender Punkte verstanden werden können. Ihnen werden jeweils Koordinaten

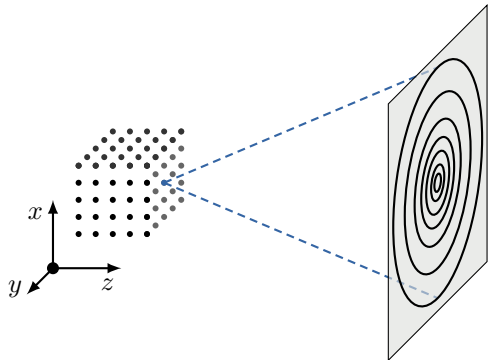


Abbildung 5.1: Punkt-basierte Propagation nach Ref. [26, Abb. 2a]: Das Objekt wird durch einzelne Punkte abgebildet, die ihrerseits einzeln auf den holographischen Schirm propagiert werden.

und Amplituden zugewiesen. Letztere stellen dabei den Wert der Welle am Ort des Punktes dar, mit der das Objekt beleuchtet wurde. Nach dem Huygens'schen Prinzip können mit diesen Amplituden dann neue Wellen ausgesendet werden, siehe Abschnitt 2.3.

Die einzelnen Punkte können als Pixel aufgefasst werden, in die das Ausgangsobjekt aufgeteilt wurde. Nehmen wir ohne Beschränkung der Allgemeinheit ein Koordinatensystem an, in deren xy -Ebene die holographische Schirmfläche liegt, so lässt sich jeder Pixel durch eine diskrete x - und y -Koordinate und einen Abstand z zum holographischen Schirm charakterisieren. Diese Pixel sind typischerweise äquidistant gewählt, so dass sie mathematisch leichter durch diskretisierte Koordinatenzähler ν und μ behandelt werden können. Auch in realen Anwendungen sind diese Pixel nahezu äquidistant. Zur industrielle Herstellung der häufig verwendeten CCD-Sensoren sei auf Ref. [8, Ahg. 1] verwiesen.

Die Pixelkoordinaten ergeben sich damit zu einer diskreten Form

$$(\nu\Delta x, \mu\Delta y, z). \quad (5.1)$$

Für spätere Anwendungen lassen sich diese über eine Koordinatentransformation verallgemeinern. Im Rahmen dieser Arbeit werden solche Transformationen bei Schirmen eingeführt, die nicht parallel zu einer Koordinatenebene liegen. Denkbar wäre aber auch die Implementierung von halbkugelförmigen Schirmen, wie sie etwa im Rahmen von Ref. [52, Kap. 2 ff.] bei der Holographie mittels Photoelektronenbeugung, englisch X-ray photoelectron diffraction (XPD), angewendet werden.

Für die Bestimmung des Hologramms ist zunächst die Summe der Kugelwellen dieser Pixel auf der Hologrammfläche zu bestimmen. Nach Addition der Referenzwelle ergibt sich das Hologramm aus dem Betragsquadrat, siehe Abschnitt 2.3. Die Summe aller beteiligten Partikularwellen sei im Folgenden bereits in den diskreten Pixel-Koordinaten mit

$$H(\nu, \mu) \stackrel{(A.2)}{=} \sum_n \psi_{o,n}(\nu\Delta x, \mu\Delta y, z) \quad (5.2)$$

bezeichnet. H bezeichnet dabei die Summe der Partikularwellen, anschaulich also die Wellenfunktion an Stelle des Hologramms vor Überlagerung mit der Referenzwelle.

Abbildung 2.3 illustrierte bereits, dass ein einzelner Objektpunkt dabei eine Fresnel-Zonenplatte abbildet. Wenn dieser sich lotrecht über dem Zentrum des holographischen Schirms im Abstand von z_0 befindet, wird dessen Abbildung durch eine Funktion

$$F(\nu, \mu; z_0) \propto \exp\left(\frac{2\pi i}{\lambda} \sqrt{(\nu\Delta x)^2 + (\mu\Delta y)^2 + z_0^2}\right) \quad (5.3)$$

beschrieben. Es handelt sich dabei anschaulich um eine Ansammlung von Kreisringen, die aufgrund ihrer elementaren Bedeutung auch Principal Fringe Pattern (PFP) genannt wird. Damit lässt sich Gleichung (5.2) nun zu

$$H(\nu, \mu) = \sum_n A_n F(\nu - \nu_n, \mu - \mu_n; z_n) \quad (5.4)$$

als Summe von Fresnelzonenplatten von selbstleuchtenden Punkten mit Amplitude A_n , diskreten Koordinaten ν_n, μ_n und Abstand z_n umschreiben. Das sind also die Parameter, die jedes Atom in der späteren Berechnung unterscheiden.

Für das Erstellen eines Hologramms der Größe N^2 muss damit für jeden der insgesamt P Pixel eines Objekts eine Berechnung mit arithmetischem Aufwand A durchgeführt werden. Damit ergibt sich die Problemkomplexität zu

$$\mathcal{O}(N^2 \cdot P \cdot A) \quad (5.5)$$

nach der in Abschnitt 4.2.3 eingeführten Landau-Notation. Die Laufzeit steigt damit im Wesentlichen quadratisch mit der Hologrammgröße und linear mit der Objektpixelanzahl, die ihrerseits bei einem dreidimensionalen Objekt kubisch mit der Objektgröße oder -auflösung wächst.

Im Folgenden sollen nun Methoden beschrieben werden, die diese Komplexität begrenzen. Denn nur mit geringerem Rechenaufwand ist eine Skalierung der Holographie auf realistischere Fragestellungen und Objektauflösungen denkbar.

5.1.1 Faltungsmethode

Handelt es sich bei dem Objekt um eine einzige Ebene im Abstand z_0 parallel zum Schirm, so lässt sich Gleichung (5.4) als Faltung

$$H(\nu, \mu; z_0) = A(\nu, \mu; z_0) * F(\nu, \mu; z_0) \quad (5.6)$$

beschreiben, wobei A die Amplitude der Objektebene angibt. Seien \tilde{A} und \tilde{F} die Fourier-Transformierten der beteiligten Funktionen, so lässt sich der Ausdruck (5.6) durch das Faltungstheorem (4.17) zunächst scheinbar umständlicher als

$$H(\nu, \mu; z_0) = \mathcal{F}^{-1} \left\{ \tilde{A}(\omega_\nu, \omega_\mu; z_0) \cdot \tilde{F}(\omega_\nu, \omega_\mu; z_0) \right\} \quad (5.7)$$

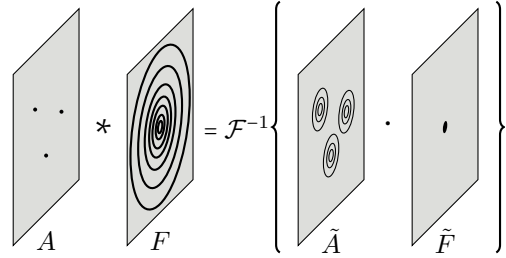


Abbildung 5.2: Der Faltungsmethode liegt das gleichnamige Theorem zugrunde. Nach der Transformation entspricht eine einfach zu berechnende Multiplikation der aufwändigeren Faltung. Die scheinbar komplexere Fourier-Transformation lässt sich aufgrund spezialisierter Algorithmen effizient anwenden.

formulieren. Grafisch ist dieser Vorgang in Abbildung 5.2 dargestellt.

Allerdings ist die Fourier-Transformation dank etwa dem in Abschnitt 4.2.3 beschriebenen FFT-Algorithmus deutlich effizienter berechenbar als die Faltung. Insbesondere mithilfe der GPU, also der Graphics Processing Unit, lassen sich so schneller und parallelisiert Hologramme berechnen, da der arithmetische Aufwand gegenüber der „einfachen“ Methode sinkt, genauer wird dies noch in Abschnitt 5.3 beschrieben.

Eine Ebene ist dabei zur Beschreibung eines dreidimensionalen Objekts selbstverständlich nicht ausreichend. Es verfügt mit der Tiefe über einen weiteren Parameter, der in dieser Darstellung Berücksichtigung finden muss. Das kann durch ein ebenenbasiertes Sampling geschehen, wie es in Abbildung 5.3 gezeigt ist. Dabei wird das Objekt in parallele Ebenen aufgeteilt, die abschließend gemäß

$$H(\nu, \mu) = \sum_n H(\nu, \mu; z_n) \quad (5.8)$$

zu summieren sind [153]. Bei wenigen Ebenen ist diese Methode damit durchaus lohnend, mit zunehmender Ebenenzahl nähert sich die

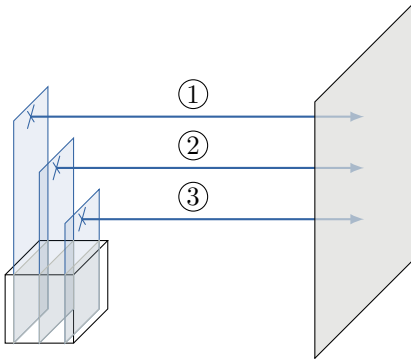


Abbildung 5.3: Ist ein Objekt senkrecht zum Schirm ausgedehnt, lässt es sich in einzelne Ebenen zerlegen. Diese lassen sich dann einzeln nach verschiedensten Methoden propagieren. Die jeweiligen Ergebnisse müssen abschließend lediglich summiert werden.

Laufzeit aber sehr der der einfachen Summe an, sodass sich der zusätzliche algorithmische Aufwand durch eingeführte Komplexität der Berechnung nicht lohnt [26].

Die Berücksichtigung der Selbst- oder Fremdüberdeckung von Objekten bei eingeschränktem Sichtbereich erlaubt den Verzicht auf die Simulation einiger Objekte oder Objektteile. Durch Einführen von Masken lässt sich dieser ebenenbasierte Prozess optimieren, indem für den Betrachter faktisch immer unzugängliche Punkte ausgeblendet werden [11, Kap. 2.4.3.5].

5.1.2 Lookup-Tabellen

In Analogie zu Logarithmentabellen, die zuvor berechnete Werte als Nachschlagewerk enthielten [154, Kap. 5 f.], können auch immer wieder getätigte Berechnungen am Computer durch das Nachschlagen einmalig bestimmter Werte signifikant beschleunigt werden.

Bei Betrachtung der Gleichung (5.4) oder Abbildung 5.4 fällt auf, dass wesentlicher Bestandteil die stets neue Berechnung derselben

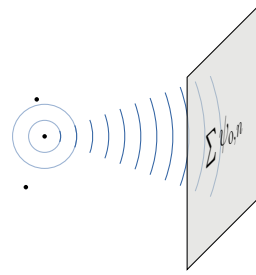


Abbildung 5.4: Der „naive“ Ansatz der Berechnung nach punktbasierter Methode sieht vor, die Summe aller Streuer einzeln zu berechnen. Das Ergebnis jedes einzelnen Summenglieds ist ein Kreisringssystem: die bereits vertraute Fresnelzonenplatte.

Funktion innerhalb einer Summe, der Fresnelzonenplattenfunktion F , ist. Man greift daher auf eine sog. Lookup Table (LUT) zurück, die zu bestimmten Werten einen Ergebniswert vorberechnet und bereithält.

Werden nun neben jedem die Atome beschreibenden Parameter ν_n , μ_n und z_n auch für die Schirmkoordinaten ν und μ alle Werte der Fresnelzonenplatten vorberechnet und in einer Tabelle gespeichert, so können zur Laufzeit der Hologrammberechnung die zugehörigen Ergebnisse einfach „nachgeschlagen“ werden. Abbildung 5.5a stellt dies anschaulich dar. Die richtige Fresnelzonenplatte ist dann nur noch mit der Amplitude des Punktstreuers zu multiplizieren. Die Komplexität der Berechnung reduziert sich dann auf

$$\mathcal{O}(N^2 \cdot P), \quad (5.9)$$

da ein Lookup gegenüber der Laufzeit der bisherigen arithmetischen Berechnungen A vernachlässigbar ist. Diese Art der LUT war bereits etabliert, bevor sie erstmals auf das konkrete Problem der Computerholographie angewandt wurde [155].

Das Berechnen dieser Tabellen bringt jedoch auch Nachteile mit sich: Ähnlich der seitenlangen Logarithmentabellen, benötigen auch

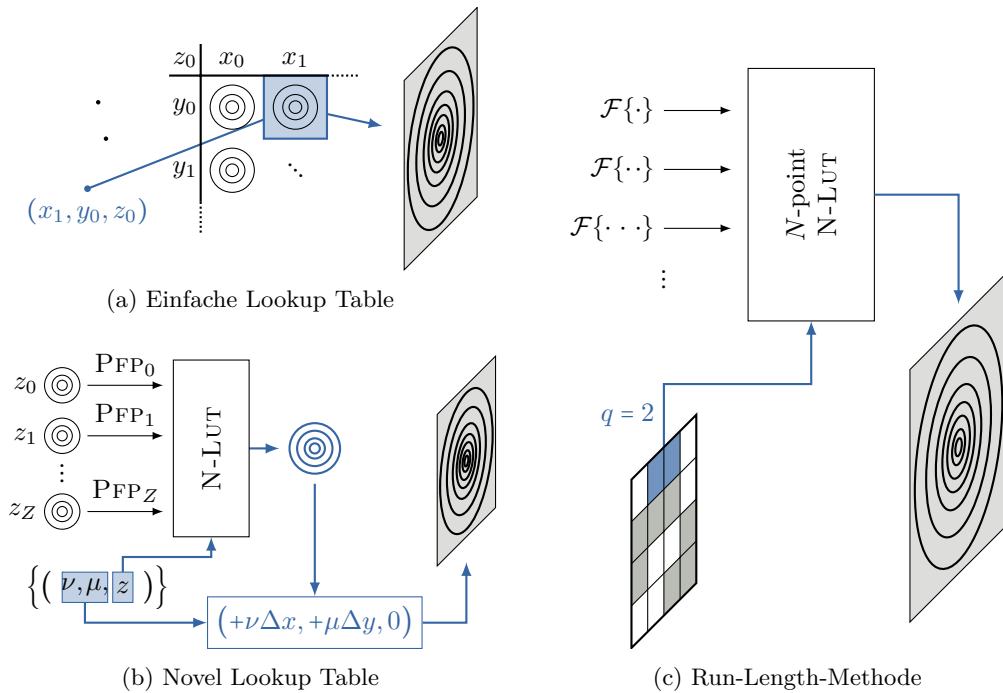


Abbildung 5.5: Verschiedene Arten der Implementierung von Lookup-Tabellen. (a) Einfache Methode, die alle Koordinaten speichert und für einzelne Streuer und Pixelkoordinaten abfragt. (b) N-LUT-Methode, in der Kreisringssysteme verschiedener Ebenen z_n verschoben werden. Darstellung nach Ref. [152, Abb. 2]. (c) Durch Vereinigung ähnlicher, nebeneinanderliegender Pixel lässt sich die Laufzeit drastisch reduzieren. Die Darstellung folgt hierbei Ref. [152, Abb. 3].

die LUT viel Speicherplatz. Hier werden beispielsweise fünf Koordinaten angenommen, sodass sich eine fünfdimensionale Datenstruktur ergibt. Sei Z die Anzahl der angenommenen z_n -Ebenen, die vorab zu berechnen sind, und ferner B die Größe des Datentyps der Ergebniswerte, so ergibt sich ein numerischer Speicherbedarf S von

$$S = N^2 \cdot N^2 \cdot Z \cdot B. \quad (5.10)$$

Da komplexe Zahlen abzubilden sind, beträgt der Speicheraufwand für Real- und Imaginärteil mindestens $B = 2B = 16$ bit, typischerweise aber 32 bit oder sogar 64 bit und damit das Achtfache. Im Falle eines typischen Hologramms mit je $N = 1024$ Pixeln Kantenlänge muss eine solche LUT bei Annahme von gerade einmal $Z = 8$ Tiefenebenen also einen Speicheraufwand von mindestens $S = 16$ TiB haben. Damit ist diese Tabelle zu groß für das Ablegen in einer schnell verfügbaren Speicherform. Selbst wenn derartige Speicherkapazitäten bereitstünden, wäre der Speicher im Rahmen eines Clusters auf mehrere physische Speichereinheiten verteilt. Etwaige Transferzeiten bei der Datenübertragung können hierbei Geschwindigkeitsvorteile durch Vorberechnung und Speicherung sehr schnell zunichte machen.

Mithilfe der in Abbildung 5.5b dargestellten Novel Lookup Table (N-LUT) nach Ref. [156] lässt sich dieser Platzbedarf stark einschränken. Ihr liegt die Idee zugrunde, die Fresnelzonenplatte als verschoben anzunehmen. Es wird also für jede Ebene z_n nur noch ein PFP berechnet, das dann lotrecht unter den jeweils untersuchten Objektpunkt geschoben wird. Anstatt für Kombinationen aus beispielsweise ν und ν_n eine LUT zu erstellen, wird diese damit effektiv nur für die Differenz $\nu - \nu_n$ erstellt. Es werden also mehr arithmetische Operationen durchgeführt, da nun immer eine Differenz zu bilden ist. Additionen sind aber günstig durchzuführen, dementsgegen steht eine Reduktion der LUT um zwei Dimensionen.

Der Speicherbedarf sinkt im oben gezeigten Beispiel um drei Größenordnungen auf

$$S = N^2 \cdot Z \cdot B = 16 \text{ GiB}. \quad (5.11)$$

Im Fall der Lichtholographie lässt sich diese Größe durch Anpassung der Auflösung der PFP in Abhängigkeit des z -Abstands und der Beachtung des Auflösungsvermögens des Betrachters weiter dramatisch reduzieren [157], im Falle der Elektronenholographie mit der Rekonstruktion am Computer hilft dieser optisch begründete Reduktionsansatz aber nicht weiter, weil hier keine konkrete „Betrachtungsperspektive“ definierbar ist. Vielmehr soll eine absolute Sicht auf jeden Pixel im Rekonstruktionsvolumen gewonnen werden.

Eine andere Möglichkeit ist die Reduzierung der anzunehmenden Punkte mithilfe eines Run-Length-Encoding, wie es in Abbildung 5.5c dargestellt ist [158]. Objektpunkte ähnlicher Amplitude liegen auf einer Fläche oft nah beieinander, etwa weil sie Teil einer homogenen Fläche sind. Daher lassen sich Bereiche solcher Punkte zu einem virtuellen Punkt zusammenfassen – die Methode wird effektiv regionenbasiert. Für das Nachschlagen in der N-LUT sind lediglich die Parameter Schirmabstand z_n und Amplitude relevant. Benachbarte Punkte werden daher zusammengefasst, wenn sie in diesen beiden Parametern nur einen geringen Unterschied aufweisen. In der praktischen Ausgestaltung findet das Zusammenfassen etwa in einer ausgezeichneten Richtung auf den Objektebenen statt. Die neu gefundenen Punkte können dann anhand der genannten N-LUT propagiert werden.

Bei Betrachten der Fresnelzonenplatte fällt weiterhin auch auf, dass sie rotationssymmetrisch ist. Diese Eigenschaft macht sich die Line-Scan-Methode nach Ref. [159] zunutze. Das Kreismuster wird auf ein Linienmuster reduziert, das über die Bestimmung eines Radius' ausgewertet werden kann. Sei R die ra-

diale Auflösung des anzunehmenden PFP, so lässt sich der Speicheraufwand auf

$$S = R \cdot Z \cdot B = 16 \text{ KiB} \quad (5.12)$$

reduzieren, wenn wir in erster Näherung annehmen, dass die radiale Auflösung der Hologrammkantenlänge $R = N$ entspricht.

Durch die starke Reduktion der Tabelle steigt aber der Rechenaufwand wieder signifikant. Zur Bestimmung des Radius muss multipliziert werden. Das Radizieren lässt sich vermeiden, wenn der quadrierte Radius als Auswahlparameter hinzugezogen wird.

Beim Erweitern der Linie in ein Kreismuster, etwa nach Ref. [160], ist ferner schnell feststellbar: Nicht jeder im rechteckigen Gitter angeordnete Hologrammpixel lässt sich über einen Radius zu einem beliebigen anderen Punkt erreichen. Es fehlen einzelne Pixel [161]. Statt auf den nächstliegenden Pixel zu runden, lässt sich genauer rechnen, wenn radial-symmetrisch interpoliert wird [162]. Abbildung 5.6c–d stellt den Unterschied beider Methoden heraus.

Eine andere Art der Speicherreduzierung des vollständigen Musters ist möglich, wenn die Berechnung der Fresnelzonenplatte (5.3) in eine horizontale und vertikale Komponente durch Verwendung eines Split-Lookup-Table-Frameworks nach Ref. [163] aufgeteilt wird. Beide Komponenten können unabhängig voneinander betrachtet und damit in getrennten LUT nachgeschlagen werden. Der Speicherbedarf reduziert sich dadurch bei gleichbleibender Berechnungsgrundlage von oben auf etwa

$$S = 2N \cdot Z \cdot B = 32 \text{ KiB}. \quad (5.13)$$

Ein solches Aufteilen der Fragestellung mit in LUT gespeicherten Ergebnissen zur Vermehrung der Anzahl an Tabellen bei gleichzeitiger Reduktion deren Speicheraufwands ist dabei bereits in der Vergangenheit angewendet worden. Im Bereich stark korrelierter

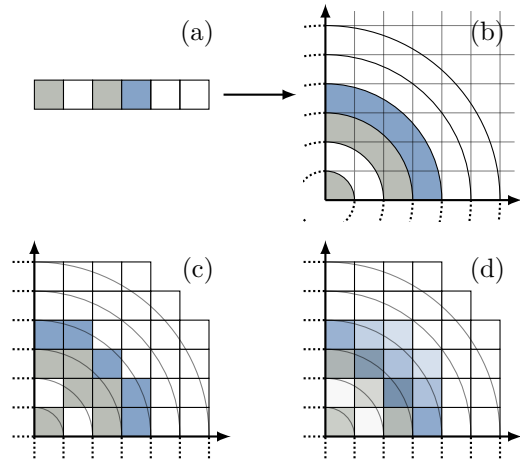


Abbildung 5.6: (a) Rotationssymmetrische Muster lassen sich als radiale Linie abspeichern. (b) Die Schirme verfügen über eine rechteckige Pixelanordnung, auf die die Expansion des Linescans nicht aufgelegt werden kann. (c) Die naheliegendste Methode ist also die Wahl des jeweils nächstliegenden Radialpixels. (d) Bessere Ergebnisse lassen sich durch Interpolation zwischen den Werten der radial benachbarten Pixel erhalten.

Materie geschah dies beispielsweise 16 Jahre früher mit den sog. Lin-Tabellen [164].

Zur weiteren Reduktion des Speicherbedarfs bei entsprechend deutlich höherem Rechenaufwand im Rahmen von jeder der vorgestellten Methoden wäre eine zusätzliche Kompression der Tabelle durch etablierte Algorithmen denkbar [165].

5.1.3 Wavefront Recording Plane

Zur weiteren Beschleunigung der punktbasierten Methoden wird in Ref. [166] erstmals eine Wavefront Recording Plane (WRP) umgesetzt. Dabei handelt es sich um eine virtuelle Projektionsebene mit möglichst geringem Abstand vor dem Objekt, wie sie in Abbildung 5.7 gezeigt wird. Da die Amplitudenwer-

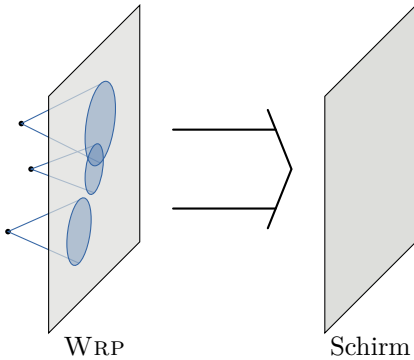


Abbildung 5.7: Die Wavefront Recording Plane ist eine virtuell eingeschobene Zwischenebene, die leichter zu berechnen und damit schneller zu propagieren ist.

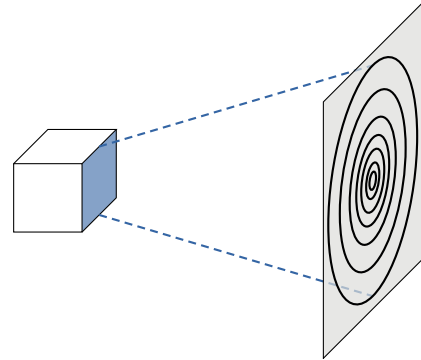


Abbildung 5.8: Im Rahmen der Polygon-Methode werden einzelne Flächen als selbstleuchtende Ebenen angenommen, aus denen direkt ein Muster berechnet wird.

te der Fresnelzonenplatte mit zunehmendem Abstand vom Ursprungspunkt stark abnehmen, kann bei einer gegenüber dem Schirm deutlich näher vor dem Objekt positionierten Ebene eine kleine Fresnelzonenplatte genähert werden. In einem zweiten Schritt wird eine Fresnel-Beugung der naheliegenden Ebene auf die ferne, eigentliche Hologrammebene berechnet.

Die Laufzeit der Berechnung der ersten Projektion ist dabei proportional zur Größe der Zonenplatte. Da sie im geringem Abstand mit kleinem Radius genähert werden kann, nimmt die Laufzeit mit dem Abstand zwischen Objekt und Projektionsebene ab. Daher eignet sich diese Methode für möglichst flache Objekte, da von der Projektionsebene weiter entfernte Objektpunkte sonst wieder große Fresnelzonenplatten mit entsprechend langer Laufzeit hervorrufen.

Spätere Ideen teilen das Objekt in zwei [167] oder mehr [168] Objektvolumina mit getrennten Ebenen. Eine andere Methode zur Beschleunigung ist die interpolierte WRP nach Ref. [169]. Hierbei wird das Objekt in gleich große Würfel geteilt, die in einen Punkt interpoliert werden – eine dreidimensionale Version des oben erwähnten Run-Length-Enco-

ding. Damit ist etwa schon die Erzeugung eines holographischen Films möglich [169].

Für alle oben genannten WRP-Methoden ist die Einbeziehung einer LUT, wie in Abschnitt 5.1.2 beschrieben, für weitere Beschleunigung denkbar [152].

5.2 Weitere Methoden

Punktbasierte Methoden bieten sich für die Betrachtung der Elektronenholographie sehr an, weil sie von punktförmigen, atomaren Streuern ausgehen. Im Rahmen der CH existieren aber viele weitere Methoden, die hier kurz thematisiert werden sollen. In einigen Fällen können sie Stärken ausspielen, die auch für die Betrachtung atomarer Streuer hilfreich sind.

Gerade bei großen zusammenhängenden Flächen von Objekten ist das Zerlegen in einzelne Pixel und deren einzelne Propagation sehr aufwändig. Stattdessen lässt sich ein Objekt auch in ebene Flächen zerlegen, die dann als selbstleuchtend angenommen werden. Abbildung 5.8 verdeutlicht das Prinzip: Diese Ebenen werden dann als Ursprung von den in Abschnitt 2.2.1 eingeführten ebenen Wellen approximiert [170].

Eine auf mehrere Ansätze der Holographie anwendbare Optimierung ist der Gedanke der *Sparsity*. Wird eine geeignete, invertierbare Transformation durchgeführt, lässt sich ein Signifikanzkriterium für die einzelnen Koeffizienten definieren. Ein solches Kriterium könnte beispielsweise über einen Betrag definiert sein, der gegenüber der ersten Ordnung verschwindend gering scheint. Dem nicht entsprechende und damit insignifikante Koeffizienten können vernachlässigt werden. Nach der Rücktransformation ist so eine gute Approximation bei deutlich geringerer Laufzeit möglich. Je nach Methode bieten sich dafür verschiedene Transformationen und Herangehensweisen an [171–174].

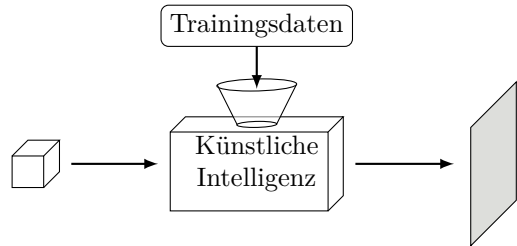


Abbildung 5.9: Methoden auf Basis einer künstlichen Intelligenz verhalten sich i. d. R. wie eine Blackbox, die durch Trainingsdaten selbstständig lernt, zwischen Objekt und Hologramm zu interpolieren und ihre Erkenntnisse auf neue Objekte anwendet.

In den letzten Jahren gewinnt auch die Verwendung des maschinellen Lernens an Bedeutung. Ein lernfähiger Algorithmus – etwa auf einem neuronalen Netz basierend – kann für die Belange dieser Arbeit als eine Black Box angenommen werden, wie sie in Abbildung 5.9 gezeigt ist. Anstatt einen konkreten Algorithmus zur Wellenpropagation zu implementieren, wird einem Programm ein Ausgangsobjekt und das daraus bereits berechnete Hologramm präsentiert. Durch *Training* des Programms anhand vieler dieser Entscheidungen soll es möglichst selbstständig lernen, welche Beziehung zwischen den beiden Bildern besteht. Nach einem erfolgreichen Training ist es dann möglich, dem Programm nur ein Ausgangsobjekt zu präsentieren, zu dem ein Hologramm selbstständig erzeugt wird.

Die Algorithmen profitieren dabei von Optimierungen aus anderen Forschungsbereichen der Bildverarbeitung. Insbesondere zeigt sich bereits, dass neuronale Netze konventionelle Algorithmen nicht nur in Geschwindigkeitsaspekten, sondern in Aufgaben der Bilderkennung auch in der Genauigkeit weit übertreffen können [175]. Weiterhin profitieren neuronale Netze von spezifischen Hardwareimplementierungen, die in Abschnitt 5.3 skizziert werden sollen.

Die sehr großen für ein Training benötigten Datensätze zählten bisher zu den gewichtigen Nachteilen dieser Methode. Zu wenig Trainings-Datensätze führen zu einer mangelhaften Anpassung des Algorithmus, die falsche Prognosen nach sich zieht. Jedoch werden immer mehr Algorithmen entwickelt, die mit weniger Daten und Ressourcen auskommen [176, 177] und sich für die Holographie anpassen ließen [178–180].

5.3 Hardware

Mit der Optimierung der Art der konkreten Hologrammberechnung ist damit die Möglichkeit zur Beschleunigung längst nicht ausgeschöpft. Stattdessen hat sich in den letzten Jahren auch immer spezifischere Hardware entwickelt, die teilweise explizit nur für die Anwendung in der Holographie genutzt werden kann. Da auch die vorliegende Arbeit von Hardwarebeschleunigung profitiert, soll an dieser Stelle in die bestehenden Möglichkeiten anhand eines historischen Abrisses nach Ref. [181] eingeführt werden.

Ursprüngliche Berechnungen fanden dazu ausschließlich auf der Central Processing Unit (CPU) statt, die entwickelt wurde, um

möglichst allgemeine Berechnungen durchzuführen. Ihre fundamentale Struktur ist bis heute unverändert geblieben [182, Kap. 5.9], weshalb auch eine echte Parallelisierung nur mit erheblichem Aufwand umsetzbar ist. 2005 wurden erste echte Dual-Core-Prozessor für Desktop-Computer vorgestellt [183, Kap. 1.1]. Diese neuen Prozessoren müssen nun weniger zwischen verschiedenen Aufgaben, die gleichzeitig abzuarbeiten sind, wechseln. Das spart Ressourcen und erlaubt die Entwicklung parallelisierter Algorithmen, die unabhängige Berechnungen gleichzeitig ausführen können [184, Kap. 1.19].

Die CPU verfügt über eine hohe Schaltfrequenz. Ein kurzer Taktzyklus erlaubt schnelle Berechnungen. Ferner ist die CPU in der Lage, Fließkommazahlen in sog. *einfacher Genauigkeit* (also einer Größe von typischerweise 32 bit) gut zu berechnen. Auf Kosten der Laufzeit lässt sich die Genauigkeit erhöhen, indem die Datentypgröße verdoppelt wird. Allerdings ist die CPU für eine serielle Abarbeitung optimiert, die Implementierung von parallelisierter Software – wie erwähnt – aufwändig und mit vielen Problemen behaftet. Diese treten etwa auf, wenn mehrere Teilprozesse in die gleiche Variable schreiben möchten. Typischerweise überschreibt der letzte Zugriff dann einen relevanten Wert. Solche Situationen werden *Race Conditions* genannt.

Für die Erarbeitung der Frameworks für die Berechnungen im Rahmen dieser Arbeit und folglich für die weiteren Ausführungen werden im Falle der CPU ausschließlich die gängige x86- bzw. 64-Architektur der Hersteller Intel und AMD zugrunde gelegt. Sie sind etabliert und alle der in Anhang B näher beschriebenen Bibliotheken bereits für diese Architektur implementiert.

Gegen Ende der 1990er Jahre kamen erste Graphics Processing Unit (GPU) als zusätzliche Chips zur Berechnung etwa von Computerspielegrafik auf [185, Kap. 1.3]. Sie sind in der Lage, mit Fließkommazahlen in

halber (16 bit), einfacher (32 bit) und doppelter (64 bit) Genauigkeit umzugehen, und arbeiten hoch parallelisiert. Im Gegensatz zur CPU mit typischerweise nicht mehr als 24 Kernen im High-End-Bereich verfügt die GPU über tausende Kerne und ist damit von Beginn an auf die parallele Berechnung ausgelegt.

Mit der Einführung der Compute Unified Device Architecture (CUDA) durch Nvidia im Jahr 2007 wurde es Entwicklern möglich, von dieser Parallelisierung mit einer deutlich geringeren Einstiegshürde zu profitieren und die GPU als Koprozessor für Berechnungen mit einzubeziehen. Daneben existieren auch weitere Frameworks, wie etwa die Radeon Open Compute Platform (ROCm) des Herstellers AMD und abstraktere Programmierschnittstellen wie die Open Computing Language (OpenCL). Die zunächst für Grafik-Rendering ausgewiesenen Grafikkarten ließen sich so deutlich universeller einsetzen. Etwa maschinelles Lernen, Bild- und Videobearbeitung, Ver- und Entschlüsselung und viele weitere Prozesse profitieren seitdem von ihrer Beschleunigung.

In dieser Arbeit wurde auf die CUDA zurückgegriffen. Sie verfügt über eine hochperformante FFT-Komponente und viele weitere portierte Bibliotheken [186, Kap. 8]. Ferner steht eine Schnittstelle für die verwendete Programmiersprache Python zur Verfügung. Gegenüber einer reinen CPU-Implementierung existiert allerdings ein Transferengpass. Die Daten müssen zur Berechnung aus der CPU in die GPU und nach der Berechnung wieder zurück transportiert werden. Insbesondere hierin liegt zusätzlicher Programmieraufwand und Optimierungspotential. Die im Rahmen dieser Arbeit entwickelte Bibliothek versucht im Rahmen der genannten Beschränkungen die Berechnungen effizient zwischen CPU und GPU aufzuteilen. Eine weitere Klasse von einsetzbaren integrierten Schaltungen ist das Field Programmable Gate Array (FPGA). Dessen konkrete

Verschaltung lässt sich nachträglich programmieren, sodass es sich um spezifisch anpassbare Hardware handelt. Dem deutlich größeren Programmieraufwand steht eine enorme Beschleunigung gegenüber. Eine Realisation findet ein Hardware-Design auf Basis von FPGA beispielsweise im mehrfach weiterentwickelten Computer HORN [187, 188].

Einsetzbar sind auch digitale Signalprozessoren, Digital Signal Processor (DSP) genannt, die auf die kontinuierliche Verarbeitung von digitalen Signalen, wie sie bereits in Abschnitt 4.1 behandelt wurden, optimiert sind [189].

Mit Intels Xeon Phi existiert ferner eine Klasse von Koprozessoren, die ähnlich wie Grafikkarten aufgebaut, aber mit dem bekannten x86-Befehlssatz der CPU programmierbar und ebenfalls für holographische Berechnungen geeignet sind [189]. Auch Hybride aus CPU und FPGA wurden für die Holographie entwickelt [190].

5.4 Eigenentwicklung

Zur Berechnung der in den Folgekapiteln vorgestellten Simulationen ist im Rahmen dieser Arbeit ein Ökosystem aus verschiedenen Bibliotheken zum Umgang mit holographischen Berechnungen am Computer entstanden. Die einzelnen Komponenten dieser *Holib* genannten Ansammlung werden detaillierter in Anhang B beschrieben. An dieser Stelle sollen die Kernkonzepte dieser Eigenentwicklung mit den zuvor beschriebenen Algorithmen verknüpft werden. Einzelne Bestandteile haben einen gewichtigen Einfluss auf die weitere Ausführung der Berechnungen, tragen zum Verständnis der noch vorzustellenden Ergebnisse bei und sollen daher an dieser Stelle knapp behandelt werden.

5.4.1 Wellenfelder

Das in Abschnitt 4.3.1 beschriebene Sampling eines Objekts erfolgt im Rahmen dieser

Bibliothek genau so wie die Beschreibung von Hologrammen. Mit dem `Wavefield` (B.1.2) wird eine Datenstruktur definiert, die einem kubischen Volumen innerhalb eines kartesischen Koordinatensystems komplexe Wellenwerte zuordnet. Die Koordinaten müssen dabei nicht äquidistant liegen und beschreiben damit eine nicht zwangsläufig homogene Verteilung von Pixeln, an denen eine Wellenfeldverteilung gemessen wird. Im Fall eines zu holographierenden Objekts entspricht der Wert einer rein realen Amplitude.

Der Propagationsoperator (4.47) ist innerhalb dieser Klasse definiert. Er ist durch einen punktbasierten Algorithmus implementiert, wie er in Abschnitt 5.1 beschrieben wurde. Er ist möglichst modular aufgebaut, sodass er sich leicht weiter optimieren lässt – etwa durch Einbau einer LUT. Insbesondere ist die Ausnutzung der vorhandenen Hardware-Ressourcen nach den in Abschnitt 5.3 beschriebenen Überlegungen optimiert worden, indem die in Anhang B beschriebenen Parallelisierungen durchgeführt wurden.

Das einfache Wellenfeld nimmt dabei die Propagation von Kugelwellen, wie sie in Abschnitt 2.2.2 beschrieben wurde, mit einer Summe über Wellen der Art

$$\psi_{\mathbf{W}}(\mathbf{r}) \stackrel{(2.26)}{=} \frac{A_{\mathbf{o}}}{k|\mathbf{r} - \mathbf{o}|} e^{-ik|\mathbf{r} - \mathbf{o}|} \quad (5.14)$$

an, wenn sich der Ort des einzelnen Streuers durch den Vektor \mathbf{o} beschreiben lässt.

Für die Elektronenwellen aus einer atomaren Probe hatten wir in Abschnitt 3.1.2 aber bereits die kegelförmige Wellenform als treffender herausgestellt. Der damit entsprechend angepasste Propagationsoperator ist im `ConicWavefield` (B.1.3) implementiert. Für alle atomaren Streuer wird dort ein anpassbarer, konstanter Winkel Φ angenommen. Der Vektor zwischen dem Ort der Referenzwelle \mathbf{r}_0 und dem Ort eines jeden Objektpixels \mathbf{o} bildet dabei die Ausrichtung des Ausbreitungskegels. Anhand des Winkels zwi-

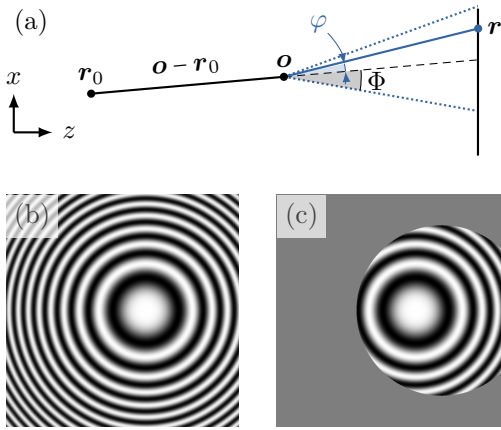


Abbildung 5.10: Ein typisches Einpunkt-Experiment veranschaulicht, was sich bei der Einführung der kegelförmigen Elektronenwelle ändert. (a) Konstruktionszeichnung zur Bestimmung, ob ein Detektorpixel \mathbf{r} aufgrund seines Winkels φ im Kegel der Elektronenwelle liegt und damit eine Wellenfeldinformation erhält. (b) Gemessene Hologrammintensität im Falle eines `Wavefield` (B.1.2) und (c) eines `ConicWavefield` (B.1.3).

schen ihm und dem relativen Ort der Auswertung

$$\begin{aligned} \varphi &:= \angle(\mathbf{o} - \mathbf{r}_0, \mathbf{r} - \mathbf{o}) \\ &= \arccos\left(\frac{(\mathbf{o} - \mathbf{r}_0) \cdot (\mathbf{r} - \mathbf{o})}{|\mathbf{o} - \mathbf{r}_0| \cdot |\mathbf{r} - \mathbf{o}|}\right) \end{aligned} \quad (5.15)$$

lässt sich, wie in Abbildung 5.10a gezeigt, bestimmen, ob der Pixel sich im Kegel befindet. Überschreitet dieser Winkel den maximalen Öffnungswinkel Φ , wird die Wellenfunktion eines Streuers des konischen Wellenfelds als null angenommen, ansonsten entspricht sie der einfachen Kugelwelle nach Gleichung (5.14). Mathematisch ausgedrückt bedeutet dies

$$\psi_C(\mathbf{r}) = \begin{cases} \psi_W(\mathbf{r}) & \text{wenn } |\varphi| \leq \Phi, \\ 0 & \text{sonst.} \end{cases} \quad (5.16)$$

Die Maskierung durch die kegelförmige Ausbreitung ist im direkten Vergleich in Abbildung 5.10 dargestellt.

Als weiteres relevantes Wellenphänomen ist die Referenzwelle als `Refwave` (B.1.1) sowohl als ebene Welle (2.14) wie auch als Kugelwelle (2.18) wählbar nach Abschnitt 2.3 implementiert. Sie benötigt ferner einen Wellenvektor \mathbf{k} , einen Referenzwellenursprung \mathbf{r}_0 und eine Ausgangsamplitude A_r . Beleuchtet sie in der Simulation ein Wellenfeld (also etwa ein zu holographierendes Objekt), werden neben den skalierten Amplituden ferner auch die Referenzwellendaten als Metadaten in das Wellenfeldobjekt übernommen. So erhalten alle folgenden Emissionen ihre Wellenlänge von der beleuchtenden Referenzwelle.

5.4.2 Koordinatentransformation

Da Wellenfelder im Sinne dieser Bibliothek nicht notwendigerweise in ihrer Ausrichtung parallel zu den kartesischen Koordinatenachsen liegen, ist eine Form der Rotation nötig. Dafür existiert in der Bibliothek das Objekt `Orientation` (B.5), das neben einer Rotation um die Koordinatenachsen in Form von Hamilton'schen Quaternionen, einer Rotationsmatrix oder Euler'schen Winkeln auch einen Bezugspunkt aufnimmt, um den die Rotation geschieht.

Die Rotation um den Ursprung lässt sich dabei mathematisch durch Rotationsmatrizen beschreiben. Für jede kartesische Achse lässt sich so eine Matrix \mathbf{R}_i zur Rotation um einen Winkel α_i nach

$$\begin{aligned} \mathbf{R}_x(\alpha_x) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha_x & \sin \alpha_x \\ 0 & -\sin \alpha_x & \cos \alpha_x \end{pmatrix}, \\ \mathbf{R}_y(\alpha_y) &= \begin{pmatrix} \cos \alpha_y & 0 & \sin \alpha_y \\ 0 & 1 & 0 \\ -\sin \alpha_y & 0 & \cos \alpha_y \end{pmatrix}, \\ \mathbf{R}_z(\alpha_z) &= \begin{pmatrix} \cos \alpha_z & -\sin \alpha_z & 0 \\ \sin \alpha_z & \cos \alpha_z & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned} \quad (5.17)$$

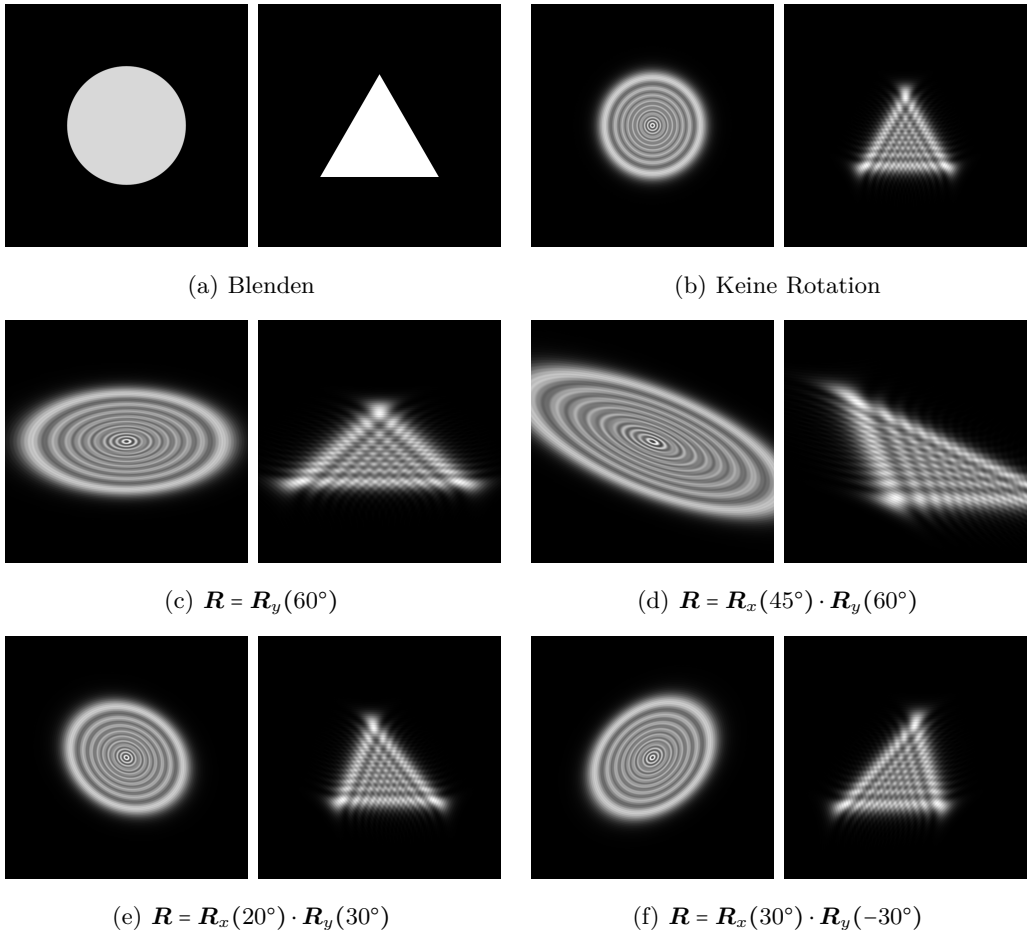


Abbildung 5.11: Beispiele der Aufnahme von Intensitätsmustern je einer kreis- und einer dreiecksförmigen Blende (a) auf Schirme mit einer Rotation um den jeweils eigenen Mittelpunkt (b)–(f) nach Beispielen aus Ref. [11, Abb. 9.11]. Alle Bilder haben eine Auflösung von $N^2 = 1024 \times 1024$ Pixeln, eine Pixelgröße von $\Delta x = \Delta y = 1 \mu\text{m}$ und wurden mithilfe simulierter HeNe-Laserwellen mit einer Wellenlänge von $\lambda = 633 \text{ nm}$ aufgenommen.

definieren [125, Kap. 4.4.2.1]. Die Gesamtrota-
tion wird dann als ein Produkt der Art

$$\mathbf{R} = \mathbf{R}_z(\alpha_z) \cdot \mathbf{R}_y(\alpha_y) \cdot \mathbf{R}_x(\alpha_x) \quad (5.18)$$

dieser Rotationsmatrizen beschrieben. Im
vorliegenden Falle wird die Rotation anhand
der Cardan-Winkel in der Reihenfolge um die
 x -, y - und zuletzt z -Achse durchgeführt. Da-
bei ist das Produkt im Allgemeinen nicht
kommutativ und die Angabe einer solchen
Rotationsmatrix auch nicht zwangsläufig ein-
deutig [125, Kap. 4.4.2.2].

Die Inverse der Rotationsmatrix lässt sich
als Transponierte oder durch leichte Winkel-
transformation gemäß

$$\mathbf{R}_i^{-1}(\alpha_i) = \mathbf{R}_i^\top(\alpha_i) = \mathbf{R}_i(-\alpha_i) \quad (5.19)$$

ausdrücken. Ferner sind die Rotationsmatrizen mit

$$\det \mathbf{R}_i(\alpha_i) \equiv 1 \quad (5.20)$$

normiert und damit orthogonal [11, Kap. 9].
Entsprechend ist auch \mathbf{R} eine orthogonale
Matrix mit

$$\mathbf{R}^{-1} = \mathbf{R}^\top. \quad (5.21)$$

Die Rotation lässt sich auf alle Koordinaten
mittels

$$\mathbf{x}' = \mathbf{R} \cdot \mathbf{x} \quad (5.22)$$

anwenden.

Allerdings findet die Rotation nun um den
Ursprung des Koordinatensystems statt. Da-
mit der Rotationspunkt frei gewählt werden
kann, wird ein Bezugspunkt \mathbf{c} definiert. Die-
ser wird durch die Bibliothek standardmäßig
als in der Mitte des Wellenfeldvolumens lie-
gend angenommen. Die Koordinatentransfor-
mation geschieht damit in drei Schritten. Zu-
nächst wird durch Subtraktion des Bezugs-
vektors \mathbf{c} dieser in den Ursprung transfor-
miert. Anschließend wird die Rotation gemäß
Gleichung (5.22) durchgeführt. Nun wird der

Bezugsvektor wieder auf alle rotierten Koor-
dinaten addiert, um das Wellenfeld wieder an
die ursprüngliche Position zu verschieben.

Ergebnisse einer Propagation nach Rotation
des empfangenden Schirms sind beispielhaft
in Abbildung 5.11 dargestellt. Um den Be-
schränkungen der Cardan-Rotation zu entge-
hen, ist die Rotation intern über die Rotati-
on mithilfe von Quaternionen implementiert.
Allerdings sind Cardan-Winkel und Quater-
nionen ineinander umrechenbar. Es sei dafür
auf Ref. [191, Kap. 8.12] verwiesen.

5.4.3 Kristallsampling

Zur Simulation eines atomaren Kristalls
steht die Bibliothek `crystals` (B.4) be-
reit, mit der sich aus einer Einheitszelle
`crystals.Crystal` (B.4.1) ein ausgedehnter
Kristall `crystals.SuperCell` (B.4.2) kon-
struieren lässt. Es ist ferner möglich, Fehl-
stellen in Form von Fremdatomen oder Lö-
chern einzubauen. Die so konstruierten Kri-
stalle lassen sich dank implementierter Kon-
verter direkt in ein gesampelttes Wellenfeld
überführen. An jedem Atomzentrum wird ein
Pixel generiert, dessen Anfangsamplitude ein-
em elementabhängigen Wert und damit der
Streuwahrscheinlichkeit a_n entspricht, wie sie
in Anhang A.1 beschrieben wird.

5.4.4 Rekonstruktion

Um zu überprüfen, inwieweit die atomare In-
formation im Hologramm vorhanden und wie-
derherstellbar ist, muss das Hologramm re-
konstruiert werden. Damit anhand der im Fol-
gekapitel noch zu spezifizierenden Kriterien
ein Vergleich dieser Rekonstruktionen mög-
lich ist, muss sie automatisiert am Computer
ablaufen.

Eine sehr einfache Methode für diese Rekon-
struktion ist die bereits in Abschnitt 4.3.2
beschriebene Rückpropagation [11, Kap. 13.2],
die auch hier implementiert wurde. Sie wird

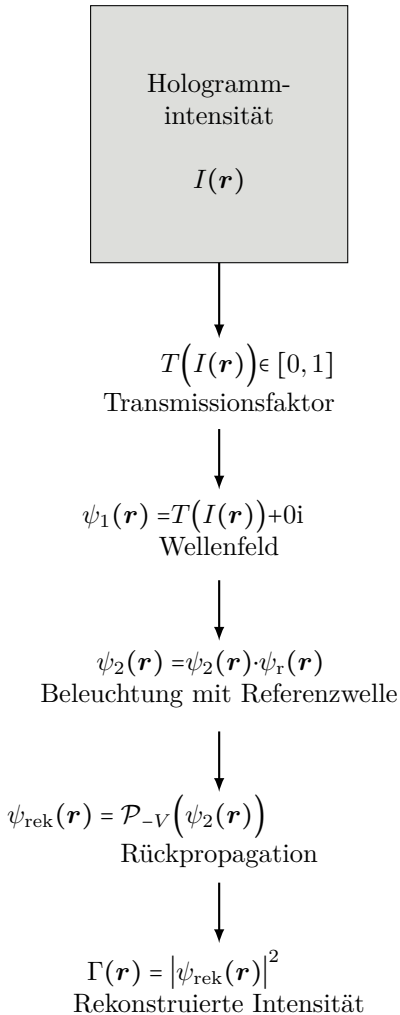


Abbildung 5.12: Schematische Darstellung des Rekonstruktionsprozesses, wie er durch den Algorithmus durchgeführt wird.

mathematisch völlig analog zur Propagation berechnet. Damit reicht die in diesem Kapitel ausführlich vorgenommene Betrachtung des Prozesses zur Simulation eines Hologramms aus, um auch die Rekonstruktion zu erschließen.

Schematisch ist die Rekonstruktion in Abbildung 5.12 dargestellt: Die Hologrammintensität nach Berücksichtigung der Transmissionsseigenschaften des virtuellen Schirms nach Gleichung (2.35) liegt bereits in korrekt gesampelter Form vor und wird in ein eigenes Wellenfeld exportiert. Die gleiche Referenzwelle wie zur Holographie oben beleuchtet auch das Hologramm. Anschließend lässt sich das Wellenfeld mit dem Invertierungsoperator markieren. Alle Folgepropagationen werden als Rückpropagation durchgeführt. Konkret ändert sich damit also das Vorzeichen im Exponenten innerhalb der Wellenfunktion, wie es in Gleichung (A.9a) gezeigt ist.

Im Anschluss wird zur Auswertung des Rekonstruktionsvolumens die Intensität im Gegensatz zum Hologramm nicht aus einem Beugungsmuster durch Überlagerung mit einer Referenzwelle gebildet, sondern es wird lediglich das Betragsquadrat der neuen Objektwelle gebildet. Das bei der Holographie als problematisch identifizierte Zwillingbild bei diesem grundsätzlich gleichen Prozess kann dabei nicht erneut auftreten. Es kam erst durch Addition mit der Referenzwelle als komplex konjugiertes Objektwellenfeld aus dem Betragsquadrat zustande [11, Kap. 13.2.1].

6 Untersuchungen

Um die Grundvoraussetzungen der Rekonstruierbarkeit mittels EBSD aufgenommener Hologramme zu überprüfen, benötigt es zunächst einen reproduzierbaren Versuchsaufbau und Methoden der Auswertung von Testmessungen. Im vorliegenden Kapitel soll dazu zunächst eine Kristallstruktur mit stets gleicher lokaler Umgebung vorgestellt werden, die sich für diese Untersuchungen anbietet. Im Anschluss wird der konkrete Aufbau unter Berücksichtigung der in Kapitel 3 beschriebenen Experimentalparameter erläutert.

Da ein großer Parameterraum untersucht werden soll, bedarf es automatisierter Auswertungswerkzeuge. Für die Bewertung des Erfolgs einer Rekonstruktion existieren jedoch nach Kenntnis des Autors keine standardisierten Gütekriterien. Gleichwohl erinnert die Einteilung in die Zustände „Atom“ oder „kein Atom“ an binäre Klassifikation, wie sie im Bereich des maschinellen Lernens schon lange etabliert ist [192, Kap. 2]. Im letzten Abschnitt sollen daher Kernkonzepte dieser Klassifizierung und ihrer Bewertung übernommen und auf das vorliegende Problem angepasst werden.

6.1 Probensystem

Für die Untersuchungen in den folgenden Kapiteln soll eine echte Kristallstruktur simuliert werden. Damit wird auch die praktische Bedeutung der zu erbringenden Erkenntnisse belegt. Im Rahmen dieser Arbeit soll dazu ein Wolframeinkristall Gegenstand der Simulation werden.

Wolfram ist ein selten auftretendes Metall von hoher technischer und wirtschaftlicher Bedeutung. Nach Kohlenstoff besitzt es den höchsten Schmelzpunkt aller Elemente und damit ebenfalls den höchsten aller Metalle bei 3410 °C [193, Kap. 27.2.1]. Innerhalb der

Gruppe der metallischen Materialien wird es darin ansonsten nur durch wenige Metallverbindungen, also Kombinationen von Elementen, wie Karbide einiger Schwermetalle und Mischkristalle in geringem Maße übertroffen [194, 195]. Gegenüber Kohlenstoff besitzt es dabei auch eine große Festigkeit. Kohlenstoff- und Tantalfäden sind häufig spröde, weshalb sich Wolfram besonders für den Einsatz in Hochtemperaturgebieten eignet und sich so früh als Leuchtmaterial beispielsweise in Glühbirnen und Röntgenröhren durchsetzen konnte. Ferner wird es aufgrund seiner Temperaturbeständigkeit in Schweißstäben, Thermoelementen oder wegen seiner großen Härte, ausreichender elektrischer Leitfähigkeit und guter Oxidationsbeständigkeit als Kontaktmaterial in der Elektrotechnik eingesetzt [196, Kap. 1].

Auf natürlichem Weg kommt Wolfram sehr selten vor, es macht etwa $5 \cdot 10^{-5} \%$ des magmatischen Gesteins der Erdkruste aus [197, Tab. 11]. Sein Vorkommen lässt sich mithilfe von Pneumatolyse erklären [198], einer Sonderform der Autometasomatose, bei der unter hohen Gasdrücken [199] eine stoffliche Umsetzung im Anschluss an die Auskristallisation des Gesteins stattfindet [200, Kap. 26.6]. Ferner tritt Wolfram in Meteoriten auf [201].

In der Erdkruste findet es sich dabei nur in Oxiden gebunden. Besonders häufig in der Form der Wolframite Ferberit (FeWO_4) und Hübnerit (MnWO_4), ferner als Scheelit (CaWO_4) und Stolzit (PbWO_4) [196, Kap. 2]. Die Gewinnung aus den geborgenen Erzen ist aufgrund der geringen Menge von lediglich 1%–3% Wolframoxid im Gestein sehr unwirtschaftlich. Daher muss es zunächst zu einer höheren Konzentration aufbereitet werden. Dafür existieren einige vielschrittige Verfahren. Beispielsweise eine magnetische Abscheidung als Trockenverfahren und ein opti-

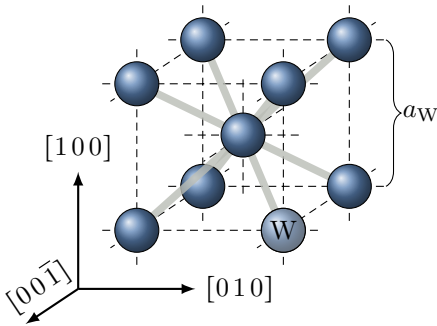
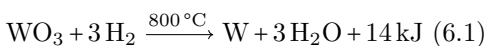


Abbildung 6.1: Aufbau einer Einheitszelle reinen Wolframs. Es handelt sich um ein kubisch raumzentriertes Gitter (bcc).

nal vorgelagertes Nassverfahren, etwa die vorübergehende Lagerung in einer Schwefelsäurelösung. Bei der anschließenden Flotation als nassmechanisches Verfahren [202, Kap. 1] wird die Eigenschaft der fein zerkleinerten Mineralien ausgenutzt, aus einer wässrigen Trübe bei Einsatz von Gasblasen durch Ankoppeln aufzuschwimmen [203, Kap. 4].

Durch Schmelzen mit Soda oder Zersetzung mithilfe anorganischer Säuren lassen sich nach Reinigung reine Wolframoxide isolieren [196, Kap. 3]. Im Anschluss lässt sich elementares Wolfram über Reduktion mit Wasserstoff bei 800 °C mittels



technisch als graues Pulver gewinnen [193, Kap. 27.2.1].

6.1.1 Kristallstruktur

Für die holographische Untersuchung betrachten wir im Rahmen der vorliegenden Arbeit rein periodische Kristalle. Daher sind für die experimentelle Forschung Einkristalle nötig. Sie sind weitestgehend homogen und weisen eine finite Anzahl gleicher lokaler Umgebungen und damit möglichst wenig Kristallite auf.

Im Fall von Wolfram lässt sich zunächst ein Einkristallfaden gewinnen, indem Wolframmetall unter Zusatz eines Bindemittels durch Diamantdüsen gespritzt wird. Der Faden wird mit hohen Temperaturen von 2400 °C–2600 °C nachbehandelt. Am noch warmen Kristall schlägt sich weiteres Wolfram aus einer Gasphase nieder, wodurch dann Kristalle beliebiger Größe angefertigt werden können [204].

Wolfram ist ein sehr hartes Metall. Angefertigte Proben lassen sich daher problemlos schrittweise bis zur feinsten Körnung hin schleifen und abschließend mithilfe von feinem Blei- oder Aluminiumoxid polieren. Eine völlig riefenfreie Oberfläche ist damit leicht realisierbar [196, Kap. 8.2].

Diese Wolframkristalle bilden ein kubisch raumzentriertes Gitter, nach dem englischen *body centered cubic* kurz auch bcc genannt, wie es in Abbildung 6.1 dargestellt ist. Die Gitterkonstante ist dabei in alle Raumrichtungen mit etwa

$$a_{\text{W}} = 3,165 \text{ \AA} \quad (6.2)$$

gleich [205]. Mit einem Atomgewicht von $m = (183,84 \pm 0,01) \text{ u}$ [206] lässt sich die Masendichte einer Einheitszelle, die nach Abbildung 6.1 zwei ganze Atome enthält, zu

$$\rho = \frac{2 \cdot m}{a_{\text{W}}^3} = 19,257 \text{ g cm}^{-3} \quad (6.3)$$

errechnen.

Der fortan angenommene Kristall hat eine rein kubische Form. Mithilfe Miller'scher Indizes lässt sich sein Schnitt als durch eine der $\{100\}$ -Ebenen verlaufend beschreiben, er verläuft also parallel zu einer der kartesischen Koordinatenebenen. Der Kristall wird mit seiner Ausrichtung daher als $\text{W}(100)$ bezeichnet.

Wolfram ist nicht nur hinsichtlich seiner dargestellten, herausragenden technischen Eigenschaften als Untersuchungsobjekt sehr relevant. Seine Festigkeit macht es leicht präpa-

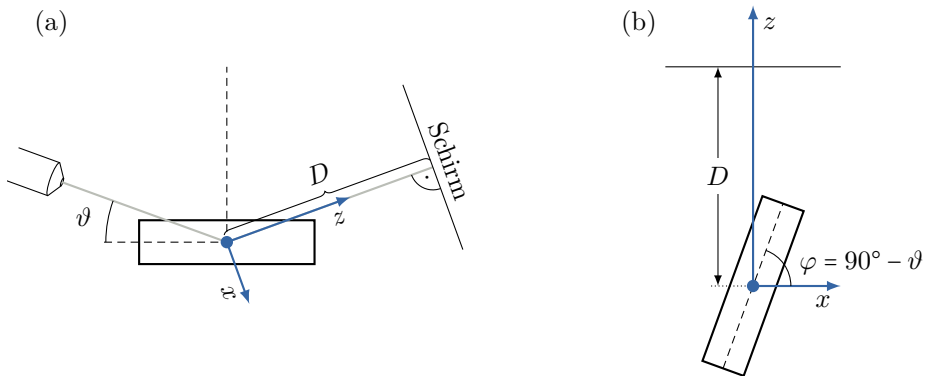


Abbildung 6.2: Der simulierte EBSD-Aufbau als (a) vereinfachte Skizze nach Abbildung 3.4 und (b) reduziert auf die für die Simulation relevanten Bestandteile mit transformiertem Koordinatensystem.

rierbar und als Probenmaterial unempfindlich. Insbesondere die Kristallstruktur birgt aber einen weiteren, entscheidenden Vorteil für die kommenden Untersuchungen: Jedes Atom besitzt exakt die gleiche lokale Umgebung. Damit reicht es nach den Ausführungen in Abschnitt 3.3 aus, ein einziges Hologramm aus einer lokalen Umgebung zu berechnen. Bei Einbeziehung der möglichen chemischen Sensitivität einer EBSD-Messung durch Energiefilter nach Ref. [97] trifft das auch auf einige weitere Kristalle zu. So verfügt Kochsalz (NaCl) zwar über eine zweiatomige Basis innerhalb eines Gitters mit kubisch flächenzentrierter Struktur (face centered cubic, kurz fcc) [193, Kap. 6.1.1.2], allerdings unterscheiden sich beide Basisatome in der Ordnungszahl und verfügen jeweils immer über die gleiche lokale Umgebung.

6.2 Aufbau

Der experimentelle Aufbau einer EBSD-Apparatur war bereits Gegenstand von Kapitel 3. Wie in Abbildung 3.4 dargestellt, besteht er im Wesentlichen aus einer Elektronenkanone, einem rotierten Kristall und einem Schirm. Nach Abschnitt 3.3 liegt für die Betrachtung der reinen Holographie deren Ausgangspunkt,

der Ursprung der Referenzwelle, aber innerhalb des Kristalls. Wie in Abbildung 3.5 dargestellt, nehmen wir den ersten Streuprozess als Quelle der Referenzwelle an. Für die vollumfängliche Simulation der EBSD-Holographie reicht folglich die Berücksichtigung von Kristall und Schirm aus. Die Elektronenkanone findet sich lediglich in der gewählten Wellenlänge wieder, da die an ihr eingestellte Elektronenstrahlenergie die Energie der ausgelösten Elektronenwellen vorgibt.

Für die Beschreibung des Experiments führen wir das in Abbildung 6.2 gezeigte Koordinatensystem ein. Der Ursprung liegt dabei im Kristall, der positioniert und rotiert wird. Der Schirm wird parallel zur neuen xy -Ebene in einem z -Abstand positioniert, der dem Proben-Schirm-Abstand D entspricht.

Der Schirm lässt sich nach Abbildung 6.3 durch seine Ausdehnung d und die Zahl der Pixel in einer Dimension N definieren. Er sei in der weiteren Beschreibung ohne Beschränkung der Allgemeinheit als symmetrisch in x - und y -Richtung angenommen. Aus diesen Parametern lassen sich auch die Pixelgrößen

$$\Delta x = \Delta y = N/d \quad (6.4)$$

berechnen.

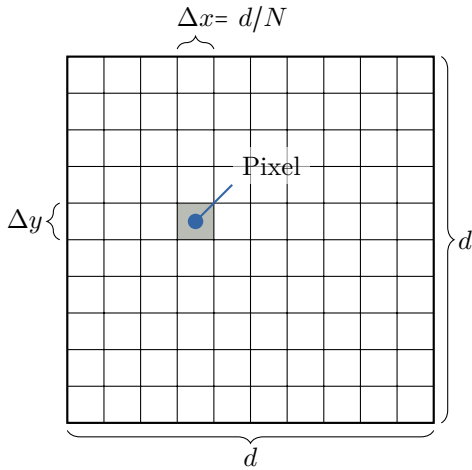


Abbildung 6.3: Der holographische Schirm ist durch seine Pixelzahl N und Ausdehnung d in jeweils einer Dimension definiert. Äquidistante Pixel messen ein Wellenfeld am Ort des Pixelzentrums.

6.2.1 Lokale Umgebung

Die Probe besteht aus einer Ansammlung von Punktstreuern, die jeweils den Kristallatomen entsprechen. Dabei wird keineswegs eine vollständige Probe simuliert. Bei einer als Würfel angenommenen Probe mit einer Kantenlänge von lediglich $l = 2$ cm müssten bereits

$$\left(\frac{l}{a_W}\right)^3 \approx 2,52 \cdot 10^{23} \quad (6.5)$$

Streuer simuliert werden. Das ist mit heutigen Methoden nicht leistbar. Allerdings ist es nach der Argumentation zur mittleren freien Weglänge und deren Bedeutung als Reichweite für die Definition der lokalen Umgebung in Abschnitt 3.3 auch gar nicht nötig. Da für die Berechnung eines Gesamthologramms nur die Simulation einer lokalen Umgebung erforderlich ist, reicht es aus, eine solche um das als Referenzwellenursprung angenommene Atom zu konstruieren. Sie ist nach Abschnitt 3.1.1 durch die Größe des IMFP begrenzt. Bei einer angenommenen kinetischen

Energie von $E_{\text{kin}} = 20$ keV ergibt sich die mittlere freie Weglänge nach Gleichung (3.2) zu $\lambda_{\text{IMFP}} \approx 158$ Å. Nach Division durch die Gitterkonstante a_W erhalten wir daraus einen Mindestradius für die lokale Umgebung. Der so simulierte Kristall soll aus Gründen der Einfachheit in der algorithmischen Beschreibung kubisch angenommen werden. Der Kristallgröße entspricht damit bei Berücksichtigung der zweiatomigen Basis nach Aufrundung eine Zahl von

$$N_{xyz} = 102 \times 102 \times 51 \times 2 \approx 106 \cdot 10^4 \quad (6.6)$$

Streuern. Da wir nach Abschnitt 3.1.2 von einer konischen Wellenausbreitung in Richtung Schirm ausgehen können, ist die Verdopplung der Kristallgröße nur in x - und y -Richtung nötig. Tieferliegende Atome haben keine Gelegenheit, ein Muster auf dem Schirm abzubilden, da ihre Kegel dem Schirm abgewandt sind. Ferner verhindert der IMFP in dieser Tiefe auch, dass die Elektronen die Oberfläche überhaupt erreichen.

Tatsächlich wirkt ein 51 Einheitszellen tiefer Kristall intuitiv sehr groß für eine Oberflächenuntersuchung. Durch den Probenwinkel φ ist bei gleichbleibendem IMFP in Strahlrichtung auch die senkrechte Tiefe zusätzlich begrenzt. Es stellt sich die Frage, ob dieses Volumen nicht möglichst klein angenommen werden sollte, da die Zeit für die Berechnung eines Hologramms mit zunehmender Kristalltiefe wächst. Ferner erfordern mehr angenommene Streuer auch eine höhere Schirmauflösung, denn zusätzliche Streuprozesse erzeugen zusätzliche feinere Beugungsmuster, die es aufzunehmen gilt.

Aus holographischen Experimenten mit sichtbarem Licht ist bekannt, dass das Einführen von mehr Partikeln in die holographische Szene die Rekonstruktion erschwert [207]. So zeigt sich bei der Verwendung von Partikeln mit einem Durchmesser von $d_p = 5 \mu\text{m} - 50 \mu\text{m}$ unter Beleuchtung mit einem Argonlaser der Wellenlänge $\lambda = 514,5$ nm, dass das Signal

Untergrund-Verhältnis primär antiproportional zu einer *Speckle Parameter* genannten Variable

$$G = \frac{\pi^3}{16} \cdot d_p^2 \cdot N_p \cdot L_V \quad (6.7)$$

ist [208]. Diese wiederum wird bestimmt von der Partikelanzahl N_p und der Probenvolumentiefe L_V . Bei Skalierung der Wellenlänge auf eine Elektronenwelle mit EBSD-typischer Energie von $E_{\text{kin}} = 20 \text{ keV}$ ergibt sich für solche Partikel ein Durchmesser von etwa $d_p = 0,84 \text{ \AA} - 8,43 \text{ \AA}$ und damit in der Größenordnung der Atomdurchmesser. Die Ergebnisse der Publikation Ref. [208] sind damit auf den vorliegenden Aufbau übertragbar.

Neben der damit wachsenden Laufzeit ist auch die Rekonstruktion bei zunehmender Kristalltiefe immer herausfordernder. Eine Vertiefung des Probenvolumens erhöht das Rauschen in der Rekonstruktion gegenüber dem Signal deutlich [209]. Im Experiment wird auch deutlich, dass sich zwischen Referenzwellenursprung und Schirm diverse Atomketten ausbilden, die eine stärkere Intensität in Richtung der Verbindungsachse hervorrufen [210]. Das unterdrückt Signale der Atome abseits dieser Linien, die zur Rekonstruktion ebenfalls nötig wären.

Grundsätzliches Ziel einer optimalen Simulation wäre somit die Reduzierung der Anzahl möglicher Streuer. Als Abschätzung soll an dieser Stelle aber dennoch der IMFP für die maximal mögliche Tiefe angenommen werden. Damit ließen sich die hier ermittelten Ergebnisse gesichert auf flachere Systeme übertragen und haben gesichert physikalische Bedeutung [20].

Die angenommene Probe ist schließlich durch ihre Orientierung definiert. Diese wird durch die in Abschnitt 5.4.2 festgelegten Parameter beschrieben. Ferner wird also auch ein Bezugspunkt und eine Verkippung zur Definition des Aufbaus benötigt.

6.2.2 Rekonstruktionsvolumen

Zur Rekonstruktion eines Hologramms ist es nicht zweckmäßig, die vollständige lokale Umgebung, wie sie oben eingeführt wurde, als Volumen für die Berechnung der rekonstruierten Wellenfunktion ψ_{rek} anzunehmen. In einem Rekonstruktionsvolumen müssen deutlich mehr Pixel als zuvor angenommene Streuer vorhanden sein, um zu überprüfen, ob der Rekonstruktionsvorgang glückt. Im Falle des real durchgeführten Experiments wäre die Information über die genaue Lage der Streuer a priori ferner erst gar nicht vorhanden. Die Pixelabstände sind deshalb möglichst klein zu wählen. Damit wird die Pixelanzahl auf einer definierten Länge aber sehr groß. Da jeder zusätzliche Pixel die Laufzeit verlängert, gilt es, einen Kompromiss zwischen der Ausdehnung und der Auflösung des Rekonstruktionsvolumens zu treffen. So würde eine Verdopplung der Pixelzahl in jede Dimension aufgrund des kubischen Volumens ansonsten eine Verachtfachung der Laufzeit nach sich ziehen.

In den nachfolgenden Untersuchungen wurde unter Abwägung der beschriebenen Einschränkungen eine volle Einheitszelle zuzüglich eines Randbereichs als Größe für das Rekonstruktionsvolumen gewählt, die Auflösung selbst kann je nach Experiment variieren. Dabei wurden immer äquidistante Pixel angenommen.

Es wäre durchaus möglich, von einer äquidistanten Pixelplatzierung abzusehen und diese stattdessen am Ort der ursprünglichen Streuer dichter zu setzen. Das würde eine allgemein geringere Auflösung erlauben, da in den Bereichen zwischen den Streuern kein Intensitätsauschlag zu erwarten ist. Allerdings geschieht damit bereits implizit eine Gewichtung: Mögliche Fehlerpixel treten aufgrund geringerer Zahl weniger auf, das macht sich in den Qualitätsbewertungen, wie sie in Abschnitt 6.3 noch thematisiert werden, durch besser bewertete Klassifikationen entgegen objektiver Kriterien bemerkbar. Eine solche

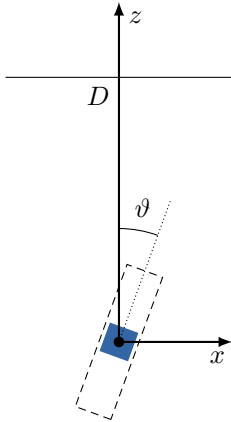


Abbildung 6.4: Lage des Rekonstruktionsvolumens innerhalb des zuvor simulierten Kristalls.

Wahl des Pixelabstands kann folglich lediglich für erste Schätzungen zur Überprüfung einer grundsätzlichen Wirksamkeit einer Rekonstruktionsmethode genutzt werden, darf jedoch keineswegs für allgemeine physikalische Aussagen angewandt werden.

Das Rekonstruktionsvolumen verfügt neben Ausdehnung und Auflösung auch über dieselbe Orientierung, mit der bereits die Probe beschrieben wurde. Der Schirm steht dann, wie in Abbildung 6.4 dargestellt, senkrecht auf der z -Achse.

6.3 Bewertungskriterien

Im Rahmen dieser Arbeit wurde ein Experiment konstruiert, das aus einem Satz von Anfangsparametern ein Ergebnis erzeugt. Hier ist das die rekonstruierte Wellenfunktion innerhalb des Rekonstruktionsvolumens über den Zwischenschritt eines Hologramms. Da im Zuge der Untersuchungen sehr viele dieser Volumen berechnet werden, bedarf es einer automatisierten Auswertung und damit einer algorithmischen „Bewertung“. Es gilt daher, ein objektives Maß der Genauigkeit der Übereinstimmung der Rekonstruktion zum Aus-

gangsbild – dem oben beschriebenen Wolframkristall – zu finden.

Eine Möglichkeit, ein solches Kriterium zu definieren, ist die Untersuchung der einzelnen rekonstruierten Pixelwerte und ihrer jeweiligen Koordinaten. Für jeden Pixel lässt sich die Frage danach, ob hier ein Atom vorliegt oder nicht, binär, also mit genau einer von zwei Möglichkeiten, beantworten. Ein verwandtes Problem ist die im maschinellen Lernen sehr geläufige Klassifikation. Beim Arbeiten mit den in Abschnitt 5.2 bereits erwähnten neuronalen Netzen, muss dem lernenden Algorithmus während der Testphase zurückgemeldet werden, ob er mit seinem berechneten Ergebnis richtig liegt.

Ein gängiges Problem des maschinellen Lernens mit binärem Ausgang ist beispielsweise die Klassifizierung von E-Mails als Spam. Schon lange sind fest programmierte, regelbasierte Systeme auf Basis von regulären Ausdrücken nicht mehr ausreichend und es braucht künstliche Intelligenz, um auf sich stets verändernde Bedrohungslagen zu reagieren. Die Klassifizierung kann dabei grundsätzlich auch in mehr als nur zwei Klassen erfolgen, wenn beispielsweise E-Mails zusätzlich nach beruflich und persönlich kategorisiert werden sollen. Damit werden aber die Grenzbedingungen weniger klar und das Problem deutlich komplexer [192, Kap. 1.1]. Es sei hier am Rande erwähnt, dass die Klassifizierung nicht zwangsläufig binär erfolgen muss, da sich diese allgemeine Form in den folgenden mathematischen Beschreibungen wiederfindet.

Die Grundprinzipien dieser Klassifizierung sollen im Folgenden schrittweise erläutert werden, da sie für das Verständnis der später eingeführten Bewertungskriterien von elementarer Bedeutung sind. Es sei bereits hier aber festgestellt: Da wir keine selbstlernenden Algorithmen anwenden, ist das Ziel einer solchen Bewertung nicht bedeutungsgleich und muss adaptiert werden. Im Kontext des hier vorliegenden Problems geht es ausdrücklich

nicht um die Bewertung eines Algorithmus, der ja nach den obigen Ausführungen lediglich die beschriebene Mathematik in Form eines Propagationsoperators umsetzt. Anhand eines Satzes von Anfangsparametern, die veränderlich sind, lässt sich mithilfe dieser Klassifizierung stattdessen bewerten, inwieweit eine Rekonstruktion unter den gegebenen Anfangsparametern möglich ist. Wir stellen fest, ob der ermittelte Kristall nach Rekonstruktion mit dem simulierten Kristall übereinstimmt. Es werden eher gewählte Parameter anstatt des eigentlichen Prozesses bewertet. Es ist daher besonders darauf zu achten, die Kernkonzepte so zu portieren, dass sie schlüssig in ein eindeutiges Bewertungskriterium münden.

6.3.1 Klassifizierung

Anhand von Ref. [192, Kap. 2] wollen wir zunächst einige grundlegende Begrifflichkeiten einführen. Dazu definieren wir eine Menge von Instanzen \mathcal{X} . Bei jedem $x \in \mathcal{X}$ handelt es sich um einen theoretisch möglichen Pixel. Eine Labelmenge \mathcal{L} umfasst alle möglichen Bezeichnungen von Instanzen. Sie entspricht in der Regel der Klassifizierungsmenge

$$\mathcal{C} = \{C_1, C_2, \dots, C_k\}, \quad (6.8)$$

die die möglichen Ausgänge des Algorithmus spezifiziert.

Im binären Fall klassifiziert der Algorithmus in einen von zwei möglichen Ausgängen, die häufig als *positiv* (engl. positive) und *negativ* (engl. negative) bezeichnet werden. Ein rekonstruierter Pixel könnte als Atom (positiv) oder eben als kein Atom (negativ) erkannt werden. Die binäre Klassifizierungsmenge lautet damit

$$\mathcal{C}_b = \mathcal{L}_b = \{\oplus, \ominus\}. \quad (6.9)$$

Da wir den holographierten Kristall a priori kennen, können wir eine Labelfunktion

$$l: \mathcal{X} \rightarrow \mathcal{L}_b, \quad x \mapsto l(x) \quad (6.10)$$

definieren, die genau dann \oplus ist, wenn der Ort des Pixels x mit dem eines Atoms übereinstimmt. Ansonsten ist $l(x) = \ominus$. Im Rahmen des maschinellen Lernens werden daraus sog. *gelabelte Instanzen* $(x, l(x))$ erzeugt, die als Trainingsdaten verwendet werden können.

Der Rekonstruktionsalgorithmus selber setzt dabei eine Klassifizierungsfunktion

$$\hat{c}: \mathcal{X} \rightarrow \mathcal{C}, \quad x \mapsto \hat{c}(x) \quad (6.11)$$

um. Durch die berechnete Rekonstruktion erhalten wir zu jedem Pixel

$$x \in \mathcal{S} \subset \mathcal{X} \quad (6.12)$$

des Schirms \mathcal{S} die Intensität der rekonstruierten Welle $\Gamma(x)$ nach Gleichung (A.12). Gegebenenfalls wird diese Funktion aufbereitet, indem etwa die Dominanz der Referenzwelle durch Division unterdrückt wird. Aus ihr definieren wir nun eine normierte Intensität

$$\Gamma_n(x) := \Gamma(x) / \max_{x \in \mathcal{S}} (\Gamma(x)), \quad (6.13)$$

die jedem Pixel einen Wert $\Gamma_n(x) \in [0, 1]$ zuordnet. Die ihr entsprechenden, tatsächlichen Amplitudenwerte des Ausgangskristalls sind nach Anhang A.3 abhängig von den ursprünglichen Streuamplituden. Bei einem uniformen Kristall mit jeweils gleicher Streuwahrscheinlichkeit wird die normierte Intensität (6.13) also für alle Pixel an Atompositionen nahe 1 sein, andernfalls gegen 0 gehen. Durch statistische Prozesse und aus einem in der Realität niemals unendlich großen Oversampling Ratio σ nach Gleichung (4.54) werden potentiell auch viele Zwischenwerte erreicht.

Die Klassifizierungsfunktion $\hat{c}(x)$, die die Instanzen x in Atom \oplus oder kein Atom \ominus einordnet, muss dies berücksichtigen. Eine naheliegende Möglichkeit ist, anhand eines problemangepassten Thresholds

$$t_s \in]0, 1] \quad (6.14)$$

größer 0 zu entscheiden. Liegt ein Pixelwert oberhalb dieses Schwellwerts, wird an seiner

Stelle ein Atom angenommen, andernfalls nicht. Die Klassifizierungsfunktion für das vorliegende Problem lautet dann

$$\hat{c}(x) = \begin{cases} \oplus & \Gamma_n(x) \geq t_s, \\ \ominus & \Gamma_n(x) < t_s. \end{cases} \quad (6.15)$$

Wir erhalten damit zu jedem Pixel eine binäre Antwort darüber, ob an dieser Stelle ein Atom vom Algorithmus „gefunden“ wurde oder nicht. Überlagern wir hier die echte, vorab simulierte Kristallstruktur, werden wir feststellen, dass an einigen Stellen eine tatsächliche Übereinstimmung vorliegt, an anderer Stelle aber fälschlich ein Atom erkannt wurde oder eines der zuvor vorhandenen Atome fehlt.

Zur weiteren Formalisierung führen wir die Indikatorfunktion $I[\cdot]$ ein. Sie ist 1, wenn ihr Argument wahr ist und sonst 0. Mit ihr lässt sich eine Zählung durchführen und die *Korrektklassifikationsrate* (engl. accuracy)

$$= \frac{1}{|\mathcal{S}|} \sum_{x \in \mathcal{S}} I[\hat{c}(x) = l(x)] \quad (6.16)$$

definieren. Dabei handelt es sich um den Anteil aller Schirmpixel, deren Klasse mit dem tatsächlichen Label übereinstimmt. Ist diese Rate 1, so wurden nach der Rekonstruktion alle Atome exakt gefunden und kein Atom an falscher Stelle entdeckt. Sie lässt sich äquivalent auch als die Wahrscheinlichkeit

$$acc = P(\hat{c}(x) = l(x)) \quad (6.17)$$

dafür verstehen, dass eine beliebige Instanz $x \in \mathcal{X}$ korrekt klassifiziert wurde.

Analog dazu lässt sich auch eine *Falschklassifikationsrate* (engl. error rate)

$$\begin{aligned} err &= \frac{1}{|\mathcal{S}|} \sum_{x \in \mathcal{S}} I[\hat{c}(x) \neq l(x)] \\ &= P(\hat{c}(x) \neq l(x)) = 1 - acc \end{aligned} \quad (6.18)$$

als Wahrscheinlichkeit dafür definieren, dass ein Pixel falsch klassifiziert wurde.

Um ein Gütekriterium für die Übereinstimmung von Kristall und Rekonstruktion herleiten zu können, ist zunächst die Einführung weiterer Begrifflichkeiten nötig. Dazu soll die Anzahl der tatsächlich im Rekonstruktionsvolumen vorhandenen, also *positiven* Atome mit

$$\mathcal{P} = \sum_{x \in \mathcal{S}} I[l(x) = \oplus] \quad (6.19)$$

bezeichnet werden. Analog dazu lässt sich die Zahl der Pixel, an denen kein Atom vorhanden ist, durch die Zahl aller *Negativen*

$$\mathcal{N} = \sum_{x \in \mathcal{S}} I[l(x) = \ominus] = |\mathcal{S}| - Pos \quad (6.20)$$

definieren.

6.3.2 Wahrheitsmatrix

Nach Anwendung der Klassifizierungsfunktion $\hat{c}(x)$ nach Gleichung (6.15) lassen sich die Pixel x in eine von vier Gruppen einteilen. Ist nach der Rekonstruktion ein Atom erkannt worden, das auch tatsächlich an dieser Stelle simuliert wurde, handelt es sich um sog. *richtig positive* (engl. true positive) Ergebnisse, deren Anzahl durch

$$\mathcal{TP} = \sum_{x \in \mathcal{S}} I[\hat{c}(x) = l(x) = \oplus] \quad (6.21a)$$

gegeben ist. Wenn der Algorithmus aber ein Atom rekonstruiert hat, an dem eigentlich keines vorhanden sein sollte, so wird von einem *falsch positiven* (engl. false positive) Ergebnis gesprochen. Die Anzahl dieser Ergebnisse berechnet sich mittels

$$\mathcal{FP} = \sum_{x \in \mathcal{S}} I[\hat{c}(x) = \oplus \neq l(x) = \ominus]. \quad (6.21b)$$

Gleiches lässt sich für Pixel definieren, die nach Rekonstruktion keinen Hinweis auf ein Atom in sich tragen. Entspricht das den simulierten Tatsachen, handelt es sich um ein *richtig negatives* (engl. true negative) Ergebnis. Sollte aber eigentlich an dieser Stelle sehr

wohl ein Atom vorhanden sein, ist von einem *falsch negativen* (engl. false negative) Ergebnis zu sprechen. Deren Anzahl lässt sich jeweils analog zu oben mit

$$TN = \sum_{x \in S} I[\hat{c}(x) = l(x) = \ominus] \quad (6.21c)$$

$$= Neg - FP,$$

$$FN = \sum_{x \in S} I[\hat{c}(x) = \ominus \neq l(x) = \oplus] \quad (6.21d)$$

$$= Pos - FN,$$

bestimmen. Zusammenfassen lassen sich diese Größen in der sog. Wahrheits- oder Konfusionsmatrix (engl. confusion matrix), wie sie Abbildung 6.5 zeigt [211].

Der Begriffsbestandteil positiv oder negativ bezeichnet in diesen vier Fällen – entgegen der Benennung der reinen Anzahlen an Positiven oder Negativen oben – den Ausgang der Rekonstruktionsberechnung, anders ausgedrückt die Prognose über die Existenz eines Atoms. Der Vorsatz richtig bzw. falsch ergänzt dann, ob diese Klassifizierung auch der Wahrheit entspricht, der Algorithmus also ein respektive richtiges oder falsches Ergebnis geliefert hat.

Die reine Zahl aller *a priori* bekannter Atome und Stellen ohne Atom lassen sich damit respektive vereinfacht auch mit

$$Pos = TP + FN \quad \text{und} \quad (6.22a)$$

$$Neg = TN + FP \quad (6.22b)$$

ausdrücken.

6.3.3 Gütekriterien

Mit diesen Größen lassen sich nun weitere Kriterien zur Klassifikationsbewertung definieren. Die *Sensitivität*

$$tpr = \frac{TP}{Pos} = P(\hat{c}(x) = \oplus \mid l(x) = \oplus) \quad (6.23)$$

gibt die Wahrscheinlichkeit dafür an, dass ein beliebiges Atom auch tatsächlich als solches

		Rekonstruiert	
		\ominus	\oplus
Atom	\ominus	richtig negativ (<i>TN</i>)	falsch positiv (<i>FP</i>)
	\oplus	falsch negativ (<i>FN</i>)	richtig positiv (<i>TN</i>)

Abbildung 6.5: Die Wahrheits- oder Konfusionsmatrix fasst die vier Ergebnisgruppen zusammen. Angepasste Darstellung nach Ref. [211, Tab. 1].

erkannt wurde. In der Literatur wird diese Größe auch häufig Richtig-Positiv-Rate (engl. true positive rate), Empfindlichkeit, Trefferquote oder Recall genannt.

Analog dazu lässt sich die *Spezifität* genannte Richtig-Negativ-Rate (engl. true negative rate)

$$tnr = \frac{TN}{Neg} = P(\hat{c}(x) = \ominus \mid l(x) = \ominus) \quad (6.24)$$

ermitteln. Sie macht Angaben über die Wahrscheinlichkeit, mit der an einer Stelle, an der sich kein Atom befindet, auch nach der Rekonstruktion keines gefunden wurde. Beide Größen geben gemeinsam Auskunft über die Genauigkeit der Klassifizierung des Algorithmus in jeweils eine der beiden Klassen an [211].

Die dem entgegengesetzten Fehlerraten der beiden Klassen sind durch die *Falsch-Negativ-Rate* (engl. false negative rate) oder Miss Rate

$$fnr = \frac{FN}{Pos} = P(\hat{c}(x) = \ominus \mid l(x) = \oplus) \quad (6.25)$$

als Wahrscheinlichkeit eines nicht erkannten Atoms und die analoge *Falsch-Positiv-Rate* (engl. false positive rate), Fallout oder Ausfallrate

$$fpr = \frac{FP}{Neg} = P(\hat{c}(x) = \oplus \mid l(x) = \ominus) \quad (6.26)$$

als Wahrscheinlichkeit für ein fälschlich erkanntes Atom an leerer Stelle gegeben. Letztere entspricht damit der Wahrscheinlichkeit für einen „Fehlalarm“.

Die zuvor definierte Korrekturklassifikationsrate (6.16) lässt sich mit diesen Größen zu

$$acc = \frac{Pos}{|S|} \cdot tpr + \frac{Neg}{|S|} \cdot tnr = \frac{TP + TN}{Pos + Neg} \quad (6.27)$$

umformulieren.

Ferner ergibt sich die *Genauigkeit*

$$\begin{aligned} ppv &= \frac{TP}{TP + FP} \\ &= P(l(x) = \oplus \mid \hat{c}(x) = \oplus) \end{aligned} \quad (6.28)$$

als Anteil der tatsächlich existenten Atome unter allen Pixeln, die nach Meinung der Klassifizierung ein Atom darstellen sollten [211]. Sie wird häufig auch positiver Vorhersagewert oder Relevanz genannt und im Englischen mit Precision oder Positive Predictive Value bezeichnet. Zu beachten ist, dass die bedingte Wahrscheinlichkeit der Genauigkeit genau umgekehrt zu der der zuvor definierten Sensitivität (6.23) notiert ist. Während dort der Anteil der positiven klassifizierten unter den simulierten Atomen (also die tatsächlich gefundenen) ermittelt wurde, bestimmt der positive Vorhersagewert, welche der vermeintlich gefundenen Atome auch tatsächlich eine simulierte Entsprechung haben.

Alle bereits genannten, im Folgenden noch erwähnten und darüber hinaus damit verwandten und relevanten Größen zur (binären) Klassifikation finden sich als Referenz übersichtlich in Anhang C dargestellt.

6.3.4 Atomkugeln

Bei der Anwendung dieser Klassifizierung auf rekonstruierte Pixelwerte, aus denen Atompositionen ermittelt werden sollen, wird deutlich, dass es schnell zu einer Uneindeutigkeit kommen kann. Es wird nicht ausschließlich der Pixel als Atom klassifiziert, der exakt an der Stelle des simulierten oder im Experiment real vorhandenen Ausgangatoms liegt. Das ist insbesondere nicht möglich und auch nicht

wünschenswert, da der Rekonstruktionspixel nicht an der exakt selben Stelle des eigentlich simulierten Atoms liegen muss. Es wäre eine unendlich hohe Rekonstruktionsauflösung nötig, um bei Erfordernis einer exakten Übereinstimmung eine Rekonstruktion zu ermöglichen.

In der Realität weicht die Rekonstruktion kugelförmig um die Position des Atoms auf. Bei einem geringen Oversampling Ratio σ nach Gleichung (4.54) deutet eher eine Häufung positiver Pixel an einer Stelle auf ein Atom hin. Bei hohem Ratio zeichnet sich das Atom hingegen schärfer ab. Abhängig ist das auch von der Lage des Atoms, etwa der Entfernung zum Referenzwellenursprung oder der Tatsache, ob das Atom innerhalb einer der in Abschnitt 6.2.1 beschriebenen Ketten liegt. Es ist also keine allgemeingültige Regel ableitbar, die beispielsweise eine gewisse Anzahl an Pixeln voraussetzt, um ein Atom als eindeutig erkannt herauszustellen.

Aus diesem Grund definieren wir einen Kugelradius r_a . Mit dessen Hilfe werden Atomkugeln $\mathcal{A}_i \subset \mathcal{S}$ um die Position der simulierten Atome gelegt. Im Rekonstruktionsvolumen definieren wir nun drei Größen neu. Als richtig positiv gemäß der Definition

$$TP = \sum_i \Theta_0 \left(\sum_{x \in \mathcal{A}_i} I[\hat{c}(x) = \oplus] \right) \quad (6.29)$$

wird nun eine dieser Atomkugeln gezählt, die mindestens einen zu positiv ausgewerteten Pixel enthält. Die dabei verwendete, allgemeine Heaviside-Funktion $\Theta_k(x)$ ist definiert als

$$\Theta_k(x) = \begin{cases} 0 & x < 0, \\ k & x = 0, \\ 1 & x > 0. \end{cases} \quad (6.30)$$

Als falsch negativ wird weiterhin eine Atomkugel bewertet, in der sich gemäß

$$FN = \sum_i I \left[\left(\sum_{x \in \mathcal{A}_i} I[\hat{c}(x) = \oplus] \right) = 0 \right] \quad (6.31)$$

kein einziger positiver Pixel befindet. Als falsch positiv

$$FP = \sum_{x \in \mathcal{S}} I[\hat{c}(x) = \oplus \wedge x \notin \mathcal{A}_i \forall i] \quad (6.32)$$

wird nun einschränkend nur noch jeder positiver Pixel bezeichnet, der sich außerhalb einer der Kugeln \mathcal{A}_i befindet.

Dabei wäre denkbar, um so erkannte Atome auch eine Kugel mit Radius r_a zu konstruieren, in deren Inneren dann weitere Pixel nicht zur Erhöhung der *FP*-Zahl beitragen. Das würde die Größenordnungen an *FP* und *TP* innerhalb eines Rekonstruktionsvolumens angleichen, bereits wenige *FP*-Pixel reichen sonst aus, um deren Zahl gegenüber der *TP* stark zu erhöhen. Im oben skizzierten Rekonstruktionsvolumen etwa ist eine vollständige Wolfram-Einheitszelle enthalten, die gerade einmal neun Atome und damit maximal $TP = 9$ aufweisen kann. Dem stehen häufig etwa $N_r^3 = 100^3 = 1 \cdot 10^6$ Rekonstruktionspixel und damit annähernd so viele mögliche *FP* gegenüber. Andererseits wirft eine Definition über Kugeln gleich mehrere Fragen auf: Anhand welchen Kriteriums wird ein Pixel als Zentrum einer solchen Kugel definiert? Ist sie so zu legen, dass die Zahl der *FP* möglichst minimiert wird? Derartige Restriktionen erhöhen den Rechenaufwand in der Auswertung enorm, ohne zusätzlichen Informationsgewinn einzubringen. Aus diesem Grund soll darauf im weiteren Verlauf verzichtet werden.

Was implizit nun bereits geschieht, und bei der Wahl des konkreten Kriteriums noch the-

matisiert werden soll, ist eine Gewichtung: Eine Rekonstruktion, die durch etwa einen hohen Threshold t_s wenig *FP* erkennt, dafür aber auch weniger *TP*, wird besser bewertet als eine Lösung, die alle *TP* erkennt, dabei als Kollateralschaden aber auch viele *FP* zulässt. Das ist im vorliegenden Fall aber durchaus erwünscht: Ein Algorithmus, der generös klassifiziert und alle *TP* auf Kosten eines sehr vollen Rekonstruktionsvolumens findet, hilft bei der automatisierten Auswertung nicht weiter und sollte daher stärker bestraft werden.

Ferner wird deutlich, dass die Definition von richtig negativen Ergebnissen sehr schwierig anzupassen ist. So könnten alle negativen Pixel außerhalb der Atomkugeln so klassifiziert werden. Damit würden aber je nach Wahl des Gütekriteriums Rekonstruktionen bestraft, die auch innerhalb der Atomkugeln sehr präzise im Zentrum ein Atom nachweisbar machen und viel Leerraum innerhalb der Atomkugel aufweisen. Eine Gewichtung nach Abstand vom Atomzentrum würde dies berücksichtigen, verkompliziert die Bewertung aber zusätzlich. Daher soll an dieser Stelle auf eine neue Definition der *TN* und auf ihre Verwendung im Folgenden ebenfalls verzichtet werden.

Nachdem die Grundideen der binären Klassifikation damit auf das Problem der Rekonstruktionsbewertung übertragen wurden, kann im nächsten Kapitel weitergehend ein objektives Kriterium für die Bestimmung der Qualität einer Rekonstruktion erarbeitet werden.

7 Ergebnisse

Nachdem in Kapitel 2 die theoretischen Grundlagen für die Holographie, in Kapitel 3 die physikalischen Prozesse der Elektronenholographie und in Kapitel 4 die Besonderheiten innerhalb der Simulation erläutert wurden, sollen im Folgenden die Ergebnisse vorgestellt werden, die durch die in Kapitel 5 beschriebenen Algorithmen ermittelt wurden. Der konkrete experimentelle Simulationsaufbau wurde dabei schon in Kapitel 6 beschrieben, sodass hier die Simulationsergebnisse ausgewertet und eingeordnet werden können.

Dazu soll zunächst ein Standardexperiment durchgeführt werden, mit dessen Hilfe die Datenaufbereitung und insbesondere ein Bewertungskriterium für die Rekonstruktion eingeführt werden kann. Im Anschluss sollen anhand dieser Kriterien verschiedene Schirmparameter automatisiert untersucht werden, um die Realisierbarkeit eines elektronenholographischen Aufbaus mittels EBSD zu überprüfen. Zur Stabilisierung experimenteller Untersuchungen soll abschließend die Addition von Mustern verschiedener Energien motiviert und qualitativ untersucht werden.

7.1 Standardexperiment

Anhand eines konkreten Beispiels soll dazu zunächst deutlich gemacht werden, wie der gesamte Prozess am Computer konkret abläuft. Es sei hierzu ein Experiment mit Standardparametern angenommen. Nach Abschnitt 6.2 wird das Experiment vornehmlich durch eine Schirmgröße von $d = 5\text{ cm}$ und eine Pixelzahl je Dimension von $N = 1000$ beschrieben. Der Abstand zwischen Schirm und Probe wird in Übereinstimmung mit den in Kapitel 3 beschriebenen experimentellen Aufbauten für alle folgenden Simulationen als $D = 2\text{ cm}$ angenommen. Die kinetische Ener-

gie der Elektronen betrage ebenfalls übliche $E_{\text{kin}} = 20\text{ keV}$.

Für die Simulation wurde der in Abschnitt 6.1.1 beschriebene Wolframeinkristall $W(100)$ genutzt. Er ist schematisch in Abbildung 7.1a dargestellt und besitzt in seiner Gesamtheit die in Abschnitt 6.2.1 erläuterte Ausdehnung von

$$N = 102 \times 102 \times 51 \quad (7.1)$$

einzelnen Einheitszellen. Die Basis ist zweiatomig. Insgesamt sind also etwa 10^6 Streuer in einer lokalen Umgebung zu berücksichtigen. Als Referenzwellenquelle wird ein in x - und y -Richtung zentral und in z -Richtung möglichst tief liegender Streuer ausgewiesen. Die lokale Umgebung muss in jede Raumrichtung um die durch den IMFP bestimmte Größe von 51 Einheitszellen ausgedehnt sein. Lediglich die negative z -Richtung zum Kristallinneren hin kann aufgrund der konischen Wellenausbreitung in Richtung Schirm entfallen.

Die Referenzwellenquelle dient nun als Ausgang einer Kugelwelle mit einer nach de-Broglie bestimmten Wellenlänge von

$$\lambda \stackrel{(3.1)}{=} 8,67\text{ pm} \approx 0,09\text{ \AA}. \quad (7.2)$$

An jedem Streuer wird eine eigene Welle ausgelöst, deren Form nach Abschnitt 3.1.2 ein Kegel sein muss. Um den zu konstruieren, wird zwischen Referenzwellenquelle und jedem einzelnen Streuer je eine virtuelle Verbindungsachse gezogen und gedanklich bis zum Schirm fortgesetzt. Am Ort der Streuer wird dann jeweils ein Kegel mit Winkel Φ um diese Achse in Richtung des Schirms geöffnet. Damit lässt sich nun die Randbedingung der Nullamplitude abseits des Öffnungswinkels Φ gemäß Gleichung (5.16) berücksichtigen:

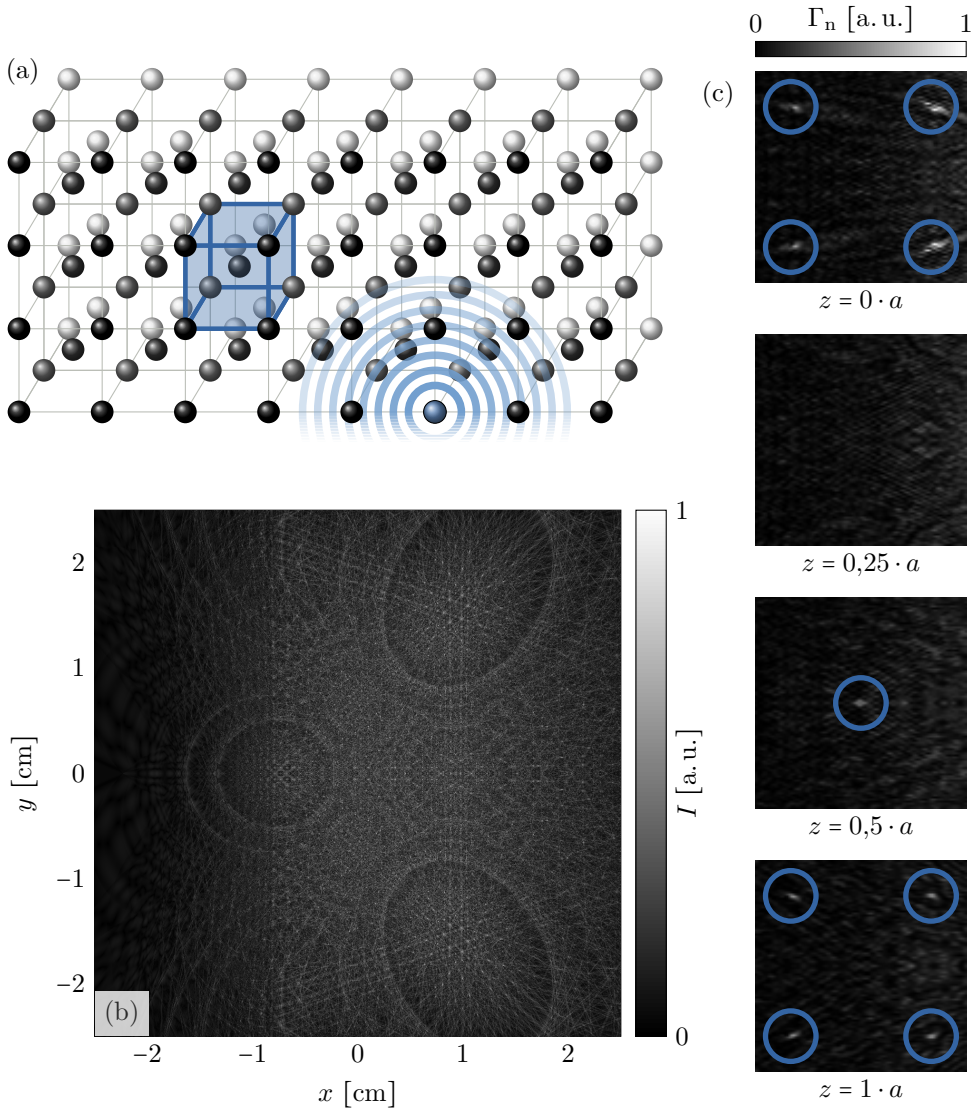


Abbildung 7.1: Beispielholographie und -rekonstruktion bei $E_{\text{kin}} = 20 \text{ keV}$, $d = 5 \text{ cm}$ und $N = 1000$. (a) Der Wolframeinkristall wird hier in einem kleineren Ausschnitt gezeigt. Hervorgehoben ist die Referenzwellenquelle und die später rekonstruierte Einheitszelle. Theoretisch ist dabei die Rekonstruktion jeder Einheitszelle im Kristall denkbar, nah an der Referenzwellenquelle liegende erzeugen aber größere Signale und sind damit deutlicher gegenüber dem Untergrund erkennbar. (b) Ein vollständiger Kristall erzeugt ein Hologramm, auf dem auch deutlich der Schattenwurf durch die Verkippung des Kristalls gegenüber dem Schirm um $\varphi = 70^\circ$ erkennbar ist. Der Bildkontrast wurde im Nachgang gemäß der Vorschrift $\sqrt[3]{I}$ angepasst, um die Mustermodulation deutlicher erkennbar zu machen. (c) Das Rekonstruktionsvolumen lässt sich ebenenweise analysieren. Hervorgehoben sind die Stellen, an denen ein Atom simuliert wurde. Eine gute Übereinstimmung mit der normierten Rekonstruktionsintensität Γ_n ist erkennbar.

Innerhalb des Kegels wird eine isotrope Welle angenommen, außerhalb verschwindet die Amplitude. Die ausgehende Welle besitzt die gleiche Wellenlänge (7.2) der Referenzwelle.

Nach Summation aller dieser komplexwertigen, kegelförmigen Sekundärwellen und der primären Referenzwelle am Schirm wird ein Betragsquadrat gebildet. Die auf diese Weise bestimmte Intensität I (A.7) wird als Hologramm bezeichnet und ist in Abbildung 7.1b dargestellt.

Vor der Rekonstruktion wird das Hologramm durch Normierung gemäß Gleichung (2.35) aufbereitet. Jedem Schirmpixel wird damit ein reeller Wert T zwischen 0 und 1 zugeordnet. Dieses normierte Hologramm wird, wie in Abschnitt 2.3.1 beschrieben, mit derselben Referenzwelle beleuchtet. Wie die Schwärzung einer Fotoplatte das durchdringende Licht abschwächt, stellen hier die Pixelwerte $T \in [0, 1]$ Skalierungsfaktoren für die Referenzwellenamplitude dar.

An jedem Ort wird die Referenzwellenamplitude bestimmt, skaliert und eine neue Partikularwelle mit der ermittelten Amplitude ausgestrahlt. Diese wird aber als sphärische Welle angenommen. Da die Amplitude außerhalb der zuvor angenommenen Kegel 0 betrug, trägt jeder Pixel nur die Abbildungsinformation einer isotropen Welle in sich. Es bedarf daher keiner gesonderten Behandlung zum Eliminieren dieser Unterscheidung zwischen Pixel innerhalb und außerhalb eines Kegels. Die Pixel werden also völlig analog zu den atomaren Streuern der Holographie selbst zu punktförmigen Strahlungsquellen mit Kugelwellen.

Die Propagation findet in diesem Falle jedoch in das Rekonstruktionsvolumen statt, dass sich wiederum innerhalb des Kristalls befindet. In Abbildung 7.1a ist gezeigt, dass es sich um eine Einheitszelle handelt, die nah an der Referenzwellenquelle liegt. Genau genommen wird in jede Raumrichtung ein Rand mit Länge $a_W/4$ symmetrisch um die Einheitszelle

herum simuliert. Insgesamt beträgt die Ausdehnung dieses kubischen Rekonstruktionsvolumens damit je kartesischer Raumrichtung

$$d_r = \frac{6}{4} \cdot a_W \approx 4,75 \text{ \AA}. \quad (7.3)$$

Die Amplitude der holographischen Partikularwelle ist nach Gleichung (2.30) proportional zur Amplitude der Referenzwellenquelle am entsprechenden Ort. Diese wiederum ist nach Gleichung (2.29) antiproportional zum Abstand zwischen Referenz- und Streuatom. Weit entfernte Atome erzeugen also nur geringe Signale auf dem Schirm, die in der Rekonstruktion entsprechend schwierig vom Untergrund zu unterscheiden sind. Das Volumen wurde aus diesen Gründen bewusst nah an die Referenzwellenquelle gelegt.

Einige Atome werden aufgrund ihrer kegelförmigen Wellenausbreitung auch gar nicht auf den Schirm abgebildet, wenn der Kegel nicht auf den Schirm gerichtet ist. Im Laufe dieses Kapitels wird diese geometrische Grundvoraussetzung noch von großer Bedeutung sein.

Weiterhin erzeugen Streuer abseits der optischen Verbindungsachse zwischen Referenzwellenquelle und Schirm höherfrequente Muster auf dem Schirm und sind damit anfälliger für die in Abschnitt 4.2.2 beschriebenen Aliasing-Effekte, die die spätere Rekonstruktion unmöglich machen [11, Kap. 14.2.2.2].

Für die im Folgenden berechneten Rekonstruktionsvolumina wurden relativ zur Referenzwellenquelle mit Ort

$$\mathbf{r}_0 = (0 \quad 0 \quad 0) \quad (7.4)$$

aus den genannten Gründen stets Werte mit der asymmetrischen Verschiebung

$$\begin{aligned} x_r/a_W &\in [-11/4, -5/4], \\ y_r/a_W &\in [-3/4, 3/4], \\ z_r/a_W &\in [3/4, 9/4] \end{aligned} \quad (7.5)$$

angenommen. x_r ist dabei bewusst völlig analog zu den Ausführungen in Abschnitt 2.3.2

abseits des Zentrums gewählt, um die Dominanz der Referenzwelle zu reduzieren.

Für die weitere Auswertung werden alle Rekonstruktionsvolumina je Raumrichtung eine Auflösung von

$$N_r = 101 \quad (7.6)$$

Pixeln haben. Die Kombination aus Ausdehnung d_r und Pixelzahl N_r wurde dabei so gewählt, dass sich die exakten simulierten Atompositionen ebenfalls in den rekonstruierten Koordinaten wiederfinden. Insbesondere daher ist eine ungerade Anzahl an Rekonstruktionspixeln je Raumrichtung nötig. So werden keine Aufbauten nachteilig bewertet, die durch eine perfekte Abbildung das Atom scharf an die ursprüngliche Position zurück projizieren. In einer experimentellen Umsetzung ist diese Position relativ zur Referenzwellenquelle a priori natürlich unbekannt und kann höchstens vermutet werden. Daher müsste hier in späteren Experimenten eine höhere Auflösung des Rekonstruktionsvolumens angenommen werden. Da im Rahmen dieser Arbeit aber viele Muster ausgewertet werden, musste ein Kompromiss für N_r mit

$$N_r^3 \approx 103 \cdot 10^4 \quad (7.7)$$

getroffen werden. Die Laufzeit skaliert bei Hinzunahme weiterer Pixel gemäß N_r^3 kubisch in N_r . Im Laufe der Musterauswertung wird aber auch klar, dass sich Atome sehr wohl in benachbarten Pixeln durch ein „Aus schmieren“ bemerkbar machen. Es ist also keine exakte Positionierung der gewählten Koordinaten nötig.

Nach der Bildung des Betragsquadrats über die rekonstruierte Welle ergibt sich auch hier eine Intensität Γ nach Gleichung (A.12), die die eigentliche Rekonstruktion darstellt. Vor der weiteren Behandlung normieren wir diese nach Gleichung (6.13) zu Γ_n . Abbildung 7.1c stellt Γ_n in einigen Schnitten durch das Rekonstruktionsvolumen dar. Es treten Häufungen von Pixeln auf, die Werte nahe $\Gamma_n = 1$

zeigen. Wir vermuten hier ein rekonstruiertes Atom. Und tatsächlich stimmen diese Anhäufungen mit den in Abbildung 7.1c hervorgehobenen, simulierten Positionen überein.

Zur Einordnung des numerischen Aufwands dieser einzelnen Simulation sei auf die verwendete Hardware verwiesen. Die Simulationen wurden alle am High Performance Computing (HPC) Cluster LiDO3 durchgeführt. I. d. R. wurden dafür jeweils 20 Kerne alloziert, die gleichzeitig gerechnet haben. Die Laufzeit für die oben beschriebene Holographie und Rekonstruktion betrug dabei in Summe nach allen Optimierungen in etwa 20 Minuten.

7.1.1 Visualisierung

Die Betrachtung von Schnitten durch das Rekonstruktionsvolumen ist für die Auswertung und Visualisierung nicht hilfreich. Stattdessen ist also eine weitere Abstraktion vonnöten. Ziel ist, dass sich die Rekonstruktion in einem dreidimensionalen Streudiagramm darstellen lässt. An der Stelle eines gefundenen Atoms soll ein Pixel eingefärbt werden, ansonsten wird er ausgeblendet. Aus diesem Grund verwenden wir die bereits eingeführte Klassifizierungsfunktion $\hat{c}(x)$ nach Gleichung (6.15), die anhand des Thresholds t_s entscheidet, ob ein Pixel als Atom gewertet werden soll oder nicht. Alle Pixelwerte, die oberhalb dieses Grenzwertes t_s liegen, werden dann im Streudiagramm markiert.

Bei willkürlicher Annahme verschiedener Thresholds t_s ergibt sich das in Abbildung 7.2 gezeigte Bild. Ferner sind dort als Gitterkugeln die Positionen dargestellt, an denen ein Atom zu erwarten gewesen wäre. Diese Kugeln besitzen bereits den in Abschnitt 6.3.4 eingeführten Radius r_a , der eine lokale Unsicherheit in der Verortung der Atome zulässt. Bei Betrachtung der Streudiagramme wird deutlich, dass der Threshold t_s gut gewählt werden muss. Ist er zu gering, werden viele

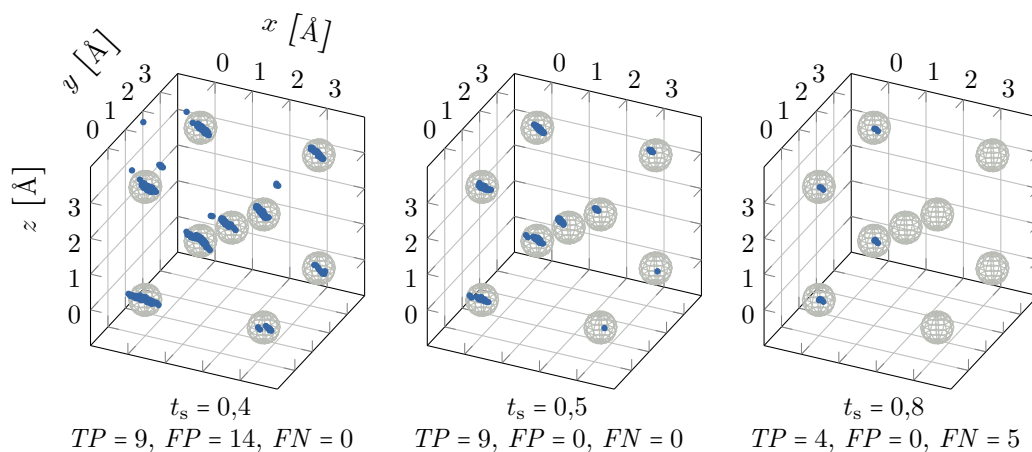


Abbildung 7.2: Der eingestellte Threshold t_s bestimmt, ab welcher Rekonstruktionsintensität Γ_n ein Pixel als rekonstruiert angesehen werden soll. Entsprechend gilt es einen idealen, problemabhängigen Threshold (hier also etwa um $t_s \approx 0,5$) zu ermitteln, der die Rekonstruktion bestmöglich abbildet. Bei der Klassifizierung helfen die in Abschnitt 6.3.2 eingeführten Klassen zu richtig und falsch Positiven bzw. Negativen. Es ist erkennbar, dass ein niedriger Threshold zu einer Überklassifikation führt, also zu viele vermeintliche Atome erkannt werden. Ein zu hoher Threshold hingegen bedingt in der Rekonstruktion einige fehlende Atome.

Pixel als Positive klassifiziert – und das auch weit über die Grenzen der durch die Atomkugelradien r_a gestatteten Unsicherheiten hinaus. Dementgegen ist ein zu großer Threshold ursächlich für eine Unterklassifikation. Einige Atome verschwinden im Untergrund und werden nicht mehr dargestellt. Ein fester Threshold für alle Rekonstruktionen lässt sich damit also nicht finden. Zu sehr sind dessen ideale Werte vom konkreten Problem abhängig.

Stattdessen soll in den meisten Fällen ein Algorithmus genutzt werden, der den Threshold optimiert, indem er ihn anhand einer Randbedingung minimiert. Das Randkriterium ist dabei die Maximierung eines Gütekriteriums, das die Übereinstimmung zwischen Ergebnis und Realität bestimmt. In Abschnitt 7.2 sollen dazu verschiedene Kriterien vorgestellt und eines als besonders geeignet herausgestellt werden. Allgemein gilt aber: Je niedriger der Threshold gewählt werden kann, unter

dem alle Atome rekonstruiert und keine Pixel fehlrekonstruiert werden, desto klarer zeichnen sich die Atome in der Rekonstruktion ab. Damit setzen sich die Atome auch klarer gegenüber dem Untergrund ab. Im realen Experiment, bei dem a priori *keine* Information über die Lage der zu erwartenden Atome vorliegt, fällt dann die Auswertung eines rekonstruierten Musters ebenfalls leichter.

Ferner wird die reine Rekonstruktion aus den oben im Rahmen der Positionierung der zu rekonstruierenden Einheitszelle dargelegten Gründen stark von der Referenzwelle dominiert. Naheliegende Atome tragen mit größerer Amplitude zum Hologramm und damit mit größerer Intensität zur Rekonstruktion bei. Qualitativ lässt sich das bereits aus der rekonstruierten Welle

$$\psi_{\text{rek}} = \left(|\psi_r|^2 + |\psi_o|^2 \right) \psi_r + |\psi_r|^2 \psi_o + |\psi_r|^2 \psi_o^* \quad (7.8)$$

ablesen, die in der vorliegenden Form aus Gleichung (2.40), (A.8) und (A.10) sinngemäß gewonnen wurde. Von besonderem Interesse ist hier der mittlere Summand, in dem die Intensität der Referenzwelle die Objektwelle skaliert. In der Rekonstruktion führt also die Abstandszunahme zum Referenzwellenursprung unmittelbar zu abnehmenden Intensitätswerten in der Rekonstruktion.

Daher werden die Daten im Rahmen dieser Arbeit aufbereitet, um diese Abstandsabhängigkeit zu eliminieren. Dies geschieht durch Normieren auf die Referenzwelle. Dass dies nicht das optimale Vorgehen ist, wird auch sofort bei Betrachtung der in Abbildung 7.2 dargestellten Streudiagramme deutlich, bei denen diese Optimierung bereits durchgeführt wurde. Atome nah an der Referenzwelle zeichnen sich nun nur noch durch einige wenige Pixel aus, die allerdings dominant sind und innerhalb der Atomkugeln zentrumsnah auftreten. Um das Zentrum herum liegende Pixel geraten durch die Normierung schnell unterhalb des gewählten Thresholds, während fern liegende Atome sich durch mehr Pixel auszeichnen, die durch die Normierung betont werden.

Der dritte Term mit konjugierter Objektwelle in Gleichung (7.8) ist ursächlich für das Zwillingbild und bereitet nach Abschnitt 3.1.2 keine Probleme. Allerdings lässt sich der erste Mischterm nicht eliminieren und die Skalierung erzeugt hier Rauschen, das die Rekonstruktionsqualität beeinträchtigen kann.

Im Rahmen der weiteren Auswertung ist das nicht weiter von großer Bedeutung. Allerdings wäre denkbar, lediglich auf einen Referenzwellenabstand r zu einer bestimmten Potenz p zu normieren, um ein Gleichgewicht zwischen beiden Effekten herzustellen. Durch die Quadratur im Rahmen der Intensitätsbildung ist im klassischen Ansatz $p = 2$. Dementgegen kann eine neuerliche Rekonstruktion Γ' durch Optimierung von p anhand eines oben bereits genannten Gütekriteriums mittels

$$\Gamma' = \Gamma/r^p, \quad p \in]0, 2] \quad (7.9)$$

problemangepasst empirisch bestimmt werden.

Durch automatische Optimierung ergibt sich für die oben beschriebene Simulation nach Normierung auf die Referenzwelle ein optimaler Threshold von $t_s \approx 0,448$. Bei dessen Anwendung ergibt sich das in Abbildung 7.3 gezeigte Rekonstruktionsvolumen.

In Abschnitt 6.2.1 wurde bereits thematisiert, dass ein tiefes Kristallvolumen in der Rekonstruktion zu Abbildungsproblemen führt. Diese äußern sich hier in Form des Aufweitens der eigentlich als kugelförmig anzunehmenden Atome entlang der Verbindungsachse zwischen Referenzwellenursprung und Schirm. Während die Kristallausdehnung in x - und y -Richtung also kein Problem darstellt, wird der Kristall in Richtung der Verkipfung als tief wahrgenommen und erzeugt so Artefakte mit übereinstimmendem Winkel. Diese Vorzugsrichtungen sind zur besseren Visualisierung in der nicht maßstäblichen Darstellung von Abbildung 7.3 ebenfalls eingezeichnet.

7.2 F_1 -Score

Um einen optimalen Threshold zu bestimmen und ferner die Rekonstruktionsergebnisse verschiedener Experimente miteinander vergleichen zu können, ist ein objektives Maß für die Übereinstimmung der Rekonstruktion mit dem tatsächlichen Kristall festzulegen. Anhand der in Abschnitt 6.3.2 ff. getätigten Überlegungen zu Klassifizierungen in verschiedene Klassen und experimenteller Unsicherheiten wollen wir nun ein einziges Bewertungskriterium wählen, das dieser Aufgabe gerecht wird. Es muss sich als Optimierungsgröße eignen, damit es für die Threshold-Untersuchung maximiert werden kann.

Im Rahmen der in Abschnitt 6.3 vorgestellten binären Klassifikation wurden viele solcher Messgrößen, *Scores* genannt, bereits definiert. Sie sind in der Lage, ein Ergebnis zu bewerten,

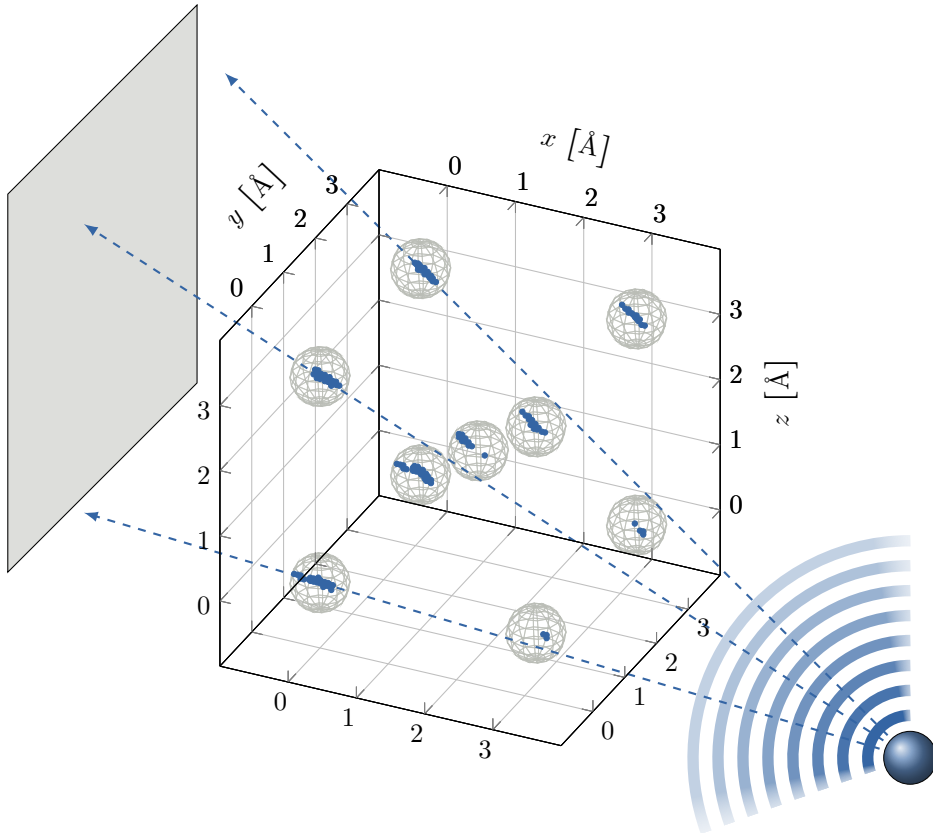


Abbildung 7.3: Rekonstruktion des beispielhaften Hologramms aus Abbildung 7.1 nach Anwendung der Klassifizierungsfunktion (6.15) mit optimalen Threshold $t_s \approx 0,448$. Die Daten wurden dabei auf die Referenzwelle normiert, um deren starke Dominanz zu eliminieren. Ferner ist im korrekten Maßstab die Referenzwellenquelle rechts vom Streudiagramm eingezeichnet, der Schirm hingegen nicht maßstäblich links. Es wird gleich mehrerlei deutlich: Atome, die ferner von der Referenzwellenquelle liegen, schmierern aus. Das geschieht in Richtung der Verbindungsachsen zwischen Referenzwellenquelle über die Streuer hinweg in Richtung Schirm, derer beispielhaft drei eingezeichnet sind. Ursächlich dafür ist die vom Schirm aus wahrgenommene Tiefe nach Beschreibung in Abschnitt 6.2.1. Ferner zeichnen sich Atome nahe an der Referenzquelle sehr scharf ab. Das wird durch die geringe Zahl von positiv markierten Pixeln nah am Zentrum der Unsicherheitskugel mit Radius r_a deutlich. In der experimentellen Auswertung bedeutet dies einerseits, dass die Rekonstruktion sich klar abzeichnet, andererseits aber auch, dass die Auflösung entsprechend hoch gewählt werden muss, um diese Pixel einzufangen. Erweiterte Darstellung nach Ref. [20, Abb. 2a].

dessen eigentlicher Ausgang zuvor bereits bekannt ist. Allerdings stehen uns dabei nicht alle eingeführten Größen zur Verfügung, die auf den Komponenten (6.21) der Wahrheitsmatrix basieren. Zu Ende von Abschnitt 6.3.4 hatten wir bereits thematisiert, dass die Definition von TN im vorliegenden Fall nicht ohne Weiteres möglich ist und daher darauf verzichtet werden muss. Daher scheiden bereits einige Scores aus, etwa der häufig genutzte und in Anhang C aufgeführte *Matthews Correlation Coefficient* nach Ref. [212].

Primäre Fragestellung bei der Wahl eines Kriteriums ist, welche Gewichtung für das vorliegende Problem angemessen ist. Durch Festlegung eines Qualitätskriteriums geschieht automatisch durch die Definition eines von der Wahrheitsmatrix abhängigen Terminus eine Priorisierung einer oder mehrerer Eigenschaften. Bei der Anwendung rein algorithmischer Bewertungen in verschiedenen Disziplinen der Wissenschaft muss diese Priorisierung wohlüberlegt sein. Insbesondere im Zusammenhang mit Menschen als Untersuchungsobjekt sind Diskussionen über die richtige Wahl eines Scores aktueller Gegenstand der Forschung. So kann Fairness ein neu zu berücksichtigendes Kriterium sein [213]. Etwa, wenn ein binärer Klassifikator möglichst genau darüber entscheiden soll, ob eine Person in der Lage wäre, einen Kredit abzubezahlen, dabei aber nicht auf Daten etwa der ethnischen Zugehörigkeit zurückgreifen soll [214].

In Anhang C finden sich weitere mögliche Scores zur Bewertung des Ausgangs einer binären Klassifikation in einer tabellarischen Übersicht. Diejenigen davon, die sich unter der oben gemachten Einschränkung – fehlende Abhängigkeit von TN – noch für die Bewertung eines Rekonstruktionsvolumens eignen, finden sich in Abbildung 7.4 dargestellt. Sie werden hier in Abhängigkeit vom gewählten Threshold t_s gezeigt.

Wir erkennen, dass sich Präzision und Sensitivität gemäß ihrer Beschreibung in Abschnitt 6.3.3 nur für jeweils einen anderen Teil

des gezeigten Spektrums eignen. Die Sensitivität als Wahrscheinlichkeit für das tatsächliche Erkennen eines beliebigen Atoms muss bei niedrigem Threshold ein gutes Ergebnis angeben. Schließlich wird für $t_s = 0$ jeder Pixel als mögliches Atom klassifiziert. Es fehlt also die Berücksichtigung der FP , also der fälschlich als Atom erkannter Pixel. Umgekehrt verhält es sich mit der Präzision, denn ihr fehlt die Berücksichtigung der FN . Ein Ergebnis, das durch hohen Threshold $t_s = 1$ keinen Pixel als Atom klassifiziert, wird hier gut bewertet, weil im Sinne der Präzision keine Falschklassifikation stattgefunden hat.

Es zeigt sich, dass der F_1 -Score hier gut zwischen beiden Größen in ihrem jeweiligen Schwerpunktgebiet zu interpolieren scheint. Der allgemeine F -Score ist dabei definiert als

$$F_\beta = (1 + \beta^2) \frac{ppv \cdot tpr}{tpr + \beta^2 \cdot ppv}, \quad (7.10)$$

hier soll er aber in seiner geläufigsten Form F_1 mit $\beta = 1$ genutzt werden [215]. Es handelt sich bei ihm dann genau um das harmonische Mittel von Präzision und Sensitivität

$$F_1 = \frac{2 \cdot ppv \cdot tpr}{ppv + tpr} \in [0, 1]. \quad (7.11)$$

Auch andere Kriterien, wie der Fowlkes-Mallows-Score nach Refs. [216, 217] und der häufig auch Threat Score genannte Critical Success Index nach Ref. [218], liegen nahe bei der durch den F_1 -Score angegebene Bewertung. Sie finden sich beide ebenfalls in Anhang C definiert. Sie belegen durch ihre gute Übereinstimmung, dass der F_1 -Score ein geeignetes Kriterium ist. Die Wahl eines Scores ist damit primär von der gewünschten Schwerpunktsetzung abhängig.

Der F_1 -Score ist insbesondere dann eine geeignete Größe, wenn die TN im Kontext der Überprüfung nicht relevant oder wie hier gar nicht definierbar sind [192, Kap. 12.1]. Allerdings tendiert er dazu, eine Klassifizierung

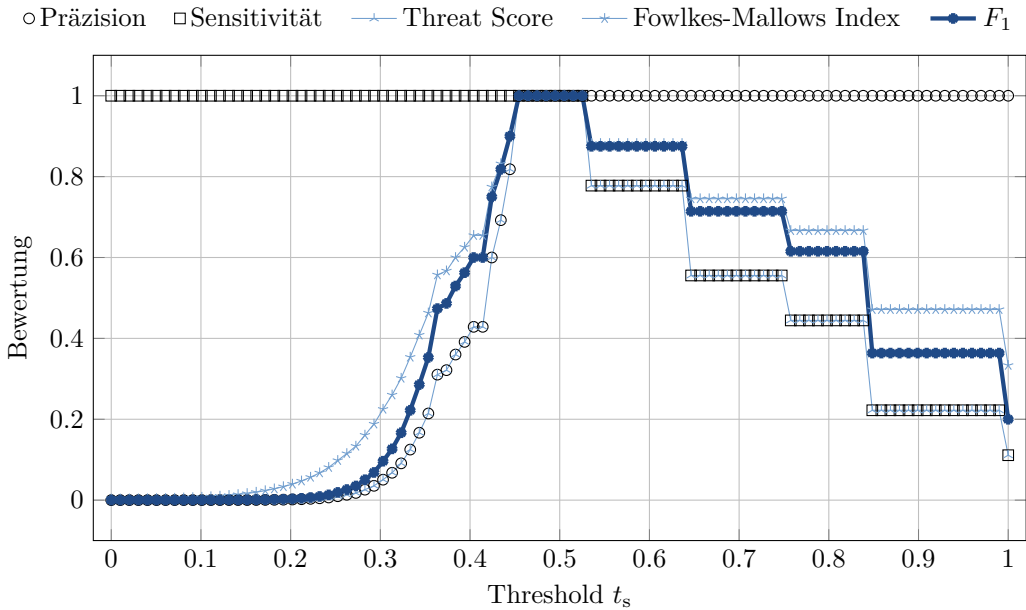


Abbildung 7.4: Die Auftragung verschiedener möglicher Bewertungskriterien eines Rekonstruktionsergebnisses in Abhängigkeit des Thresholds t_s zeigt, dass Sensitivität und Präzision nur für jeweils eine Hälfte des Spektrums brauchbare Ergebnisse liefern. Der hervorgehobene F_1 -Score hingegen interpoliert gut zwischen diesen beiden Größen. Threat Score und Fowlkes-Mallows Index sind zwei weitere etablierte Kriterien, die dem Verlauf von F_1 gut folgen.

zu bevorzugen, die alle Pixel als Negative angibt – was in Abbildung 7.4 auch anhand der asymmetrischen Zu- und Abnahme entlang des Thresholds t_s deutlich wird. Es ist mit abnehmender Klassifizierung als Positive damit tendenziell einfacher, einen besseren Score zu erhalten [192, Kap. 12.1]. Da die Threshold-Anpassung in der vorliegenden Simulation aber die Ausgabe von mindestens einem Atom erzwingt, ist keine reine Negativ-Rekonstruktion möglich. Diese Schwerpunktsetzung des F_1 -Score stellt also keine Einschränkung für die vorliegende Arbeit dar.

7.3 Massenauswertung

Mit dem nun definierten F_1 -Score kann jeder Rekonstruktion eine einzige reelle Zahl zwi-

schen 0 und 1 zugewiesen werden. Dabei bedeutet $F_1 = 1$ eine perfekte Übereinstimmung zwischen dem berechneten Rekonstruktionsvolumen und dem ursprünglich angenommenen Kristall. Dementgegen beschreibt $F_1 = 0$ keine erkennbare Entsprechung.

So lässt sich nun zu jeder Rekonstruktion ein optimaler Threshold zur Optimierung des F_1 -Score finden und in einer Grafik visualisieren. Im Rahmen dieser Arbeit wurden dazu insbesondere die Schirmgröße d^2 und dessen Auflösung N^2 untersucht. Dazu wurden Schirmausdehnungen von

$$d \in \{1 \text{ cm}; 2 \text{ cm}; \dots; 20 \text{ cm}\} \quad (7.12)$$

je Dimension und eindimensionale Auflösungen von

$$N \in \{100; 200; \dots; 2000\} \quad (7.13)$$

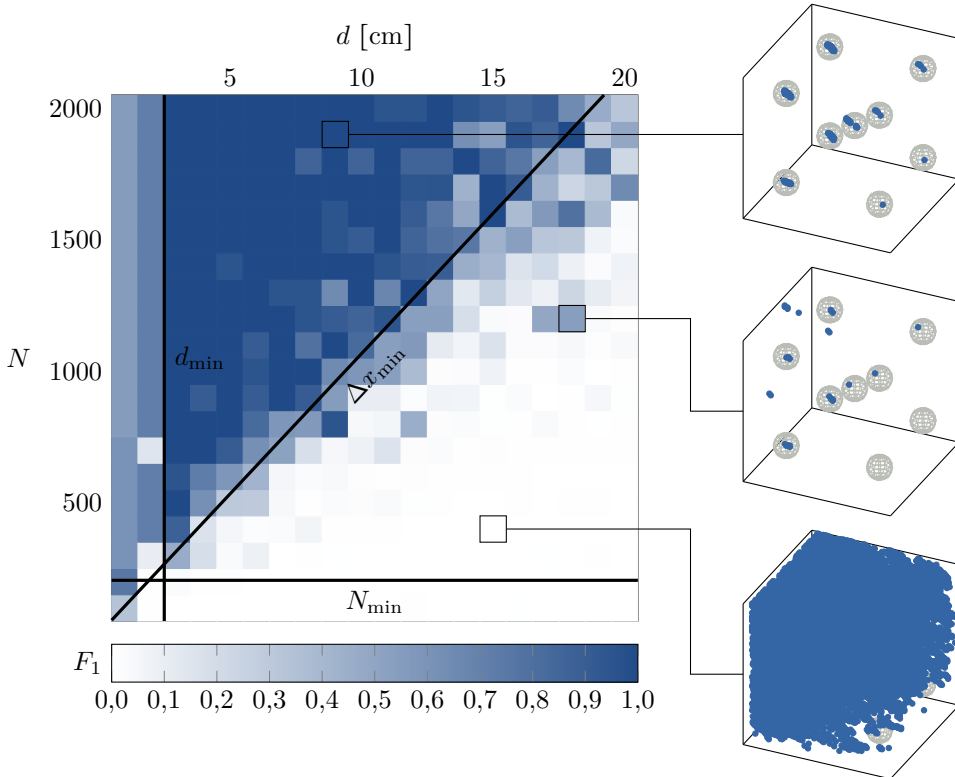


Abbildung 7.5: Jedem Pixel im Plot entspricht ein Kristall, der bei $E_{\text{kin}} = 20 \text{ keV}$ simuliert und rekonstruiert wurde. Dem Rekonstruktionsergebnis lässt sich dann ein F_1 -Score derart zuordnen, dass die Übereinstimmung mit der simulierten Einheitszelle bewertet werden kann. Die gezeigten Beispiele belegen die Entsprechung bei hohem Score. Eingezeichnet sind beschriftete Geraden als Mindestvoraussetzungen für die erfolgreiche Rekonstruktion. Im Einzelnen die Pixelgröße Δx_{\min} nach Abtasttheorem und die durch das Oversampling Ratio σ_{\min} nach Gleichung (7.15) gegebene Pixelzahl N_{\min} . Ferner können aus den im Text beschriebenen geometrischen Überlegungen nur Schirme mit einer Kantenlänge von mindestens d_{\min} überhaupt rekonstruieren. Erweiterte Darstellung nach Ref. [20, Abb. 2b].

angenommen. Mit jeweils 20 Werten eines jeden Parameters wird ein Parameterraum aus 400 (N, d) -Paaren untersucht. Alle weiteren Experimentalparameter wurden wie zuvor bereits beschrieben festgelegt und die kinetische Energie mit $E_{\text{kin}} = 20 \text{ keV}$ angenommen. Die Auswertung aller 400 Paare ergibt dann ein zweidimensionales Histogramm, wie es in Abbildung 7.5 dargestellt ist. Zur Verdeutlichung der Datenmenge sei betont, dass sich hinter jedem Pixel eine vollständige Rekonstruktion mit jeweils N_r^3 Pixeln nach Gleichung (7.7) findet. Jedes Rekonstruktionsvolumen hat selbst im komprimierten Speicherformat netCDF4, wie es in Anhang B näher beschrieben ist, einen Speicherbedarf von etwa 16,5 MiB. Insgesamt finden sich in Abbildung 7.5 damit etwa 6,6 GiB Messdaten, die innerhalb von etwa 8 Tage reiner LiD03-Rechenzeit berechnet wurden. Zur Verdeutlichung und späterem Vergleich sind drei beispielhafte Rekonstruktionsvolumina mit unterschiedlichem F_1 -Score visualisiert.

Es bilden sich klare Bereiche heraus, in denen eine Rekonstruktion gut möglich ist, während sie in anderen Teilen des (N, d) -Parameterraums unmöglich wird. Insbesondere ist ein Dreieck erkennbar, das zu größeren Schirmen d bei geringerer Auflösung N beschränkt wirkt.

Ferner gilt es, die Kante für $d \leq 2 \text{ cm}$ zu rechtfertigen. Dies lässt sich durch einfache geometrische Überlegungen erläutern. Damit eine Rekonstruktion von Streuern überhaupt möglich ist, müssen diese Streuer in irgendeiner Form eine Information auf den holographischen Schirm abbilden. Durch die in Abschnitt 3.1.2 beschriebene kegelförmige Welle ist ihre Ausbreitung aber begrenzt. Aufgrund der in Abbildung 7.1a dargestellten Lage der Referenzwelle und des konkreten Rekonstruktionsvolumens gelangen bei der angenommenen Verkippung von $\varphi = 90^\circ - \vartheta = 70^\circ$ nur begrenzt Beugungsinformationen auf dem Schirm. Bei einer Ausdehnung von $d = 2 \text{ cm}$ sind das lediglich zwei der insgesamt

neun rekonstruierten Streuer, bei $d = 1 \text{ cm}$ dann keiner mehr. Entsprechend sind die Ergebnisse ungeachtet der Pixelzahl für diese beiden Werte entsprechend schlecht bewertet, also nahe $F_1 = 0$. Siehe auch die Erläuterung zur Güte in Abschnitt 7.2.

Für die folgenden Untersuchungen sollen daher die Daten mit $d < d_{\text{min}} = 3 \text{ cm}$ ausgeklammert werden. Lassen sich in einem Experiment kleinere Schirmgrößen nicht vermeiden, ist insbesondere auf ein passendes Experimentsdesign und die Wahl des richtigen Rekonstruktionsvolumens zu achten. Mit zunehmender Schirmgröße wird die Wahl des Rekonstruktionsvolumens freier. Bei gleichbleibend zu wählender Pixelgröße nimmt die Pixelzahl und damit die Rechenzeit demgegenüber aber entsprechend zu.

7.3.1 Oversampling Ratio

Die Mindestanforderung an das Oversampling Ratio (4.54) nach Refs. [144, 145] gibt eine untere Grenze für die Anzahl der Schirmpixel an. Deren Erfordernis wurde über die Zahl der Unbestimmten innerhalb des Gleichungssystems der Rekonstruktion hergeleitet. Allerdings wird ebendort von der Rekonstruktion durch Fourier-Transformation ausgegangen, die Bestimmung und Herleitung der Zahl der relevanten Pixel aus der Anzahl der beteiligten Fourier-Koeffizienten entsprechend leicht. Wird die Berechnung auch hier lediglich anhand der Zahl der Rekonstruktionspixel durchgeführt, erhalten wir ein Ratio von

$$\sigma = \frac{N_{\text{scr}}}{N_{\text{rec}}} = \frac{N^2}{N_r^3} = \frac{N^2}{101^3} \geq 2 \quad (7.14)$$

und damit als Bedingung $N \geq 1436$. Allerdings zeigt Abbildung 7.5 zweifelsfrei, dass auch unterhalb dieser Pixelzahl eine Rekonstruktion möglich ist. Ferner würde diese naheliegende Annahme $N_{\text{rec}} = N_r^3$ zur unintuitiven Konsequenz führen, dass die für die Mustersaufnahme zu wählenden Schirmparameter

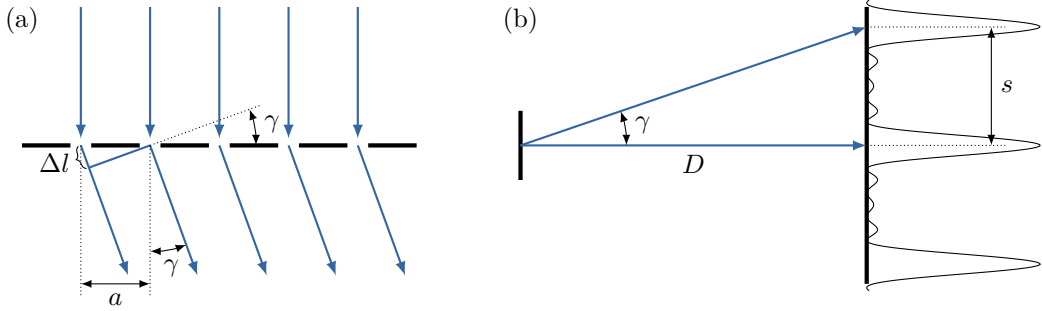


Abbildung 7.6: Geometrische Überlegungen zur Herleitung einer Mindestpixelgröße nach Abtasttheorem erfordern, zunächst den Abstand zwischen Intensitätsmaxima festzustellen. (a) Der Abstrahlwinkel γ bestimmt den Gangunterschied. Darstellung erfolgt nach Ref. [219, Abb. 10.43]. (b) Derselbe Winkel dient dann zur Bestimmung der Position auf dem Schirm.

damit abhängig von der erst im Anschluss stattfindenden Rekonstruktion und ihrer Auflösung sind.

Die hier vorgestellten Daten legen eher eine Abhängigkeit von der Zahl der eingangs einfließenden Streuer nahe. Deren Oversampling Ratio lautet

$$\sigma = \frac{N^2}{N_{xyz}} \stackrel{(6.6)}{=} \frac{N^2}{102 \times 102 \times 51 \times 2} \geq 2 \quad (7.15)$$

und ergibt eine untere Grenze von $N_{\min} = 205$, die gut mit den gezeigten Daten übereinstimmt. In Abschnitt 6.2.1 wurde bereits erläutert, dass eine Erhöhung der Zahl an Streuern N_{xyz} die Komplexität der Rekonstruktion erhöht. Es liegt also nahe, dass damit die Zahl der Schirmpixel und die Menge an gespeicherter Information steigen muss. Das ist durch die Berechnung nach Gleichung (7.15) gegeben.

7.3.2 Abtasttheorem

Neben der soeben erläuterten Grenze für kleine N , die durch $\sigma(N)$ gegeben ist, fällt in Abbildung 7.5 eine diagonal verlaufende Abhängigkeit auf. Sie scheint durch d/N , also

durch die Pixelgröße bestimmt, und hat entscheidenden Einfluss auf die Rekonstruktionsfähigkeit. Das wird intuitiv nachvollziehbar, wenn die Erläuterungen zum Abtasttheorem aus Abschnitt 4.2.1 hinzugezogen werden.

Der Zusammenhang zwischen Abtasttheorem und der erforderlichen Mindestpixelgröße ist dabei nicht unmittelbar zu fassen [220]. Es stellt sich nämlich die Frage nach der Wellenlänge der aufzunehmenden Muster. Die Größe der Pixel kann rein physikalisch nicht der subatomaren Wellenlänge nach Gleichung (7.2) der eingesetzten Elektronenwellen entsprechen. Stattdessen ist die Modulation der aufzunehmenden Information durch die Fluktuation des durch die Beugung hervorgerufenen Intensitätsmusters gegeben. Diese können wir als wesensverwandt mit den Interferenzen eines Gitters annehmen, wenn wir die Verkippung des Kristalls zunächst vernachlässigen.

Für die Beugung am Gitter ergibt sich nach Ref. [219, Kap. 10.5.3] der in Abbildung 7.6a gezeigte Gangunterschied von

$$\Delta l = a \sin \gamma. \quad (7.16)$$

Auf dem Schirm addieren sich die Wellenfunktionen genau dann zur maximalen Intensität,

wenn dieser Gangunterschied einem ganzzahligen Vielfachen der Wellenlänge entspricht. Zur konstruktiven Interferenz kommt es also bei

$$a \sin \gamma = m \cdot \lambda \quad (7.17)$$

mit einer Zähllkonstante $m \in \mathbb{N}$ und Gitterkonstante a . Sei nun s der Abstand zwischen dem Hauptmaximum nullter und erster Ordnung (also für $m = 1$) auf dem Schirm und der Abstand zwischen Gitter und Schirm mit D bezeichnet, gilt für den Winkel γ auch die geometrische Beziehung

$$\sin \gamma = s/D, \quad (7.18)$$

die sich leicht aus Abbildung 7.6b ablesen lässt. Durch Einsetzen in Gleichung (7.17) erhalten wir dann als Abschätzung für die Abstände zwischen den Maxima

$$s = \lambda \cdot D/a. \quad (7.19)$$

Um diese Maxima auflösen zu können, ist nach Abtasttheorem (4.37) also – nach Einsetzen der Experimentalparameter von oben – mindestens eine zu s halbe Pixelgröße von

$$\Delta x_{\min} = \frac{s}{2} \approx 0,274 \text{ mm} \quad (7.20)$$

nötig. Bei Berücksichtigung der Verkipfung des Kristalls um $\varphi = 70^\circ$ erhalten wir damit näherungsweise

$$\Delta x_{\min} = \frac{s}{2} \cdot \cos(70^\circ) \approx 0,0937 \text{ mm}. \quad (7.21)$$

Die aus dieser Pixelgröße resultierende Diagonale ist in Abbildung 7.8 eingezeichnet und kann den Bereich unterhalb dieser Gerade mit $F_1 = 0$ gut erklären. Eine Rekonstruktion in diesem Bereich ist also durch eine rein physikalische untere Grenze undenkbar.

Das Histogramm der Abbildung 7.5 lässt sich zur weiteren Untersuchung des Einflusses der Pixelgröße entfalten. Für jedes Tupel (d, N) mit $d \geq d_{\min}$ berechnen wir die Pixelgröße

$$\Delta x(d, N) = d/N \quad (7.22)$$

und ordnen dem Tupel den ermittelten F_1 -Score zu. Dadurch entstehen sehr viele eng beieinander aber nicht äquidistant liegende Pixelgrößen. Da mit zunehmender Pixelgröße nur noch Schirme mit großem N und gleichzeitig kleinem d beitragen können, werden entsprechende Δx -Werte rar. Um diesem Umstand gerecht zu werden, wird weiterhin eine logarithmische Skalierung angenommen. Nah beieinander liegende Werte werden in Bins in entsprechend logarithmischem Abstand zusammengefasst und gemittelt. Wir erhalten damit das in Abbildung 7.7a dargestellte Diagramm, das auch die Standardabweichung innerhalb der im zugehörigen Bin liegenden Daten nach Mittlung zeigt.

Es bestärkt die hergeleitete Abhängigkeit der Pixelgröße vom Abtasttheorem. Für eine Pixelgröße $\Delta x \leq \Delta x_{\min}$ ist eine Rekonstruktion damit möglich, unterhalb der Mindestgröße hingegen nicht. Im theoretischen Idealfall zeigt sich eine Abhängigkeit der invertierten Stufenfunktion

$$F_1(\Delta x) = 1 - \Theta_1(\Delta x - \Delta x_{\min}) \quad (7.23)$$

mit der Heaviside-Funktion Θ_k nach Gleichung (6.30).

Um deren experimentelle Flanke und damit die Mindestpixelgröße Δx_{\min} zu ermitteln, bestimmen wir in einer ersten Schätzung die Fläche oberhalb der Messwerte für alle Pixelgrößen links eines als Mindestpixelgröße angenommenen Wertes und rechts davon die Fläche unterhalb der Messwerte. Diese in Abbildung 7.7c gezeigte Gesamtfläche ergibt minimiert einen optimalen Wert Δx_{\min} als erste Schätzung.

Eine weitere Möglichkeit zur Bestimmung der Mindestpixelgröße aus den Messdaten besteht aus einer algorithmischen Ausgleichsrechnung der Messdaten gegen eine Funktion. Dabei eignet sich die reine Stufenfunktion nach Gleichung (7.23) für eine Parameterschätzung nicht. Stattdessen muss ein „Aufweichen“, wie es in Abbildung 7.7b gezeigt

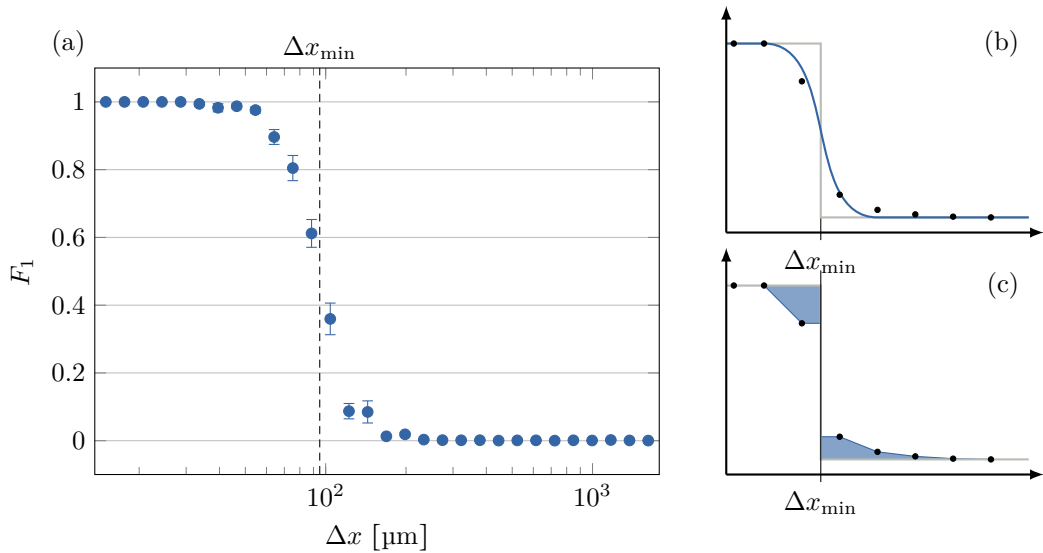


Abbildung 7.7: Untersuchung, inwieweit die erfolgreiche Rekonstruktion durch die Pixelgröße bedingt ist. (a) Nach Mittelung über verwandte Pixelgrößen ergibt sich die gezeigte Datenlage, die sehr an eine Stufenfunktion erinnert. Daher werden zur exakten Bestimmung der minimalen Pixelgröße Δx_{\min} nach Abtasttheorem zwei Wege zur validierten Bestimmung eingeschlagen. (b) Der Fit durch eine Faltung aus Stufen- und Gauß-Funktion ergibt die Pixelgröße als Fitparameter. (c) Sie lässt sich aber auch als Parameter der Fläche zwischen idealer Stufenfunktion und Messdaten verstehen und so minimieren. Beide Ansätze liefern hierbei ähnliche Ergebnisse.

ist, mit aufgenommen werden, um die experimentelle Ungenauigkeit berücksichtigen zu können. Als Ausgleichsfunktion wurde daher

$$F_1(\Delta x, \sigma) = (1 - \Theta_0(\Delta x - \Delta x_{\min})) * g(\Delta x; 1, \Delta x_{\min}, \sigma) \quad (7.24)$$

angenommen. Da wir von normalverteilten Abweichungen von der idealen Stufenfunktion ausgehen, wird hier mit der Gauß-Funktion oder Dichtefunktion der Normalverteilung

$$g(x; A, \mu, \sigma) = \frac{A}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) \quad (7.25)$$

mit Amplitude A , der Standardabweichung σ und dem für die Faltung unerheblichem Zen-

trum μ [125, Gl. 16.70b] gefaltet. Es handelt sich dabei um eine sog. *Weierstraß-Transformation* [221, Kap. 7]. Solches „Aufweichen“ durch Faltung wird häufig durch eine Gauß-Funktion realisiert. Im Rahmen der digitalen Signalverarbeitung wird hierbei von einem *Filter* gesprochen. Bei Bildern realisiert ein Gauß-Filter i. d. R. ein Weichzeichnen, wie es etwa zur Kaschierung von Details bei der Verkleinerung von Bildern genutzt wird [222, Kap. 5.2.2].

Es ergibt sich in beiden Fällen eine Mindestpixelgröße von etwa

$$\Delta x_{\min} = 0,095 \text{ mm} \quad (7.26)$$

unterhalb der eine Rekonstruktion nicht mehr möglich ist. Diese stimmt recht genau mit der

oben physikalisch ermittelten Größe (7.21) überein.

7.4 Musteraddition

Um die Rekonstruktion eines Hologramms zu verbessern, wurde in der Vergangenheit bereits mehrfach vorgeschlagen, mehrere Muster zu verschiedenen Energien aufzunehmen und diese dann zu addieren [52] oder zu mitteln [104]. Dazu werden Muster mit genau gleichen geometrischen Parametern, aber sich unterscheidender kinetischer Energie genutzt. In Abschnitt 3.3 wurde die Differenzbildmethode vorgestellt, in der zwei solcher Muster voneinander subtrahiert werden, um die holographische Information freizustellen. Dabei wurde über die stets gleiche Lage der Beugungserscheinungen erster Ordnung begründet, dass deren Subtraktion diese im Hologramm eliminiert. Bei der Rekonstruktion nutzen wir den gegenteiligen Prozess aus. Die Rekonstruktionen verschiedener Differenzbilder sollten an den gleichen Stellen Atome nachweisen. Bei deren Addition verstärken sich also die Intensitäten an der Stelle der Atome, während der Untergrund statistisch ist und sich damit nicht verstärkt. Im vorliegenden Fall addieren wir die komplexwertigen Rekonstruktionsvolumina, bevor die Rekonstruktionsintensität Γ nach Gleichung (A.12) gewonnen wird.

Im Rahmen dieser Arbeit wurden die in Abschnitt 7.3 vorgestellten Untersuchungen von insgesamt jeweils 400 Mustern für eine feste kinetische Energie von $E_{\text{kin}} = 20 \text{ keV}$ für insgesamt 21 äquidistante Energien zwischen 10 keV und 20 keV, also

$$E_{\text{kin}} \in \{10,0 \text{ keV}; 10,5 \text{ keV}; \dots; 20,0 \text{ keV}\} \quad (7.27)$$

durchgeführt. Insgesamt wurden damit also insgesamt $400 \times 21 = 8400$ Kristalle rekonstruiert und damit 139 GiB an Rekonstruktionsdaten innerhalb von 167 Tage reiner Rechenzeit generiert.

Im Folgenden bezeichne nun \mathcal{N}' die Zahl der Rekonstruktionsvolumina, die zur Auswertung addiert wurden. Am Beispiel von $\mathcal{N}' = 14$ findet sich in Abbildung 7.8 dieselbe Auswertung, die bereits im Rahmen von Abbildung 7.5 dargestellt ist. Insbesondere finden sich dort auch die gleichen Parameter (N, d) in ihrer konkreten Rekonstruktion gezeigt. Der Vergleich macht sofort deutlich, wie viel stabiler eine Rekonstruktion zu funktionieren scheint. Der Score nimmt allgemein für fast jedes Parameterpaar zu und selbst bei bereits maximiert gewesenen Rekonstruktionsvolumen zeigt sich, dass die Atome sich durch viel stärkere Messwerte auszeichnen – sich also deutlicher vom Untergrund abheben und auch im Experiment deutlich leichter zu bestimmen wären.

Der Vergleich mit Abbildung 7.5 zeigt, dass sich einige der in Abschnitt 7.3 beschriebenen Parametergrenzen verschieben. Die geometrisch gegebene Mindestanforderung d_{min} an die Schirmgröße bleibt davon aufgrund der unveränderten Geometrie unberührt.

Die Grenze des Oversampling Ratios wird allerdings deutlich unterschritten. Aus der Herleitung in Ref. [144] wird dies sofort klar. Es werden viel mehr Daten aufgenommen, die in die gleiche Zahl von Streuern entfaltet werden. Ein Ratio, das diese gesteigerte Zahl an Pixeln berücksichtigt, kann zu

$$\sigma = \frac{\mathcal{N}' \cdot N^2}{N_{xyz}} \geq 2 \quad (7.28)$$

abgeschätzt werden und skaliert damit gegenüber dem oben angegebenen Ratio für ein einzelnes Muster (7.15) linear in \mathcal{N}' . Damit ergibt sich im Falle der Anzahl mindestens benötigter Schirmpixel je Dimension

$$N \geq \sqrt{\frac{2N_{xyz}}{\mathcal{N}'}}. \quad (7.29)$$

In Abbildung 7.8 ist die dadurch ermittelbare Mindestanzahl von $N_{\text{min}} = 55$ eingezeichnet, die aufgrund ihrer geringen Größe nicht mehr

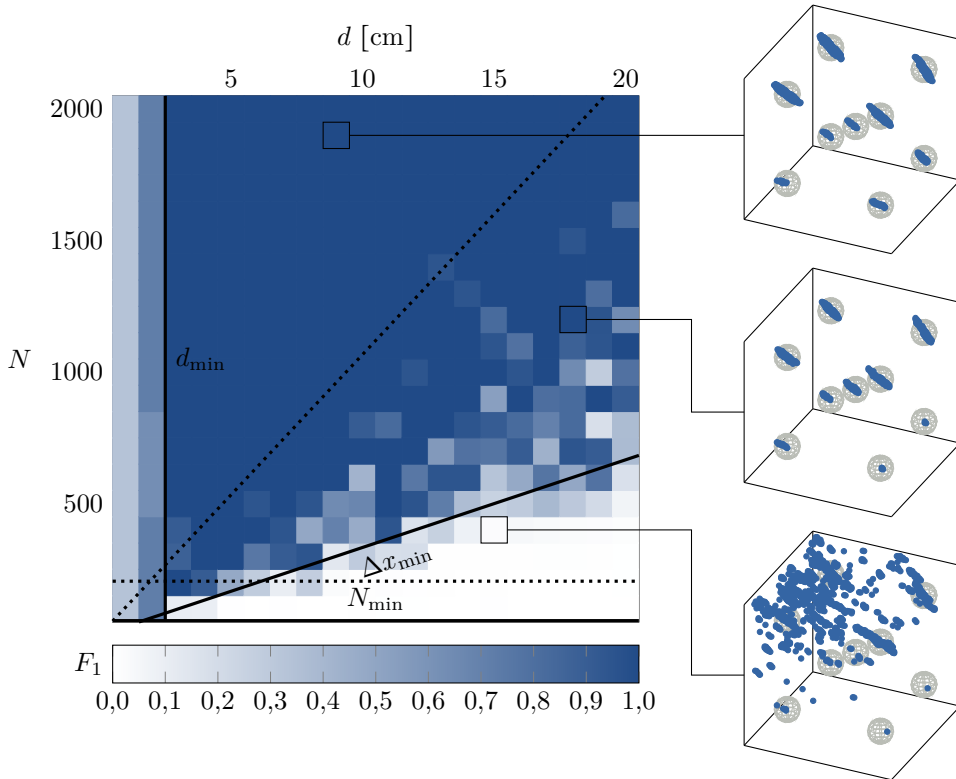


Abbildung 7.8: Nach der Addition von $\mathcal{N}' = 14$ Rekonstruktionsvolumen verschiedener Energien wird eine deutliche Steigerung in der Rekonstruktionsqualität erkennbar. Die ausgezeichneten Beispiele zeigen im direkten Vergleich mit Abbildung 7.5, dass die Rekonstruktion deutlich besser möglich ist. Viel deutlicher zeichnen sich die einzelnen Atome ab. Ferner wird die erforderliche Pixelgröße Δx_{\min} deutlich größer, die Zahl der rekonstruierbaren Szenarien steigt damit ebenfalls bedeutend. Erweiterte Darstellung nach Ref. [20, Abb. 2d].

relevant für die hier getätigten Betrachtungen mit stets $N \geq 100$ ist.

Der auffälligste Unterschied zwischen Abbildung 7.5 und 7.8 äußert sich allerdings in der stark gestiegenen Mindestpixelgröße Δx_{\min} . Das beispielhaft mittig gezeigte Muster mit $N = 1200$ und $d = 18$ cm ließ sich im Falle von $\mathcal{N}' = 1$ nur teilweise rekonstruieren, während es in Abbildung 7.8 mit $\mathcal{N}' = 14$ klar und deutlich erkennbar ist.

Die theoretische Abschätzung einer Mindestpixelgröße über das Abtasttheorem stößt hier an ihre Grenzen und kann zur Erklärung nicht mehr herangezogen werden, sodass wir uns auf die qualitative Untersuchung der vorliegenden Daten beschränken. In Abbildung 7.9 sind dafür wieder die Pixelgrößen gegen den F_1 -Score nach der in Abschnitt 7.3.2 beschriebenen Aufbereitung dargestellt. Auch hier ergeben eine Flächenabschätzung und eine Ausgleichsrechnung anhand aller Messdaten jenseits der geometrischen Grenze $d \geq d_{\min}$ in guter Übereinstimmung einen Wert von

$$\Delta x_{\min} = 0,3 \text{ mm.} \quad (7.30)$$

Er ist damit etwa um den Faktor 3,16 größer, was der Skalierung der Pixelzahl je Dimension um $\sqrt{14} \approx 3,7$ zumindest nahe kommt.

7.4.1 $\zeta_{\mathcal{N}'}$ -Score

Qualitativ haben wir damit zeigen können, dass die Hinzunahme weiterer Rekonstruktionen anderer Energien für die Berechnung der Rekonstruktion diese auch für Parameter (N, d) ermöglicht, wo sie zuvor nicht möglich schien. Auch im „Kernbereich“, der durch die in Abschnitt 7.3 beschriebenen Grenzen für Schirmgröße, Mindestpixelgröße und Pixelanzahl gegeben ist, sind die Scores im Schnitt höher und fast durchgehend $F_1 = 1$. Die Rekonstruktion wird also stabilisiert und deutlich unempfindlicher gegenüber statistischem

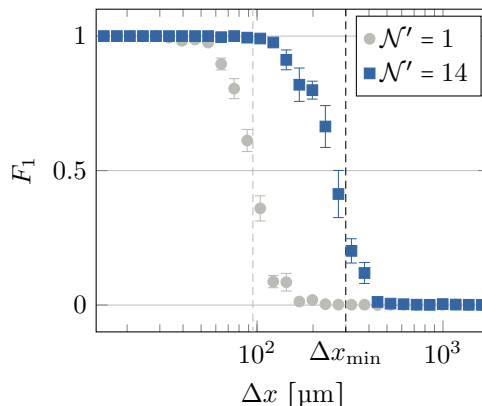


Abbildung 7.9: Die Auswertung der Scores nach Pixelgröße mit der additiven Methode zeigt deutlich, dass die minimal erforderliche Pixelgröße von $\mathcal{N}' = 1$ zu $\mathcal{N}' = 14$ um etwa den Faktor 3,16 und damit nahe $\sqrt{\mathcal{N}'}$ steigt.

Rauschen. Um diese bisher nicht näher fassbare „bessere“ Rekonstruktionsfähigkeit quantifizieren zu können, führen wir einen weiteren, kumulierten Score auf Basis des bereits bekannten F_1 -Scores ein.

Dazu seien alle Rekonstruktionen bei einer festen kinetischen Energie E_{kin} in einem arithmetischen Mittel über alle Scores durch

$$\bar{F}_1(E_{\text{kin}}) = \frac{1}{n_N \cdot n_d} \sum_{\{N\}} \sum_{\{d\}} F_1(N, d; E_{\text{kin}}) \quad (7.31)$$

zusammengefasst. Mit n_N und n_d seien jeweils die Anzahl an berücksichtigten Werten von N und d respektive bezeichnet. Hier gilt also wie oben beschrieben $n_N = 20$ und abweichend für $n_d = 18$, weil zwei Schirmgrößen $d < d_{\min}$ aus den o. g. Gründen für die weiteren Untersuchungen nicht relevant sind.

Um über die erhobenen Daten möglichst gut mitteln zu können, werden die verschiedenen Energiewerte kombinatorisch permutiert.

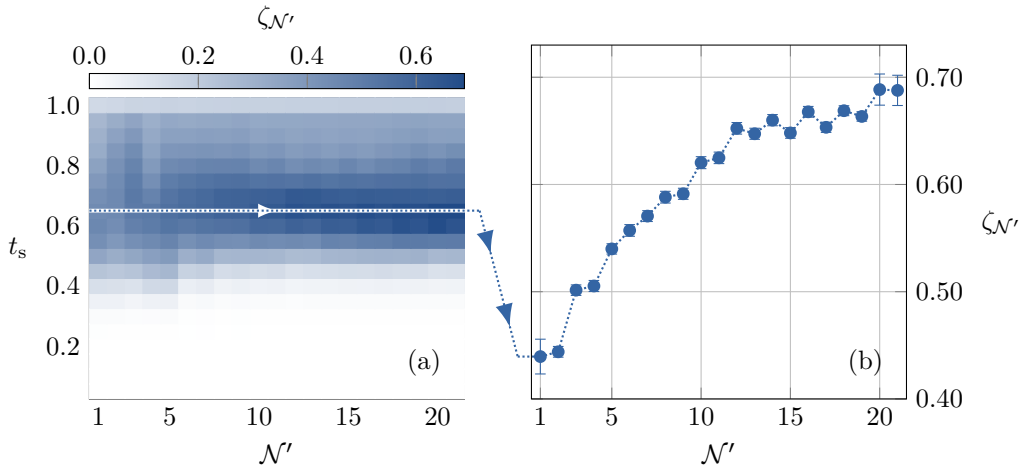


Abbildung 7.10: Beim Gegenüberstellen des gemittelten $\zeta_{\mathcal{N}'}$ gegen die Anzahl an Rekonstruktionsvolumina \mathcal{N}' fällt auf, dass diese zunehmen. (a) Das Histogramm belegt, dass das für alle gezeigten Thresholds t_s annähernd gilt. (b) Wir konzentrieren uns auf den hier optimalen Threshold von $t_s = 0,65$. Nach starkem Anstieg des Scores gelangt die Kurve sukzessive in Sättigung. Angepasste Darstellung nach Ref. [20, Abb. 3] mit neu ausgewerteten Daten.

Wir definieren dafür zunächst die Menge aller verfügbaren Energiemesswerte nach Gleichung (7.27) zu

$$E_{\text{kin}} \in \mathcal{E} := \{10,0 \text{ keV}; 10,5 \text{ keV}; \dots; 20,0 \text{ keV}\}. \quad (7.32)$$

Insgesamt sind also $|\mathcal{E}| = 21$ Energiewerte in dieser Menge enthalten. Für $\mathcal{N}' = 1$ existieren dann 21 kumulierte F_1 -Scores $\bar{F}_1(E_{\text{kin}})$.

Wollen wir nun \mathcal{N}' Energiewerte addieren, so gilt es, die Anzahl der möglichen Kombinationen zu maximieren. So können wir über alle möglichen Additionen mitteln und erhalten ein zuverlässigeres Ergebnis. Eine Kombination von \mathcal{N}' Energien sei im Folgenden mit $p^{\mathcal{N}'}$ bezeichnet, die i -te Energie dieser Folge mit $p_i^{\mathcal{N}'}$. Insgesamt existieren

$$\binom{|\mathcal{E}|}{\mathcal{N}'} = \frac{|\mathcal{E}|!}{\mathcal{N}'! \cdot (|\mathcal{E}| - \mathcal{N}')!} \quad (7.33)$$

verschiedene Kombinationen von Energiewerten, in denen keine ungeachtet der Reihenfol-

ge zwei Mal vorkommt. Sie seien zusammengefasst in der Permutationsmenge $P^{\mathcal{N}'}$, die gemäß

$$p^{\mathcal{N}'} \in P^{\mathcal{N}'} \not\subseteq \mathcal{E}^{\mathcal{N}'} \quad (7.34)$$

eine echte Teilmenge von $\mathcal{E}^{\mathcal{N}'}$ ist.

Damit lässt sich für \mathcal{N}' Energiewerte das arithmetische Mittel über alle Permutationen mit

$$\zeta_{\mathcal{N}'} = \frac{1}{\binom{|\mathcal{E}|}{\mathcal{N}'}} \sum_{p^{\mathcal{N}'} \in P^{\mathcal{N}'}} \frac{1}{\mathcal{N}'} \sum_i \bar{F}_1(p_i^{\mathcal{N}'}) \quad (7.35)$$

ermitteln. Wir suchen zu einer gegebenen Anzahl \mathcal{N}' an zu addierenden Energiemustern also alle Permutationen. Für jede kinetische Energie innerhalb dieser Permutationen werden für alle (N, d) die Rekonstruktionen nach der oben beschriebenen Threshold-Optimierung ermittelt und durch einen F_1 -Score bewertet. Aus Gründen der Übersichtlichkeit wurde das in der mathematischen Beschreibung vernachlässigt und findet implizit in \bar{F}_1

Berücksichtigung. Über alle diese ermittelten Scores wird dann das arithmetische Mittel $\zeta_{\mathcal{N}'}$ gebildet. Damit maximieren wir die Datenbasis, die in die Bildung dieses Mittels einfließt, wodurch es zuverlässiger bestimmt werden kann.

Wird der Threshold nun für alle Rekonstruktionsvolumina festgelegt und in dessen Abhängigkeit $\zeta_{\mathcal{N}'}$ bestimmt, ergibt sich das in Abbildung 7.10a gezeigte Histogramm. Ungeachtet des Thresholds t_s nimmt der gemittelte Score zu. Optimal scheint er bei etwa $t_s = 0,65$ zu sein. Für diesen Spezialfall ist $\zeta_{\mathcal{N}'}$ in Abbildung 7.10b gezeigt.

Deutlich ist eine stete Zunahme des Scores mit steigendem \mathcal{N}' . Zunächst stark, jedoch immer geringer ausfallend, bis der Wert eine Sättigung bei etwa $\zeta_{\mathcal{N}'} = 0,7$ erfährt. Der globale Score zwischen $\mathcal{N}' = 1$ und 21 nimmt dabei etwa um den Faktor 1,75 zu.

Mit den erhobenen Daten können wir einen Rückschluss zu den in Kapitel 3 beschriebenen Apparaturen zur Elektronenrückstreuung ziehen. Dort wurden Experimente beschrieben, die typische Parameter von etwa

$$d^2 = 5 \times 5 \text{ cm}^2, \quad N^2 = 1000 \times 1000 \quad (7.36)$$

als Schirm aufweisen. Alle weiteren Parameter wie Verkipfung φ und Schirmabstand D waren für die gesamte Experimentierreihe bereits aus den in Abschnitt 7.1 beschriebenen, üblichen Parametern übernommen.

Bereits mit Blick auf Abbildung 7.5 ist erkennbar, dass die Rekonstruktion für $d = 5 \text{ cm}$ und $N = 1000$ im Falle von W(100) gut gelingt. Allerdings darf die Pixelzahl nicht deutlich abnehmen, da bereits für $N < 800$ schlechtere Scores realisiert werden, gleiches gilt für

einen unwesentlich größeren Schirm mit einer Kantenlänge von $d > 8 \text{ cm}$ und gleichbleibender Pixelzahl. Theoretisch ist die Elektronenholographie mit heute verwendeten EBSD-Apparaturen also durchführbar.

Deutlich stabiler und zuverlässiger werden die Ergebnisse allerdings durch Verwendung der Additionsmethode im Nachgang. Abbildung 7.8 zeigt eine deutlich höhere Toleranz gegenüber der Parameterwahl. So reichen bereits Pixelzahlen von $N \geq 300$ bei gleicher Schirmgröße aus, um erfolgreich rekonstruieren zu können. Die Schirmgröße selbst sollte bei $N = 1000$ jedoch durch $d < 18 \text{ cm}$ begrenzt sein, was allerdings einer Verdreifachung der ursprünglich möglichen Schirmdimension gleichkommt.

In allen Parametern erhalten wir also verbesserte Ergebnisse durch die Addition verschiedener Muster. Dem steht ein zusätzlicher Aufwand in Aufnahme von Daten und Auswertung mittels der holographischen Algorithmen gegenüber, der linear in \mathcal{N}' skaliert.

Für die hier behandelten Parameter haben wir im Falle der einfachen Rekonstruktion aus nur einem Muster in Abschnitt 7.3.1 f. quantitativ die Mindestanforderungen ausweisen können. Für die deutlich geringeren Parameteranforderungen im Falle der addierten Rekonstruktionen verschiedener Energien liegt eine analoge physikalische Erklärung bisher nicht vor. Gleichwohl haben wir zeigen können, dass die Untergrenzen der einfachen Rekonstruktion niemals unterschritten werden.

8 Zusammenfassung und Ausblick

Damit eine holographische Abbildung von Kristallen gelingen kann, sind bereits aus lightholographischen Überlegungen nach Kapitel 2 einige Grundvoraussetzungen bekannt. So braucht es einen passend konstruierten Prozess mit einer definierten Referenzwelle. Damit diese überhaupt eine Struktur auflösen kann, darf sie die Größenordnung der einzelnen Objektpixel nicht übersteigen.

Die Elektronenholographie ist eine bereits etablierte Methode zur Auflösung atomarer Strukturen mit anwendungsabhängigen Nachteilen. So hat beispielsweise die Holographie mittels Low-Energy Electron Diffraction (LEED) eine geringe räumliche Auflösung [104], bei Verwendung hochenergetischer Elektronen rauschende Störeinflüsse in der Rekonstruktion [18], siehe auch Abschnitt 3.3.

Ferner werden weitere Anforderungen an einen holographischen Aufbau mittels Elektronenwellen gestellt. So muss aus in Abschnitt 7.1 skizzierten Gründen die Referenzwelle nah an den atomaren Streuern liegen. Neben der starken Intensitätsabnahme der Referenz-Kugelwelle führen große Abstände zu Aliasing-Effekten [11, Kap. 14.2.2.2]. Für die Lightholographie konnte im Allgemeinen immer eine isotrope Welle angenommen werden, was auf Elektronenwellen nicht mehr zutrifft [22]. Damit eine Elektronenholographie gelingen kann, ist also zunächst, wie in Abschnitt 3.1.2 beschrieben, eine geeignete Form der Wellenbeschreibung nötig.

Alle diese Voraussetzungen finden sich in der in Abschnitt 3.2 beschriebenen Electron Backscatter Diffraction (EBSD) erfüllt. Sie ist eine etablierte Methode zur Materialanalyse [67, 76, 80] und gibt bereits einen klaren Einblick in Kristalleigenschaften, etwa in die Orientierung seiner Kristalliten [19]. Bei typischen Wellenlängen von $\lambda_B \approx 0,055 \text{ \AA}$

nach Gleichung (3.8) liegt es also nahe, auch EBSD für die Elektronenholographie heranzuziehen. Die EBSD-Holographie wurde dabei bereits für den Spezialfall sphärischer Muster im Rahmen von Ref. [108] vorgestellt. Die experimentelle Umsetzung gestaltete sich bisher aber schwierig [223].

Im Rahmen von Abschnitt 3.3 wurden die Grundlagen beschrieben, die zur Anwendbarkeit der EBSD für die Holographie nötig sind. So lassen sich für die typischen Elektronenenergien von etwa $E_{\text{kin}} = 20 \text{ keV}$ die Elektronenwellen als annähernd kegelförmig mit energieabhängigem Öffnungswinkel Φ beschreiben [64, 65]. Wie in Abschnitt 3.1.2 dargestellt, kommen sie einer isotropen Welle im für die Holographie relevanten Maße hinreichend nahe.

Die holographische Aufnahme im Experiment wird allerdings stark dominiert von der geometrischen Information. Die Rekonstruktion wird deutlich durch deren Eliminierung vereinfacht. Dies gelingt durch eine an die Differential Photoelectron Holography (DPH) nach Ref. [116] angelehnte Methode, der in Abschnitt 3.3.1 ausgeführten Differenzbildmethode. Dabei werden zwei Muster unterschiedlicher kinetischer Energien E_{kin} voneinander subtrahiert [52].

In der Informatik gewinnt die Computerholographie als Fachgebiet immer mehr an Bedeutung. Etwa bei der Entwicklung echter 3D-Displays [17]. Kapitel 4 beschrieb die numerische Holographie, wie sie in diskreten Systemen berechnet werden kann. Weiterhin wurden zu beachtende Grenzen des Auflösungsvermögens und durch Unterabtastung auftretendes Aliasing hergeleitet und aufgezeigt.

Das sich anschließende Kapitel 5 stellte sodann übliche Algorithmen zur Berechnung der Wellenpropagation vor. In der vorliegenden Arbeit wurde eine Bibliothek auf Basis

dieser Prinzipien mit dem Ziel entwickelt, die EBSD-Holographie auf die geforderten Rahmenbedingungen zu portieren. Dabei haben wir gängige Apparaturen mit planem Schirm simuliert, die näher in Abschnitt 6.2 beschrieben sind.

Genauere Messungen wurden in Abhängigkeit mehrerer Apparaturparameter an einem W(100)-Kristall, wie er in Abschnitt 6.1 beschrieben wurde, als beispielhaftem Probensystem durchgeführt. Aufbauend auf die in Abschnitt 6.3 vorgestellten Kriterien zur binären Klassifikation existieren viele Bewertungskriterien für die Übereinstimmung zwischen Rekonstruktion und Kristall, die sich ferner in Anhang C aufgelistet finden.

Durch die Einführung des F_1 -Score in Abschnitt 7.2 als objektives Maß eines Rekonstruktionserfolgs ist es möglich, die Rekonstruktion anhand dieser Parameter zu klassifizieren und in der Folge zu bewerten. Konkrete Simulationen zeigen quantitativ, dass die zuvor theoretisch herleitbaren Grenzen für eine erfolgreiche Rekonstruktion echte physikalische Grenzen darstellen. Im Einzelnen haben wir dabei neben einer minimalen Kantenlänge d_{\min} des Schirms aus geometrischen Überlegungen im konkret vorliegenden Experiment auch allgemeingültige Mindestanforderungen an das Oversampling Ratio σ und eine minimale Pixelgröße Δx_{\min} nach dem Abtasttheorem (4.37) herleiten können.

Durch die in Abschnitt 7.4 beschriebene Addition mehrerer Rekonstruktionen verschiedener kinetischer Energien E_{kin} lassen sich die Ergebnisse deutlich verbessern und rauschende Einflüsse unterdrücken. So sinken die Anforderungen an Oversampling Ratio σ_{\min} und Pixelgröße Δx_{\min} deutlich und die Rekonstruktion wird stabiler.

Die in Kapitel 7 vorgestellten Ergebnisse können also als Anhaltspunkt zum Entwurf eines Experiments zur Holographie mittels EBSD herangezogen werden und zeitgleich Hinweise darauf geben, wie aus diesem Experiment

durch Hinzunahme weiterer Daten zuverlässiger eine genaue Rekonstruktion errechnet werden kann. Dies gelingt durch die Aufnahme mehrerer Muster verschiedener kinetischer Energien. Im Rahmen dieser Arbeit wurden dazu insgesamt 21 verschiedene kinetische Energien zu je 400 Rekonstruktionen berechnet und ausgewertet. Durch Simulation bereits heute verfügbarer Apparaturen zeigt sich, dass eine EBSD-Holographie durchführbar ist und sie deutlich von der Aufnahme zusätzlicher Muster profitiert.

Dabei wurde im Rahmen der theoretischen Beschreibung der EBSD in Abschnitt 3.3 die experimentelle Umsetzbarkeit thematisiert. Naheliegende Zweifel, etwa an der Annahme einer einzelnen Referenzwelle, konnten mithilfe von Abschätzungen über die Wahrscheinlichkeit konkurrierender Streuprozesse zerstreut werden. Ferner stützt die jeweils gleiche lokale Umgebung innerhalb der Kristalliten das Ergebnis und verstärkt die holographische Information [20].

Insbesondere ist aber davon auszugehen, dass die Hologramme in der experimentellen Umsetzung gegenüber der Simulation von einer gleichen lokalen Umgebung sogar profitieren. Muster der gleichen Umgebung mit unterschiedlicher, aber konstanter Phase sind äquivalent zueinander. Die Addition solcher Muster mit zufälliger Phase erhöht damit die Stabilität der Rekonstruktion zusätzlich merklich [224], während dieser Aspekt in der Simulation keine Berücksichtigung fand.

Mit der in Anhang B genauer vorgestellten *Hololib* liegt nun ein universell nutzbares Werkzeug für die Simulation holographischer Experimente vor, das grundsätzlich nicht auf die Elektronenholographie beschränkt ist. Vielmehr ist es damit möglich, experimentelle Aufbauten jedweder Art zu simulieren und jeden Parameter einzeln reproduzierbar zu untersuchen. So ist der Abstand zwischen Probe und Schirm D ein bisher nicht näher beschriebener, kritischer Parameter. Aufgrund

des modularen Aufbaus und Schwerpunktsetzung auf parallelisierte Berechnung insbesondere mithilfe von CPU-Clustern oder einer GPU sind Experimentierreihen leicht skalierbar. Auch die Untersuchung mehrdimensionaler Parameterräume, wie hier mit den zwei Parametern Schirmkantenlänge d und Schirmauflösung N^2 geschehen, sind damit verhältnismäßig schnell umsetzbar.

Auch die Art der Probe kann variiert werden, um zu überprüfen, ob sich die Ergebnisse etwa auf Kristalle mit mehratomigen Basen übertragen lassen. Die Untersuchungen an sphärischen Schirmen nach Ref. [108] legen das allerdings bereits nahe. Die Miteinbeziehung experimenteller Effekte, wie des winkelabhängigen Wirkungsquerschnitts und damit variierender Streuintensität, erlaubt ferner die Untersuchung des Einfallswinkels.

Bisher ist im Rahmen dieser Arbeit von der in Abschnitt 6.1.1 angenommenen, strikten Periodizität der Probe ausgegangen worden. Dies ist insbesondere hinsichtlich des in Abschnitt 6.2.1 erwähnten Arguments der lokalen Umgebung nötig. In nachfolgenden hierauf aufbauenden Arbeiten gilt es nun zu untersuchen, inwieweit Fehlstellen trotzdem zulässig sind, ohne dass die Rekonstruktion

dadurch scheitert. Eine Fehlstelle entspricht im Rahmen des hier vorgestellten Vorgehens dann der Addition von Hologrammen verschiedener lokaler Umgebungen mit jeweils einer Fehlstelle an einem anderen Gitterplatz. In deren Mittel wird kein signifikanter Effekt messbar sein. Allerdings wird ab einer gewissen Zahl an Fehlstellen die Rekonstruktion problematisch. Für eine realistische Umsetzung ist die Annahme solcher Defekte aber unumgänglich. Zu untersuchen ist also weiterhin, welchen Grad an Unreinheit eine Probe aufweisen darf, damit die hier angenommenen Parametergrenzen gültig bleiben oder inwieweit sich Fehlstellen durch Aufnahme weiterer Daten kompensieren lassen. Eine Datenredundanz kann dabei sowohl durch Erhöhung der Pixelzahl N^2 als auch durch die Addition weiterer Muster \mathcal{N}' erreicht werden.

Für zukünftige Forschungen gilt es ferner ebenfalls, die Ergebnisse dieser Arbeit einer experimentellen Untersuchung zu unterziehen. Erfolgversprechende experimentelle Parameter wurden bereits in Abschnitt 7.3 f. vorgestellt. Die Ergebnisse der hier durchgeführten Simulationen lassen sich dabei auch auf andere Methoden, wie etwa das zur Adsorbatmessung genutzte RHEED, übertragen.

Anhang

A Holographischer Prozess

Nicht alle angewendeten Rechnungen haben ihren Platz im Hauptteil der vorliegenden Arbeit gefunden. Gerade für das Verständnis des holographischen Prozesses ist eine oberflächliche Betrachtung unter Vernachlässigung vieler weiterer Effekte hilfreich, denn die Berechnungen werden schnell komplex und unübersichtlich.

Aus diesem Grund sollen im Allgemeinen angewendete Berechnungen an dieser Stelle aufgenommen werden.

A.1 Intensität mehrerer Streuer

Bei der Berechnung der Intensität in Abschnitt 2.3 wurde nur ein Streuzentrum berücksichtigt. Die dort berechnete Intensität soll hier verallgemeinert werden.

Für die Referenzwelle (2.29) sei hier wieder ohne Beschränkung der Allgemeinheit angenommen, dass die Phasenverschiebung $\varphi = 0$ ist und das Referenzwellenzentrum bei $\mathbf{0}$ liegt. Zur Vereinfachung der Berechnungen definieren wir die Parameter α_r für die Amplitude und δ_r für die Phase. Die Referenzwelle lautet dann

$$\psi_r(\mathbf{r}) = \frac{A_r}{kr} e^{-ikr} =: \alpha_r e^{-i\delta_r}. \quad (\text{A.1})$$

Das Huygens'sche Prinzip erlaubt, die Objektwelle verallgemeinert als Summe beliebiger Partikularwellen umzuformulieren [50, Kap. 8.2]. Eine derartige Superposition verschiedener Partikularwellen lässt sich schreiben als

$$\psi_o(\mathbf{r}) = \sum_n \psi_{o,n}(\mathbf{r}). \quad (\text{A.2})$$

Diese Partikularwellen $\psi_{o,n}$ lauten analog zur einzelnen Objektwelle (2.26)

$$\begin{aligned} \psi_{o,n}(\mathbf{r}) &= \psi_r(\mathbf{o}_n) \cdot \frac{a_n}{k|\mathbf{r} - \mathbf{o}_n|} e^{-i(k|\mathbf{r} - \mathbf{o}_n| - \varphi_n)} = \underbrace{\frac{A_r \cdot a_n}{k \cdot o_n}}_{=: A_n} \cdot \frac{1}{k|\mathbf{r} - \mathbf{o}_n|} e^{-i[k(|\mathbf{r} - \mathbf{o}_n| + o_n) - \varphi_n]} \\ &= \frac{A_n}{k|\mathbf{r} - \mathbf{o}_n|} e^{-i[k(|\mathbf{r} - \mathbf{o}_n| + o_n) - \varphi_n]} =: \alpha_n e^{-i\delta_n}. \end{aligned} \quad (\text{A.3})$$

Dabei ist a_n ein streuerabhängiger Amplitudenfaktor, \mathbf{o}_n der Ort des Streuzentrums. Die definierten Parameter α_n und δ_n bezeichnen auch hier respektive die Amplituden und Phasen der beteiligten Exponentialfunktionen und vereinfachen so die weiteren Berechnungen.

Die Intensität auf dem Schirm lautet dann

$$\begin{aligned}
 I(\mathbf{s}) &= |\psi_r(\mathbf{s}) + \psi_o(\mathbf{s})|^2 = \left| \psi_r(\mathbf{s}) + \sum_n \psi_{o,n}(\mathbf{s}) \right|^2 = \left| \alpha_r e^{-i\delta_r} + \sum_n \alpha_n e^{-i\delta_n} \right|^2 \\
 &= \left(\alpha_r e^{-i\delta_r} + \sum_n \alpha_n e^{-i\delta_n} \right) \cdot \left(\alpha_r e^{+i\delta_r} + \sum_n \alpha_n e^{+i\delta_n} \right) \\
 &= \alpha_r^2 e^{-i\delta_r} e^{+i\delta_r} + \alpha_r e^{-i\delta_r} \sum_n \alpha_n e^{+i\delta_n} + \alpha_r e^{+i\delta_r} \sum_n \alpha_n e^{-i\delta_n} + \sum_n \sum_m \alpha_n \alpha_m e^{-i\delta_n} e^{-i\delta_m} \quad (\text{A.4}) \\
 &= \alpha_r^2 + \underbrace{\alpha_r \sum_n \alpha_n e^{-i(\delta_r - \delta_n)} + \alpha_r \sum_n \alpha_n e^{-i(\delta_n - \delta_r)}}_{=:\chi_1} + \underbrace{\sum_n \sum_m \alpha_n \alpha_m e^{-i(\delta_n - \delta_m)}}_{=:\chi_2}.
 \end{aligned}$$

Beide Terme χ_1 und χ_2 werden getrennt voneinander betrachtet

$$\chi_1 = \sum_n \alpha_n e^{-i(\delta_r - \delta_n)} + \sum_n \alpha_n e^{+i(\delta_r - \delta_n)} = \sum_n \alpha_n \cdot 2 \cos(\delta_r - \delta_n), \quad (\text{A.5a})$$

$$\begin{aligned}
 \chi_2 &= \sum_n \sum_m \alpha_n \alpha_m e^{-i(\delta_n - \delta_m)} = \sum_n \alpha_n^2 + \sum_{n \neq m} \alpha_n \alpha_m e^{-i(\delta_n - \delta_m)} \quad (\text{A.5b}) \\
 &= \sum_n \alpha_n^2 + \sum_{n \neq m} \alpha_n \alpha_m (e^{-i(\delta_n - \delta_m)} + e^{-i(\delta_m - \delta_n)}) = \sum_n \alpha_n^2 + 2 \sum_{n < m} \alpha_n \alpha_m \cos(\delta_n - \delta_m).
 \end{aligned}$$

Insgesamt ergibt sich die allgemeine Intensität damit zu

$$I = \alpha_r^2 + 2\alpha_r \sum_n \alpha_n \cos(\delta_r - \delta_n) + \sum_n \alpha_n^2 + 2 \sum_{n < m} \alpha_n \alpha_m \cos(\delta_n - \delta_m). \quad (\text{A.6})$$

A.2 Allgemeine Rekonstruktion

Ganz allgemein lässt sich die Rekonstruktion auch aus den beteiligten Wellenfunktionen phänomenologisch herleiten. Die Intensität lautet im Allgemeinen

$$\begin{aligned}
 I(\mathbf{s}) &= |\psi_r(\mathbf{s}) + \psi_o(\mathbf{s})|^2 = \left| \psi_r(\mathbf{s}) + \sum_n \psi_{o,n}(\mathbf{s}) \right|^2 \\
 &= \left(\psi_r(\mathbf{s}) + \sum_n \psi_{o,n}(\mathbf{s}) \right) \cdot \left(\psi_r^*(\mathbf{s}) + \sum_n \psi_{o,n}^*(\mathbf{s}) \right) \quad (\text{A.7}) \\
 &= \psi_r(\mathbf{s}) \psi_r^*(\mathbf{s}) + \psi_r(\mathbf{s}) \sum_n \psi_{o,n}^*(\mathbf{s}) + \psi_r^*(\mathbf{s}) \sum_n \psi_{o,n}(\mathbf{s}) + \sum_n \sum_m \psi_{o,n}(\mathbf{s}) \psi_{o,m}^*(\mathbf{s}).
 \end{aligned}$$

Im Folgenden wird auf die Angabe des Ortsvektors als Wellenfunktionsargument aus Gründen der Übersichtlichkeit verzichtet. Bei der Beleuchtung mit der Referenzwelle folgt entsprechend

$$\begin{aligned}
 I \cdot \psi_r &= \psi_r \psi_r^* \psi_r + \psi_r \sum_n \psi_{o,n}^* \psi_r + \psi_r^* \sum_n \psi_{o,n} \psi_r + \sum_n \sum_m \psi_{o,n} \psi_{o,m}^* \psi_r \\
 &= \left(|\psi_r|^2 + \sum_n \psi_{o,n} \sum_m \psi_{o,m}^* \right) \psi_r + \psi_r^2 \sum_n \psi_{o,n}^* + |\psi_r|^2 \sum_n \psi_{o,n} \quad (\text{A.8}) \\
 &= \left(|\psi_r|^2 + \psi_o \psi_o^* \right) \psi_r + \psi_r^2 \psi_o^* + |\psi_r|^2 \psi_o = \left(|\psi_r|^2 + |\psi_o|^2 \right) \psi_r + \psi_r^2 \psi_o^* + |\psi_r|^2 \psi_o.
 \end{aligned}$$

Auch hier sind analog zu Gleichung (2.40) wieder drei Terme ersichtlich: Eine gedämpfte Referenzwelle im ersten Summanden, ein konjugiertes Bild im zweiten und schließlich die eigentliche Objektwelle, skaliert mit der Intensität der Referenzwelle.

A.3 Rückpropagation

In einem lightholographischen Experiment würde eine Linse genutzt werden, um das rekonstruierte Bild auf einem Schirm sichtbar zu machen. Im Falle einer Computersimulation ist die Imitation dieses Verhaltens durch zusätzliche Berechnungen nicht erforderlich. Sie führt ansonsten lediglich zu zusätzlicher Komplexität.

Stattdessen wird eine Rekonstruktionswelle definiert, die rückwärtsgewandt in das ursprüngliche Objektvolumen läuft. Rücklaufend bedeutet, dass sie das gegenteilige Vorzeichen zur Objektwelle in der Phase trägt. Nachdem der Schirm erneut mit der Referenzwelle beleuchtet wurde, lässt sich jedem Punkt auf ihm eine komplexe Amplitude zuordnen. Nach dem Huygens'schen Prinzip ist die Rekonstruktionswelle wieder als Summe vieler Partikularwellen mit den genannten Amplituden als

$$\psi_{\text{rek}}(\mathbf{r}) = \iint_{\mathbf{s}} I(\mathbf{s}') \cdot \psi_{\text{r}}(\mathbf{s}') \cdot \frac{1}{k|\mathbf{r} - \mathbf{s}'|} e^{i(k|\mathbf{r} - \mathbf{s}'| - \varphi_{\text{rek}})} d\mathbf{s}' \quad (\text{A.9a})$$

zu verstehen.

Im Falle eines diskreten Aufnahmegitters, wie es in Simulationen oder Experimenten mit Halbleitersensoren der Fall ist, kann dieses Integral als diskrete Summe über die Menge aller Pixel $\{\mathbf{s}\}$ ausgedrückt werden. Die Rekonstruktionswelle lautet dann

$$\psi_{\text{rek}}(\mathbf{r}) = \sum_{\{\mathbf{s}\}} I(\mathbf{s}) \cdot \psi_{\text{r}}(\mathbf{s}) \cdot \frac{1}{k|\mathbf{r} - \mathbf{s}|} e^{i(k|\mathbf{r} - \mathbf{s}| - \varphi_{\text{rek}})} =: \sum_{\{\mathbf{s}\}} \psi_{\text{rek},\mathbf{s}}(\mathbf{r}). \quad (\text{A.9b})$$

Für die weitere phänomenologische Untersuchung seien die konstanten Phasenverschiebungen vernachlässigt. Ferner seien nun die einzelnen Summanden

$$\psi_{\text{rek},\mathbf{s}} \stackrel{(2.40)}{=} \left[\underbrace{\tilde{A}(\mathbf{s}) \cdot \psi_{\text{w}}(\mathbf{s})}_{\chi_1} + \underbrace{\hat{A}(\mathbf{s}) e^{ik(|\mathbf{s} - \mathbf{o}| - 2s)}}_{\chi_2} + \underbrace{\hat{A}(\mathbf{s}) e^{-ik|\mathbf{s} - \mathbf{o}|}}_{\chi_3} \right] \cdot \frac{1}{k|\mathbf{r} - \mathbf{s}|} e^{i(k|\mathbf{r} - \mathbf{s}| - \varphi_{\text{rek}})} \quad (\text{A.10})$$

in den Mittelpunkt der Untersuchung gestellt.

Der Term χ_3 war bereits in Abschnitt 2.3.1 als derjenige ausgemacht, der zum tatsächlichen Bild beiträgt. Die wesentliche Ortsabhängigkeit sei als $\xi_3(\mathbf{s}, \mathbf{r})$ isoliert zu

$$\frac{\chi_3}{k|\mathbf{r} - \mathbf{s}|} e^{i(k|\mathbf{r} - \mathbf{s}| - \varphi_{\text{rek}})} = \frac{A_{\text{o}} A_{\text{r}} A_{\text{w}}}{k^3 |\mathbf{s} - \mathbf{o}| s^2} e^{-ik|\mathbf{s} - \mathbf{o}|} \cdot \frac{1}{k|\mathbf{r} - \mathbf{s}|} e^{ik|\mathbf{s} - \mathbf{r}|} \propto e^{-ik(|\mathbf{s} - \mathbf{o}| - |\mathbf{s} - \mathbf{r}|)} =: \xi_3(\mathbf{s}, \mathbf{r}). \quad (\text{A.11})$$

Von der Rekonstruktionswelle (A.9b) wird im Falle der Rekonstruktion auch die Intensität bestimmt. Sie lautet

$$\Gamma(\mathbf{r}) = |\psi_{\text{rek}}(\mathbf{r})|^2 = \psi_{\text{rek}}(\mathbf{r}) \cdot \psi_{\text{rek}}^*(\mathbf{r}). \quad (\text{A.12})$$

Aus diesem Betragsquadrat ergeben sich bei N^2 Pixeln auf dem Schirm und damit N^2 Streuzentren, die es zu berücksichtigen gilt, insgesamt $(3 \cdot N^2)^2$ Summanden, die jeweils auf die drei beteiligten Terme χ_1 , χ_2 und χ_3 in Gleichung (A.11) zurückzuführen sind.

Sei nun von lediglich zwei Orten auf dem holographischen Schirm σ und ρ ausgegangen, so erhalten wir im Betragsquadrat insbesondere zwei Summanden

$$\xi_3(\sigma, \mathbf{r}) \cdot \xi_3^*(\rho, \mathbf{r}) = \exp[-ik(|\sigma - \mathbf{o}| - |\sigma - \mathbf{r}| - |\rho - \mathbf{o}| + |\rho - \mathbf{r}|)] \quad \text{und} \quad (\text{A.13a})$$

$$\xi_3^*(\sigma, \mathbf{r}) \cdot \xi_3(\rho, \mathbf{r}) = \exp[-ik(|\sigma - \mathbf{r}| - |\sigma - \mathbf{o}| - |\rho - \mathbf{r}| + |\rho - \mathbf{o}|)] \quad (\text{A.13b})$$

aus der Multiplikation der Objektwellenterme, hier reduziert auf ξ_3 .

Es ist ersichtlich, dass diese beiden Terme für $\mathbf{r} = \mathbf{o}$ maximal werden, am Ort des eigentlichen Objekts also eine Rekonstruktion ermöglichen.

Analog zu ξ_3 definieren wir nach dem Ansatz (A.11) auch ξ_1 und ξ_2 zu

$$\xi_1(\mathbf{s}, \mathbf{r}) := \exp[-ik(s - |\mathbf{s} - \mathbf{r}|)] \quad \text{und} \quad (\text{A.14a})$$

$$\xi_2(\mathbf{s}, \mathbf{r}) := \exp[-ik(2s - |\mathbf{s} - \mathbf{o}| - |\mathbf{s} - \mathbf{r}|)]. \quad (\text{A.14b})$$

Die Summanden der Form $\xi_i \xi_i^*$ sind nicht weiter relevant, da sie keine Ortsabhängigkeit aufweisen. Für sie gilt ganz allgemein

$$\xi_i(\mathbf{s}, \mathbf{r}) \cdot \xi_i^*(\mathbf{s}, \mathbf{r}) = |\xi_i(\mathbf{s}, \mathbf{r})|^2 = 1, \quad i \in \{1; 2; 3\}. \quad (\text{A.15})$$

Dabei kann für \mathbf{s} sowohl σ als auch ρ eingesetzt werden, sodass sich die Zahl der von Gleichung (A.15) betroffenen Summanden auf sechs verdoppelt.

In zwölf weiteren Summanden ist keine Abhängigkeit in \mathbf{r} erkennbar. Sie lauten

$$\xi_1(\mathbf{s}, \mathbf{r}) \cdot \xi_2^*(\mathbf{s}, \mathbf{r}) = e^{-ik(s - |\mathbf{s} - \mathbf{r}| - 2s + |\mathbf{s} - \mathbf{o}| + |\mathbf{s} - \mathbf{r}|)} = \exp[-ik(|\mathbf{s} - \mathbf{o}| - s)], \quad (\text{A.16a})$$

$$\xi_1(\mathbf{s}, \mathbf{r}) \cdot \xi_3^*(\mathbf{s}, \mathbf{r}) = e^{-ik(s - |\mathbf{s} - \mathbf{r}| - |\mathbf{s} - \mathbf{o}| + |\mathbf{s} - \mathbf{r}|)} = \exp[-ik(s - |\mathbf{s} - \mathbf{o}|)], \quad (\text{A.16b})$$

$$\xi_2(\mathbf{s}, \mathbf{r}) \cdot \xi_1^*(\mathbf{s}, \mathbf{r}) = e^{-ik(2s - |\mathbf{s} - \mathbf{o}| - |\mathbf{s} - \mathbf{r}| - s + |\mathbf{s} - \mathbf{r}|)} = \exp[-ik(s - |\mathbf{s} - \mathbf{o}|)], \quad (\text{A.16c})$$

$$\xi_2(\mathbf{s}, \mathbf{r}) \cdot \xi_3^*(\mathbf{s}, \mathbf{r}) = e^{-ik(2s - |\mathbf{s} - \mathbf{o}| - |\mathbf{s} - \mathbf{r}| - |\mathbf{s} - \mathbf{o}| + |\mathbf{s} - \mathbf{r}|)} = \exp[-ik(2s - 2|\mathbf{s} - \mathbf{o}|)], \quad (\text{A.16d})$$

$$\xi_3(\mathbf{s}, \mathbf{r}) \cdot \xi_1^*(\mathbf{s}, \mathbf{r}) = e^{-ik(|\mathbf{s} - \mathbf{o}| - |\mathbf{s} - \mathbf{r}| - s + |\mathbf{s} - \mathbf{r}|)} = \exp[-ik(|\mathbf{s} - \mathbf{o}| - s)] \quad \text{und} \quad (\text{A.16e})$$

$$\xi_3(\mathbf{s}, \mathbf{r}) \cdot \xi_2^*(\mathbf{s}, \mathbf{r}) = e^{-ik(|\mathbf{s} - \mathbf{o}| - |\mathbf{s} - \mathbf{r}| - 2s + |\mathbf{s} - \mathbf{o}| + |\mathbf{s} - \mathbf{r}|)} = \exp[-ik(2|\mathbf{s} - \mathbf{o}| - 2s)]. \quad (\text{A.16f})$$

Sie sind damit also auch nicht für die Rekonstruktion relevant und tragen lediglich zu einem Untergrund bei.

Bei Betrachtung des Summanden aus reinen Referenzwellenbestandteilen

$$\xi_1(\sigma, \mathbf{r}) \cdot \xi_1^*(\rho, \mathbf{r}) = \exp[-ik(\sigma - |\sigma - \mathbf{r}| - \rho + |\rho - \mathbf{r}|)] \quad (\text{A.17})$$

fällt sofort auf, dass er für $\mathbf{r} = \mathbf{0}$ maximiert wird. Die Referenzwelle weist an ihrem Ursprung eine wesentliche Dominanz in der Rekonstruktion auf. Da σ und ρ vertauschbar sind, gilt das analog auch für $\xi_1^*(\sigma, \mathbf{r}) \cdot \xi_1(\rho, \mathbf{r})$.

Schwerer zu fassende Kreuzterme kommen durch die Überlagerung der beiden unterschiedlichen Wellen verschiedener Schirmpositionen zustande und sind nicht weiter aufzulösen. Sie tragen nicht zum Verständnis bei und werden daher hier nicht weiter behandelt.

B Implementierung

Im Zuge der vorliegenden Arbeit ist eine umfassende Bibliothek mit Namen *Hololib* entstanden, die die Berechnung von holographischen Prozessen im Rahmen der Skriptsprache *Python 3* [225] ermöglicht und soweit abstrahiert, dass ein möglichst natürlicher Umgang mit relevanten Objekten möglich ist. Das vorliegende Kapitel soll die Grundideen und wesentlichen Bestandteile skizzieren, für eine ausgiebige Nutzungshilfe sei auf die reichhaltige Dokumentation der einzelnen Bestandteile innerhalb des Quellcodes verwiesen.

Die *Hololib* erweitert bereits etablierte wissenschaftliche Werkzeuge, die den Umgang mit Daten vereinfachen. Insbesondere werden viele Berechnungen in die Bibliothek *numpy* ausgelagert, die große Datenmengen durch auf Fortran basierenden Methoden schnell berechnen kann [226]. Als Abstraktionsebene über deren Datenstruktur wird die Bibliothek *xarray* für die eigentliche Organisation von Wellenfeldern eingesetzt. Hier können einzelnen Werten Koordinaten zugewiesen werden, dem gesamten Datenfeld zusätzlich auch Metadaten [227]. Beispielsweise die in Abschnitt 5.4.2 beschriebene Orientierung wird in diesen Metadaten zwischengespeichert. Für den Austausch zwischengespeicherter Wellenfelder wird das eigentlich in den Geowissenschaften etablierte, binäre Dateiformat *netCDF4* genutzt [228]. Weitere standardisierte Berechnungen, wie etwa eine euklidische Norm, sowie den Umgang mit Si-Einheiten ermöglicht die Nutzung der Bibliothek *scipy* [229].

Ein wesentlicher Bestandteil der Bibliothek stellt die Auslagerung der Berechnung der Wellenpropagation in parallelisierte Bestandteile dar. Eine Optimierung auf die vorliegende Hardware nach den in Abschnitt 5.3 beschriebenen Kriterien wird vorgenommen. Dafür sind zwei gleichwertige Berechnungswege implementiert, beide basieren auf dem Just-In-Time-Compiler *numba* [230]. Existiert eine CUDA-fähige Grafikkarte, wird standardmäßig der Teil aktiv, der auf der GPU rechnet. Dabei wird auch möglichst effizient frühzeitig in den Speicher der GPU kopiert und möglichst spät ein Ergebnis zurück in die CPU kopiert.

Fehlt eine kompatible GPU, wird ein zweiter Bestandteil aktiv, der versucht, möglichst effektiv auf der CPU zu parallelisieren, indem die Berechnungen der einzelnen Pixel in einzelnen Threads getrennt voneinander ausgeführt werden. Liegt auch dafür keine kompatible Hardware vor, wird automatisch iterativ mit entsprechend deutlich längerer Laufzeit gerechnet. Damit ist die Bibliothek maximal portabel und ließe sich auch nachträglich an andersartige Hardware durch Implementierung weiterer Berechnungskomponenten anpassen.

Für die Modellierung der Objekte steht zusätzlich zu Hilfsmethoden zur Konstruktion eines Wellenfelds (s. u.) auch eine eigene Bibliothek zur Verfügung, die ihrerseits die Komponente *crystals* des deutlich umfangreichen, für die Arbeit mit Elektronenbeugungsdaten entwickelte Projekts *scikit-ued* [231] erweitert.

Im Folgenden sollen die einzelnen Bestandteile der Bibliothek in logischer Reihenfolge kurz skizziert werden, um die Ausführungen in Kapitel 5 ff. zu vertiefen. Hierbei werden im Hauptteil dieser Arbeit genannte Aspekte mitunter wiederholt genannt, da dieses Kapitels selbstkonsistent und als Referenz für spätere Arbeit an der Bibliothek dienend sein soll.

B.1 WaveObject

Diese abstrakte Klasse fasst alle wellenartigen Objekte zusammen, die in irgendeiner Weise eine Propagation ermöglichen. Dafür steht eine abstrakte versteckte Methode zur Verfügung, die von anhängigen Wellenobjekten zu implementieren ist. Sie wird durch eine öffentliche Implementierung des Propagationsoperators zur Verfügung gestellt und kümmert sich insbesondere darum, dass eine Rückwärtspropagation möglich ist. Das Objekt lässt sich über den Invertierungsoperator Pythons `~` so markieren, dass aus ihr folgende Propagationen rückwärts ausgeführt werden sollen. Zur Zeit existieren die drei folgenden Implementierungen des Wellenobjekts.

B.1.1 Refwave

Die Referenzwelle wird durch ein Objekt simuliert, das zeitgleich ebene (Abschnitt 2.2.1) und sphärische (Abschnitt 2.2.2) Wellen abbilden kann. Ein Typ-Parameter eines eigens dafür eingeführten Datentyps `Refwave.TYPE` bestimmt dabei die konkrete Wellenart. Ferner werden Position und Amplitude des Erregerzentrums gespeichert. Im Falle der ebenen Welle hat die Ausgangsposition keine Bewandtnis. Über Attributmethoden lassen sich Wellenlänge, Wellenvektorbetrag und Wellenvektorrichtung aus dem eigentlich gespeicherten Wellenvektor abfragen und ihn durch Änderung dieser abgeleiteten Parameter ebenfalls anpassen.

Hilfsmethoden stehen zur Erzeugung einer Referenzwelle aus einer (Elektronen-)Energie nach (3.1) und umgekehrt zur Ermittlung der Teilchenenergie bereit. Da die Referenzwelle bei Anwendung auf ein Wellenfeld dort immer für deren weitere Propagation in Metadaten gespeichert wird, ist auch der Export in ein dafür kompatibles Format und umgekehrt die Erzeugung eines Referenzwellenobjekts aus diesem Format implementiert.

Abschließend können als Konstanten typische Wellen zur Verfügung gestellt werden. So modelliert `Refwave.RED` eine rote Kugelwelle mit der Wellenlänge des roten HeNe-Lasers mit $\lambda = 633\text{ nm}$, wie sie auch für die lichthographischen Beispielgrafiken dieser Arbeit genutzt wurde.

B.1.2 Wavefield

Jedes Objekt, Bild oder Rekonstruktionsvolumen wird zur Berechnung als Wellenfeld abgebildet. Es besteht im Wesentlichen aus einem Feld komplexer Amplituden- und Phasenwerte, denen eine Koordinate auf einem kubischen Gitter zugeordnet ist. Das Wellenfeld ist dabei in drei Dimensionen implementiert, aber auch für zweidimensionale Objekte, wie etwa flache Schirme, anwendbar. In Metadaten werden Informationen über die Referenzwelle und weitere Einstellungen abgespeichert. So kann ein Wellenfeld eine `Orientation` (B.5) besitzen. Über Attribute ist das Ermitteln einer Grobumschreibung des Volumens möglich: In möglichst kompakter Form werden Ausdehnung und Auflösung in alle drei Richtungen mit der Orientierung ausgegeben. Bei einer Propagation wird dann diese Grobinformation über das Ausgangsvolumen im Zielvolumen ebenfalls abgespeichert, was die spätere Definition eines Rekonstruktionsvolumens erleichtert. Das Ausgangsvolumen kann dann ohne weitere Parameter oder Bearbeitung aus einem Attribut als leeres Wellenfeld unter Annahme äquidistanter Pixelabstände ausgelesen werden.

Weitere Attribute bilden Real- und Imaginärteile, Intensität und normierte Intensität, sowie die komplexen Wellenfeldwerte direkt ab. Die in den Metadaten hinterlegte Referenzwelle lässt sich ebenfalls durch ein Attribut konstruiert ermitteln. Eine Liste aller Koordinaten, auf die die genannte Orientierung bereits angewandt wurde, erleichtert die parallelisierte Berechnung aller Pixel.

Zum vereinfachten Rechnen mit Wellenfeldern wurden dabei viele der Standardoperatoren Pythons überladen. So ermittelt der Betragsoperator die Beträge der Wellenfeldwerte, drei verschiedene Arten des Additionsoperators berechnen die Propagation eines fremden Wellenobjekts ins eigene Volumen. Die drei Multiplikationsoperatortypen simulieren eine Beleuchtung der eigenen Pixel mit einer anderen Welle und multiplizieren deren Amplituden mit den eigenen. Analog sind auch Subtraktion und Division definiert.

Die eigentliche Propagation ist nach punktbasierter Methode unter Annahme von Kugelwellen und den oben genannten Optimierungen implementiert. Weitere Methoden erlauben das Kopieren, Maskieren (also das Reduzieren eines größeren Volumens auf die Größe eines kleineren) und Speichern in eine Datei. Klassenmethoden erlauben die Konstruktion eines Wellenfelds aus der Angabe von Grenzen für ein kubisches Volumen, insbesondere für lightholographische Experimente aus Bildern, einer zuvor gespeicherten Datei, einer Liste von Streuern oder als Kopie mit einem konstanten Wert mit gleichen Koordinaten eines anderen Wellenfelds.

B.1.3 ConicWavefield

Das konische Wellenfeld erweitert das klassische `Wavefield` (B.1.2) insbesondere um die konische Ausbreitung, wie sie in Abschnitt 3.1.2 beschrieben wurde. Dazu wird zusätzlich zu den bereits genannten Attributen ein Ausbreitungswinkel Φ des Kegels definiert. Über ein weiteres Attribut wird der Ausgangspunkt der Referenzwelle direkt verfügbar und anpassbar gemacht.

Um den dadurch festgesetzten Verbindungsvektor vom Referenzwellenursprung zum Emitter, der durch einen Pixel mit entsprechenden Koordinaten in diesem Wellenfeld abgebildet wird, legt sich ein virtueller Kegel mit oben definiertem Winkel Φ . Innerhalb dieses Kegels wird die isotrope Wellenausbreitung wie in der Kugelwelle angenommen, außerhalb wird sie als Null angenommen.

In diesem Zusammenhang wurden einige weitere Hilfsmethoden implementiert. So ist das konische Wellenfeld in der Lage, zu ermitteln, welche Emitter überhaupt auf einen übergebenen Schirm abbilden würden und kann das Volumen dann bei Bedarf verkleinern. Das spart bei Berechnung der Propagation erheblich Zeit. Ferner können auch die Streuer ermittelt werden, deren Muster ein Hauptmaximum auf dem Schirm abbildet. Ist kein Hauptmaximum sichtbar, wird im Nachgang auch keine Rekonstruktion mehr möglich sein. Zur Optimierung der Datenaufnahme und -speicherung können auch lediglich die durch Streuer hervorgerufenen Modulationen ermittelt werden. Werden sie maximiert, können Ungenauigkeiten in der Datenaufnahme durch nicht optimale Quantisierung kompensiert werden.

B.2 accessors

In diesem Paket werden sogenannte *Accessoren* zusammengefasst. Dabei handelt es sich um selbst erstellte Objekte zur Erweiterung der Bibliothek *xarray*. Sie werden im Rahmen dieser Arbeit insbesondere für die Datenauswertung genutzt. Die einzelnen Funktionen sollen hier nur kurz beschrieben werden.

Mit dem `accessors.ActualAccessor` wird eine Datensammlung als Ergebnis einer Messung ausgewiesen. Damit ist ein anschließender Vergleich möglich, insbesondere stellt er bereits Methoden zur binären Klassifizierung bereit.

Der `accessors.ExpectedAccessor` markiert Daten als erwartetes oder tatsächliches Ergebnis. Er ermöglicht einen späteren Vergleich, stellt ebenfalls Methoden zur binären Klassifizierung bereit und erlaubt insbesondere die Einführung des in Abschnitt 6.3.4 beschriebenen Kugelradius r_a als Zugeständnis an experimentelle Ungenauigkeiten individuell um jede zuvor simulierte Streuerposition.

Der `accessors.ConicExpectedAccessor` erweitert den letztgenannten Accessor um einen weiteren Radius: Statt einer Kugel werden hier zwei Radien entlang und senkrecht zur Verbindungsachse zwischen Referenzwellenzentrum und Streuerposition zur Konstruktion von Ellipsoiden definiert. Gerade für die Auswertung der `ConicWavefield` (B.1.3) ist dies praktisch, da bei tiefen Kristallen Messwerte entlang dieser Achsen „ausschmieren“. Intern werden beide Radien durch eine komplexe Zahl je Streuerposition beschrieben.

Die bisher genannten Accessoren lassen sich mithilfe eines `accessors.CompareAccessors` in Relation setzen und vergleichen. So ist eine einfache binäre Auswertung möglich. Insbesondere werden hier u. a. die Klassifizierungsgrößen aus der Wahrheitsmatrix nach Abschnitt 6.3 implementiert. Es werden weiterhin die wichtigsten Größen, die sich in Anhang C gesammelt finden, bereitgestellt. Der in Abschnitt 7.2 als für diese Arbeit wichtig herausgestellte F_1 -Score, der zur Berechnung eines Punktwerts für eine Rekonstruktion herangezogen wird, wird hier ebenfalls umgesetzt. Es sei angemerkt, dass für die schnelle Konstruktion eines solchen Vergleichsaccessors Hilfsmethoden in allen zuvor genannten Accessoren und den Wellenfeldern zur Verfügung stehen.

Der `accessors.HoloPlotAccessor` stellt einige Methoden zur schnellen Konstruktion und Speicherung von Grafiken der Hologramme und Rekonstruktionsvolumen bereit. Im einzelnen ermöglicht er die Erstellung von Histogrammen, eines `CyclePlot` (B.3) und Streudiagrammen. Ferner erleichtert er das Einzeichnen der Atomkugeln mit Radius r_a in bestehende Streudiagramme.

Im `accessors.HoloAccessor` werden viele Hilfsmethoden zusammengefasst, die an anderer Stelle in der Arbeit mit der zugrundeliegenden Datenstruktur verwendet werden. So können Normen, (quadrierte) Punktabstände und Masken berechnet werden. Ferner stehen zur Datenauswertung ein Fourier-Hochpass zur Minimierung eines homogenen Untergrunds – etwa durch die Referenzwelle – und eine Methode zur Bestimmung eines Wertgradienten zur Verfügung.

B.3 CyclePlot

Ein Rekonstruktionsvolumen ist vor Anwendung einer Klassifizierungsfunktion schwer zu visualisieren. In einem dreidimensionalen Koordinatensystem erhält jeder Punkt einen Zah-

lenwert, den es darzustellen gilt. Diese Klasse versucht die Visualisierung dahingehend zu ermöglichen, dass sie das Rekonstruktionsvolumen virtuell in Scheiben schneidet. Jede Scheibe entspricht einer Ebene mit festem Wert einer beliebigen Dimension, die verbleibenden zwei Dimensionen werden mithilfe eines Histogramms dargestellt. Die Klasse selbst ermöglicht nun das scheibenweise Abspeichern in einzelne Bilder oder die interaktive Visualisierung, bei der das Drücken von Pfeiltasten das beidseitige Durchgehen der einzelnen Ebenen am Bildschirm ermöglichen.

B.4 crystals

Dieser Bibliotheksbestandteil hält Objekte und Methoden zur Modellierung von Kristallstrukturen bereit und basiert dabei weitgehend auf dem gleichnamigen Bestandteil des Pakets *scikit-ued* [231].

In Abschnitt 3.1.1 wurde die mittlere freie Weglänge λ_{IMFP} in Form von Gleichung (3.2) eingeführt, die auch durch den Parameter N_V bestimmt ist. Der dort als Zahl an Valenzelektronen eingeführte Parameter entspricht genauer eigentlich einer Liste an festen Konstanten, deren Entsprechung sich nur in erster Näherung mit der Valenzelektronenzahl identifizieren lässt. Dieser Bibliotheksbestandteil stellt dafür eine genaue Datenbank nach Ref. [61] zur Verfügung.

Ferner werden zwei folgende Objekte als Erweiterungen zur Kristallbeschreibung implementiert.

B.4.1 crystals.Crystal

Die Klasse zur Modellierung von Kristallen erweitert die gleichnamige Klasse aus *scikit-ued* um einzelne, funktionale Attribute. Im Einzelnen ist der schnelle Zugriff auf die Gitterkonstante, atomare Masse einer Einheitszelle, Dichte und die Zahl der Valenzelektronen wie oben beschrieben möglich. Diese Parameter erleichtern alle die Berechnung der mittleren freien Weglänge λ_{IMFP} nach Gleichung (3.2), für deren Bestimmung ebenfalls eine Methode existiert. Weiterhin ist für den leichteren Zugriff eine Methode definiert, die den Inelastic Mean Free Path (IMFP) direkt als Vielfaches der Gitterkonstanten angibt.

Zur eigentlichen Berechnung muss die Einheitszelle in eine größere Struktur erweitert werden. Über eine Methode lässt sich die Einheitszelle in eine „Supercelle“ expandieren.

B.4.2 crystals.SuperCell

Die Erweiterung der gleichnamigen Klasse aus *scikit-ued* war insbesondere nötig, da die Liste der Atome dort unveränderbar war. Zum Einbringen von Fehlstellen ist eine solche Veränderung aber nötig. Daher implementiert die hier vorgestellte Klasse die Speicherung in einer bedingt durch die Problemstellung ineffizienteren, dafür aber veränderlichen Datenstruktur neu.

Ferner stellt sie Methoden zum Hinzufügen und Entfernen von einzelnen Atomen zur Dotierung bereit. Aus dem so modellierten Kristall lässt sich über eine Methode unmittelbar ein Wellenfeld konstruieren. Es ist ferner möglich, den Kristall in standardisierte *xyz*-Dateien zwischenspeichern und daraus neu einzuladen.

B.5 Orientation

Die Aufbauten in Experimenten befinden sich nicht zwangsläufig in völliger Übereinstimmung parallel zu den kartesischen Achsen eines spezifischen Koordinatensystems. Daher ist die Implementierung einer Rotation nötig. Dafür existiert ein Orientierungs-Objekt, das wie in Abschnitt 5.4.2 beschrieben funktioniert und neben der Angabe einer Rotation auch einen Drehpunkt definiert, um den die Rotation stattfinden soll.

Dabei wird die Rotation über ein gleichnamiges Objekt der *scipy*-Bibliothek umgesetzt. Sie kann etwa durch Angabe dreier Euler'scher Winkel, einer Rotationsmatrix oder Hamilton'scher Quaternionen definiert werden.

Mithilfe von angebotenen Methoden ist es möglich, die Orientierung möglichst kompakt und in einem für Metadaten der Wellenfelder kompatiblen Datenformat zu serialisieren und sie daraus wieder als Objekt zu konstruieren.

Weitere Methoden führen auf Basis der gewünschten Rotation eine Transformation der Koordinaten eines Wellenfelds durch. Dabei können diese Ursprungskordinaten als Volumenobjekt oder als Liste einzelner Koordinaten eingegeben werden.

B.6 utils

Kleinere unterstützende Methoden und Konstanten werden an dieser Stelle zusammengefasst, damit sie in anderen Bereichen der Bibliothek genutzt werden können. Eine Datenbank stellt SI-Präfixe und ihre Faktoren gegenüber, um in einer anderen Methode die Interpretation einer ausgeschriebenen Zahl aus einer Zeichenkette in SI-Einheiten zuzulassen. Verwendet werden kann diese Methode dann beispielsweise, um Experimentalparameter enthaltende Dateinamen von Rekonstruktionsergebnissen oder über die Kommandozeile übergebene Argumente auszuwerten.

Eine weitere Methode berechnet den *R*-Faktor nach Ref. [232] zum Vergleich der Übereinstimmung von simulierten und gemessenen Intensitätsmustern. An dieser Stelle ist auch die eigentliche Methode zur Berechnung der IMFP-Berechnung in zweifacher Ausführung beheimatet: einmal in den Einheiten von Ref. [60], ein zweites Mal in SI-Einheiten.

Ferner erzeugt eine Methode eine inhomogene Verteilung von Werten in einem angegebenen Intervall, die sich um ausgezeichnete Sammelpunkte häuft. Diese Verteilung ist für die Erstellung der Koordinaten eines Rekonstruktionsexperiments hilfreich, wenn im Umkreis von erwarteten Atomen genauer gerechnet werden soll. Damit lässt sich für erste Approximationen gröber und damit schneller rechnen.

B.7 typing

Abschließend befinden sich in diesem Bestandteil Abkürzungen, die die Beschreibung von Datentypen vereinfachen. Zwar verfügt Python als Programmiersprache über keine strenge Typisierung, dennoch setzt die *Hololib* konsequent auf Spezifizierung der Datentypen. Das verhindert Datentypfehler ohne Einbußen in der Performance und ist einer vollständigen Dokumentation zuträglich.

C Klassifizierungsgrößen

Die hier zusammengetragenen Klassifizierungsgrößen und ihre äquivalenten Formulierungen finden sich in Abschnitt 6.3 hergeleitet, ferner entspringen sie Ref. [192, Tab. 2.3] und Ref. [211]. Der Critical Success Index (häufig auch Threat Score genannt) wird häufig im Zusammenhang mit Wetterprognosen verwendet [218]. Der Fowlkes-Mallows Index [216] wird in einer angepassten Darstellung nach Ref. [217] aus der Wahrheitsmatrix gewonnen. Der Matthews Correlation Coefficient unterscheidet sich von den bisherigen Indizes insbesondere dadurch, dass er Werte $FM \in [-1, 1]$ ermittelt [212]. Der für diese Arbeit besonders wichtige F_1 -Score stammt ursprünglich aus Ref. [27].

Name	Definition	Entsprechung	Wahrscheinlichkeit
Positive	$Pos = \sum_x I[l(x) = \oplus]$		
Negative	$Neg = \sum_x I[l(x) = \ominus]$	$ \mathcal{S} - Pos$	
Richtig Positive	$TP = \sum_x I[\hat{c}(x) = l(x) = \oplus]$		
Richtig Negative	$TN = \sum_x I[\hat{c}(x) = l(x) = \ominus]$		
Falsch Positive	$FP = \sum_x I[\hat{c}(x) = \oplus \neq l(x) = \ominus] = Neg - TN$		
Falsch Negative	$FN = \sum_x I[\hat{c}(x) = \ominus \neq l(x) = \oplus] = Pos - TP$		
Korrektklassifikationsrate	$acc = (TP + TN)/ \mathcal{S} $	$(TP + TN)/(Pos + Neg)$	$P(\hat{c}(x) = l(x))$
Falschklassifikationsrate	$err = (FP + FN)/ \mathcal{S} $	$1 - acc$	$P(\hat{c}(x) \neq l(x))$
Sensitivität	$tpr = TP/Pos$		$P(\hat{c}(x) = \oplus l(x) = \oplus)$
Spezifität	$tnr = TN/Neg$		$P(\hat{c}(x) = \ominus l(x) = \ominus)$
Ausfallrate	$fpr = FP/Neg$	$1 - tnr$	$P(\hat{c}(x) = \oplus l(x) = \ominus)$
Falsch-Negativ-Rate	$fnr = FN/Pos$	$1 - tpr$	$P(\hat{c}(x) = \ominus l(x) = \oplus)$
Präzision	$ppv = TP/(TP + FP)$		$P(l(x) = \oplus \hat{c}(x) = \oplus)$
Critical Success Index	$TS = TP/(TP + FN + FP)$		
Fowlkes-Mallows Index	$FM = \sqrt{\frac{TP}{TP + FP} \cdot \frac{TP}{TP + FN}} = \sqrt{ppv \cdot tpr}$		
Matthews Correlation Coefficient	$MCC = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TN + FN) \cdot Pos \cdot Neg}}$		
F_1 -Score	$F_1 = 2 \cdot (ppv \cdot tpr)/(ppv + tpr) = 2TP/(2TP + FP + FN)$		

Symbolverzeichnis

Naturkonstanten

a_W	Gitterkonstante von Wolfram
c	Lichtgeschwindigkeit im Medium
c_0	Vakuumlichtgeschwindigkeit
e	Elementarladung
h	Plancksches Wirkungsquantum
m_e	Elektronenmasse
N_A	Avogadrokonstante

Funktionen

$F(x, y)$	Fresnel-Zonenplatte
$\mathcal{F}\{s\}$	Fourier-Transformation
$\mathcal{F}^{-1}\{\tilde{s}\}$	Inverse Fourier-Transformation
$s \circ \tilde{}$	Symbolhafte Fourier-Transformation
$f \in \mathcal{O}(g)$	Landau-Symbol, f wächst nicht schneller als g
$I[\cdot]$	Indikatorfunktion
$\text{sgn}(x)$	Signums-Funktion
$\delta(x)$	Delta-Distribution
$\text{III}(x)$	Einheits-Deltakamm-Funktion
$\Theta_k(x)$	Heavisidefunktion mit Wert $\Theta_k(0) = k$
$\Pi(x)$	Rechteckfunktion
∇	Nabla-Operator
∇^2	Laplace-Operator
\square	D'Alembert-Operator

Experimentparameter

a	Gitterkonstante
D	Abstand zwischen Objekt und Schirm
d	Kantenlänge des holographischen Schirms
d_r	Kantenlänge des Rekonstruktionsvolumens
N	Anzahl der Pixel des holographischen Schirms in einer Dimension
N_r	Anzahl der Pixel des Rekonstruktionsvolumens in einer Dimension
\mathcal{N}'	Anzahl addierter Rekonstruktionsvolumen
ϑ	Winkel zwischen einfallendem Elektronenstrahl und Probe
φ	Verkipfung der Probe, entspricht $90^\circ - \vartheta$
Φ	Öffnungswinkel der als Kegel angenommenen Wellenfunktion

Wellenbeschreibung

A	Allgemeine Amplitude
I	Intensität

\mathbf{k}	Wellenvektor
\mathcal{P}	Propagationsoperator
t_s	Rekonstruktionsthreshold
Γ	Rekonstruierte Intensität
γ	Kohärenzgrad
λ	Wellenlänge
λ_B	de-Broglie-Wellenlänge
λ_{IMFP}	Inelastic Mean Free Path (IMFP)
φ	Allgemeine Phasenverschiebung
ψ	Allgemeine Wellenfunktion
ψ_o	Objektwelle
ψ_r	Referenzwelle
ψ_w	Wiedergabewelle

Signaltheorie

$s(t)$	Zeitabhängiges Signal
$s_a(t)$	Abgetastetes Signal
$s_r(t)$	Rekonstruiertes Signal
$s_q(t)$	Quantisiertes Signal
$\tilde{s}(\omega)$	Frequenzspektrum des Signals
Δt	Abtastabstand
$\Delta\omega$	Abtastfrequenz
σ	Oversampling ratio
σ'	Linearer oversampling ratio
ω	Frequenz
ω_{NY}	Nyquist-Frequenz

Klassifikation

\oplus	Klassifizierungslabel „positive“
\ominus	Klassifizierungslabel „negative“
\mathcal{A}_i	Menge der Pixel in einer Atomkugel
\mathcal{C}	Klassenmenge
\mathcal{L}	Labelmenge
\mathcal{S}	Menge aller Pixel auf einem Schirm als Teilmenge der Instanzmenge
\mathcal{X}	Instanzmenge
$\hat{c}(x)$	Klassifizierungsfunktion
$l(x)$	Labelfunktion
r_a	Atomradius in der Klassifizierung
F_1	F_1 -Score
\bar{F}_1	Kumulierter F_1 -Score für eine Energie
$\zeta_{\mathcal{N}'}$	Kumulierter F_1 -Score für \mathcal{N}' Energien
FN	Anzahl an False Negatives
FP	Anzahl an False Positives
Neg	Anzahl Pixel ohne Atome
Pos	Anzahl vorhandener Atome
TN	Anzahl an True Negatives

<i>TP</i>	Anzahl an True Positives
<i>acc</i>	Accuracy
<i>err</i>	Error Rate
<i>fnr</i>	False negative rate
<i>fpr</i>	False positive rate
<i>ppv</i>	Präzision
<i>tnr</i>	True negative rate
<i>tpr</i>	True positive rate

Weitere

<i>E</i>	Energie
<i>E_B</i>	Bindungsenergie
<i>E_g</i>	Bandlückenenergie
<i>E_{kin}</i>	Kinetische Energie
<i>E_p</i>	Plasmonenenergie
<i>R</i>	R-Faktor nach Pendry et al.
<i>m</i>	Atomare Masse
<i>ρ</i>	Massendichte

Akronymverzeichnis

bcc	body centered cubic	66
BKD	Backscatter Kikuchi Diffraction	22, <i>siehe</i> EBSD
CCD	Charge Coupled Device, Häufig verwendete Bauart für Fotosensoren	24, 31, 45, 50
CG	Computer Graphics, Dreidimensionale Objektdarstellung am Computer	42, 43, 49
CGH	Computer Generated Hologram	42, 49
CH	Computer Holography, Durchführung von Holographie und/oder Rekonstruktion am Computer	42, 43, 45, 49, 56
CPU	Central Processing Unit	57–59, 99, 107
CUDA	Compute Unified Device Architecture, Programmierschnittstelle des Herstellers Nvidia zur Berechnung von Programmbestandteilen auf der GPU	58, 107
DFT	Discrete Fourier Transformation	41
DH	Digital Holography, Digitalisierung eines optischen Hologramms	42
DPH	Differential Photoelectron Holography, Von Omori et al. vorgeschlagene Methode zur Verbesserung von Photoelektronenholographie	29, 97
DSP	Digital Signal Processor	59
EBSD	Electron Backscatter Diffraction, Elektronenrückstreubeugung	21–29, 31, 49, 65, 67, 69, 77, 95, 97, 98
fcc	face centered cubic	67
FFT	Fast Fourier Transformation	41, 51, 58
FPGA	Field Programmable Gate Array	58, 59
GPU	Graphics Processing Unit	51, 58, 99, 107
HPC	High Performance Computing	80
IMFP	Inelastic Mean Free Path, Mittlere freie Weglänge	19, 21, 25, 26, 68, 69, 77, 111, 112
LEED	Low-Energy Electron Diffraction	25, 97
LiDO3	Linux Cluster an der TU Dortmund	80, 87
LUT	Lookup Table	52–56, 59
MED	Mean Escape Depth, Maximale Tiefe der bei der Spektroskopie detektierten Elektronen	19, 20
N-LUT	Novel Lookup Table	53, 54
OpenCL	Open Computing Language	58
PFP	Principal Fringe Pattern, Bezeichnung für die Fresnelzonenplatte	50, 54, 55
RHEED	Reflection High-Energy Electron Diffraction	23, 24, 27, 99

ROCm	Radeon Open Computer Platform, Programmierschnittstellen für die Arbeit mit GPU des Herstellers AMD	58
TKD	Transmission Kikuchi Diffraction	25
WKS	Kotelnikov, Whittaker und Shannon, Namensgeber des WKS-Abtasttheorems	39
WRP	Wavefront Recording Plane, Virtuelle Ebene für Propagationsmethode	55, 56
XPD	X-ray photoelectron diffraction	50

Abbildungsverzeichnis

2.1	Mikroskop nach Abbe'scher Theorie	6
2.2	Aufnahme von Lichthogrammen	7
2.3	Phasenprojektion einer Kugelwelle	10
2.4	Konstruktion zum holographischen Prozess	11
2.5	Beispielhologramme	13
2.6	Schematischer Aufbau der Holographie und Rekonstruktion	15
3.1	Universelle Kurve	20
3.2	Isotrope und konische Wellenform	21
3.3	Geometrische Darstellung der Bragg-Bedingung	22
3.4	EBSD-Aufbaus	23
3.5	Referenz- und Streuwelle im Falle von EBSD	26
3.6	Experimentelles EBSD-Muster und Differenzbild	28
4.1	Rechteckfunktion und ihre Fourier-Entwicklung	32
4.2	Quantisierung und Abtastung	35
4.3	Visualisierung der Unterabtastung	36
4.4	Visualisierung der Unterabtastung mithilfe des Fourier-Spektrums	37
4.5	Funktionsapproximation durch Spaltfunktionen	38
4.6	Unterabtastung einer ebenen Welle	40
4.7	Unterabtastung eines Hologramms	41
4.8	Flussdiagramm zur (Computer-)Holographie	42
4.9	Projektion und Propagation	43
4.10	Beugung an einer Blende und Propagation	44
5.1	Punktbasierte Propagation	50
5.2	Faltungsmethode	51
5.3	Ebenenbasiertes Sampling	52
5.4	Punktbasierte Propagation im Detail	52
5.5	Lookup-Tabellen	53
5.6	Line-Scan-Methode	55
5.7	Wavefront Recording Plane	56
5.8	Polygonbasierte Propagation	56
5.9	Propagation mit künstlicher Intelligenz	57
5.10	Kugel- und kegelförmige Propagation	60
5.11	Intensitätsmuster nach Rotation	61
5.12	Flussdiagramm Rekonstruktionsprozess	63
6.1	Einheitszelle Wolfram	66
6.2	Simulierter EBSD-Aufbau	67
6.3	Parameter des holographischen Schirms	68
6.4	Lage des Rekonstruktionsvolumens	70

6.5	Wahrheitsmatrix	73
7.1	Beispielholographie und -rekonstruktion	78
7.2	Rekonstruktionsvolumen nach Threshold	81
7.3	Beispielhafte Rekonstruktion	83
7.4	Gegenüberstellung verschiedener Scores	85
7.5	$F_1(N, d)$ für $E_{\text{kin}} = 20 \text{ keV}$	86
7.6	Abschätzung der Mindestpixelgröße nach Abtasttheorem	88
7.7	F_1 -Score nach Pixelgröße	90
7.8	$F_1(N, d)$ für $\mathcal{N}' = 14$	92
7.9	Vergleich Pixelgröße für verschiedene \mathcal{N}'	93
7.10	Der gemittelte Score $\zeta_{\mathcal{N}'}$ gegen \mathcal{N}'	94

Literaturverzeichnis

- [1] E. Hecht. *Optik*.
Übers. von K. Lippert. 7. Aufl. Berlin, Boston: De Gruyter, 2018
(siehe S. 1, 7, 9).
- [2] E. Hoppe. *Geschichte der Optik*.
Unveränderter Neudruck der Ausgabe von 1926. Wiesbaden: Sändig, 1967
(siehe S. 1).
- [3] J. M. Eder. *History of Photography*.
Übers. von E. Epstean. New York: Columbia University Press, 1945 (siehe S. 1).
- [4] T. Wright. *Visual Impact. Culture and the Meaning of Images*.
Oxford, New York: Berg, 2008 (siehe S. 1).
- [5] S. R. Wilk. *Ibn al-Haytham: 1,000 Years after the Kitāb al-Manāẓir*.
In: *Optics and Photonics News* 26.10 (2015), S. 42–48 (siehe S. 1).
- [6] H. Gernsheim. *Geschichte der Photographie. Die ersten hundert Jahre*.
Propyläen Kunstgeschichte Sbd. 3.
Frankfurt am Main, Berlin, Wien: Propyläen Verlag, 1983 (siehe S. 1).
- [7] W. Püschel. *Die Farbphotographie*.
In: *Chemie in unserer Zeit* 4.1 (1970), S. 9–15 (siehe S. 1).
- [8] A. J. P. Theuwissen. *Solid-State Imaging with Charge-Coupled Devices*.
Neudruck mit Korrekturen. Solid-State Science and Technology Library Bd. 1.
Dordrecht: Kluwer Academic Publishers, 1996 (siehe S. 1, 50).
- [9] D. Brewster. *The Stereoscope. Its History, Theory, and Construction*.
London: John Murray, 1856 (siehe S. 2).
- [10] F. Devernay und P. Beardsley. *Stereoscopic Cinema*.
In: *Image and Geometry Processing for 3-D Cinematography*.
Hrsg. von R. Ronfard und G. Taubin. Geometry and Computing Bd. 5.
Berlin, Heidelberg: Springer, 2010, S. 11–51 (siehe S. 2).
- [11] K. Matsushima. *Introduction to Computer Holography. Creating
Computer-Generated Holograms as the Ultimate 3D Image*.
Series in Display Science and Technology. Cham: Springer, 2020
(siehe S. 2 f., 9, 14, 36 f., 40–44, 46, 52, 61–63, 79, 97).
- [12] F. Zernike. *Diffraction and Optical Image Formation*.
In: *Proceedings of the Physical Society* 61.2 (1948), S. 158–164 (siehe S. 2, 6, 12).
- [13] F. Zernike. *Das Phasenkontrastverfahren bei der mikroskopischen Beobachtung*.
In: *Zeitschrift für technische Physik* 11 (1935), S. 454–457 (siehe S. 2, 6).
- [14] D. Gabor. *A New Microscopic Principle*.
In: *Nature* 161.4098 (1948), S. 777–778 (siehe S. 2, 6, 11).
- [15] W. E. Kock. *Nobel Prize for Physics: Gabor and Holography*.
In: *Science* 174.4010 (1971), S. 674–675 (siehe S. 2).

- [16] W. Rankl und W. Effing.
Handbuch der Chipkarten. Aufbau – Funktionsweise – Einsatz von Smart Cards.
5. Aufl. München: Carl Hanser, 2008 (siehe S. 2).
- [17] E.-S. Kim. *Three-Dimensional Projection Display System.*
In: *Digital Holography and Three-Dimensional Display. Principles and Applications.* Hrsg. von T.-C. Poon. Boston: Springer, 2006. Kap. 11, S. 293–332
(siehe S. 2, 31, 97).
- [18] J. M. Cowley. *Twenty forms of electron holography.*
In: *Ultramicroscopy* 41.4 (1992), S. 335–348 (siehe S. 2, 25, 97).
- [19] A. J. Wilkinson und T. B. Britton.
Strains, planes, and EBSD in materials science.
In: *Materials Today* 15.9 (2012), S. 366–376 (siehe S. 2, 24, 97).
- [20] M. Gianfelice und C. Westphal.
Electron holography by planar electron backscattered diffraction patterns.
In: *Journal of the Optical Society of America A* 39.2 (2022), A1–A6
(siehe S. 2 f., 23, 26, 69, 83, 86, 92, 94, 98).
- [21] M. Gianfelice und C. Westphal.
Simulated Imaging of Atomic Surface-Structures via Holographic Approaches by Planar Electron Backscattered Diffraction Patterns.
In: *OSA Technical Digest* JTu5A.22 (2021) (siehe S. 2).
- [22] Y. Chen, F. J. García de Abajo, A. Chassé, R. X. Ynzunza, A. P. Kaduwela,
M. A. Van Hove und C. S. Fadley. *Convergence and reliability of the Rehr-Albers
formalism in multiple-scattering calculations of photoelectron diffraction.*
In: *Physical Review B* 58.19 (1998), S. 13121–13131 (siehe S. 2, 21, 97).
- [23] H. Nyquist. *Certain Topics in Telegraph Transmission Theory.*
In: *Transactions of the American Institute of Electrical Engineers* 47.2 (1928),
S. 617–644 (siehe S. 3, 39).
- [24] C. E. Shannon. *A Mathematical Theory of Communication.*
In: *The Bell System Technical Journal* 27.3 (1948), S. 379–423 (siehe S. 3, 39).
- [25] R. H. Landau, M. J. Páez und C. C. Bordeianu.
Computational Physics. Problem Solving with Computers.
2. Aufl. Weinheim: Wiley-VCH, 2007 (siehe S. 3).
- [26] J.-H. Park.
Recent progress in computer-generated holography for three-dimensional scenes.
In: *Journal of Information Display* 18.1 (2017), S. 1–12 (siehe S. 3, 50, 52).
- [27] N. Chinchor. *MUC-4 evaluation metrics.*
In: *MUC4 '92: Proceedings of the 4th conference on Message understanding*
(1992), S. 22–29 (siehe S. 3, 113).
- [28] E. Abbe.
Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung.
In: *Archiv für Mikroskopische Anatomie* 9 (1873), S. 413–468 (siehe S. 5).

-
- [29] H. Boersch. *Über das primäre und sekundäre Bild im Elektronenmikroskop. I. Eingriffe in das Beugungsbild und ihr Einfluß auf die Abbildung.*
In: *Annalen der Physik* 418.7 (1936), S. 631–644 (siehe S. 6).
- [30] E. Menzel, W. Mirandé und I. Weingärtner. *Fourier-Optik und Holographie.*
Wien, New York: Springer, 1973 (siehe S. 5, 31).
- [31] M. Wolfke. *Über die Möglichkeit der optischen Abbildung von Molekulargittern.*
In: *Physikalische Zeitschrift* 21.18 (1920), S. 495–497 (siehe S. 5).
- [32] H. Schober. *Die Direktwahrnehmung von Röntgenstrahlen durch den menschlichen Gesichtssinn.*
In: *Vision Research* 4.3–4 (1964), S. 251–269 (siehe S. 5).
- [33] H. Kiemle und D. Röss. *Einführung in die Technik der Holographie.*
Technisch-Physikalische Sammlung Bd. 8.
Frankfurt am Main: Akademische Verlagsgesellschaft, 1969 (siehe S. 6 f., 17).
- [34] H. Boersch. *Zur Bilderzeugung im Mikroskop.*
In: *Zeitschrift für technische Physik* 11 (1938), S. 337–338 (siehe S. 6).
- [35] H. Boersch. *Holographie und Elektronenoptik.*
In: *Physikalische Blätter* 23.9 (1967), S. 393–404 (siehe S. 6).
- [36] W. L. Bragg. *A New Type of 'X-Ray Microscope'.*
In: *Nature* 143.3625 (1939), S. 678 (siehe S. 6).
- [37] W. L. Bragg. *The X-Ray Microscope.*
In: *Nature* 149.3782 (1942), S. 470–471 (siehe S. 6).
- [38] M. J. Buerger. *The Photography of Atoms in Crystals.*
In: *Proceedings of the National Academy of Sciences* 36.5 (1950), S. 330–335 (siehe S. 6).
- [39] M. J. Buerger. *Generalized Microscopy and the Two-Wave-Length Microscope.*
In: *Journal of Applied Physics* 21.9 (1950), S. 909–917 (siehe S. 6).
- [40] D. Gabor. *Wellenfront-Rekonstruktion oder „Holographie“.*
In: *Physikalische Blätter* 22.6 (1966), S. 256–265 (siehe S. 6).
- [41] E. Verdet. *Étude sur la constitution de la lumière non polarisée et de la lumière partiellement polarisée.*
In: *Annales scientifiques de l'École normale supérieure* 1.2 (1865), S. 291–316 (siehe S. 7).
- [42] E. N. Leith und J. Upatnieks. *Wavefront Reconstruction with Continuous-Tone Objects.*
In: *Journal of the Optical Society of America* 53.12 (1963), S. 1377–1381 (siehe S. 7).
- [43] E. N. Leith und J. Upatnieks. *Wavefront Reconstruction with Diffused Illumination and Three-Dimensional Objects.*
In: *Journal of the Optical Society of America* 54.11 (1964), S. 1295–1301 (siehe S. 7, 17).

- [44] E. N. Leith und J. Upatnieks.
Reconstructed Wavefronts and Communication Theory.
In: *Journal of the Optical Society of America* 52.10 (1962), S. 1123–1130
(siehe S. 7).
- [45] T. H. Maiman. *Stimulated Optical Radiation in Ruby*.
In: *Nature* 187.4736 (1960), S. 493–494 (siehe S. 7).
- [46] W. Nolting. *Grundkurs Theoretische Physik 3. Elektrodynamik*.
8. Aufl. Berlin, Heidelberg: Springer, 2007 (siehe S. 7, 9, 31 f., 41).
- [47] J. Eichler und H. J. Eichler. *Laser. Bauformen, Strahlführung, Anwendungen*.
5. Aufl. Berlin, Heidelberg: Springer, 2003 (siehe S. 9).
- [48] W. Martienssen. *Grundlagen und Anwendungen der Holographie*.
In: *VDI-Nachrichten* 38 (1967), S. 4–5 (siehe S. 10 f.).
- [49] D. Halliday, R. Resnick und J. Walker. *Physik*.
Übers. von S. W. Koch et al. Weinheim: Wiley-VCH, 2005 (siehe S. 10).
- [50] M. Born und E. Wolf. *Principles of Optics*.
7. Aufl. Cambridge: Cambridge University Press, 2019 (siehe S. 11, 16, 103).
- [51] T. Latychevskaia und H.-W. Fink.
Solution to the Twin Image Problem in Holography.
In: *Physical Review Letters* 98.23 (2007), S. 233901 (siehe S. 14, 27).
- [52] T. Lühr.
Elektronenholographie zur dreidimensionalen Abbildung von Atomstrukturen.
Dissertation. Technische Universität Dortmund, 2015
(siehe S. 15, 26–28, 50, 91, 97).
- [53] J. B. J. Fourier. *Théorie analytique de la chaleur*.
Paris: Firmin Didot, père et fils, 1822 (siehe S. 16, 32).
- [54] Y. N. Denisjuk und D. I. Stasel'ko.
Possibility of Obtaining Holograms Using a Reference Beam Whose Wavelength Differs from the Wavelength of the Light Scattered by an Object.
In: *Soviet Physics Doklady* 12.10 (1968), S. 954–955 (siehe S. 17).
- [55] L. de Broglie. *Recherches sur la théorie des Quanta*.
In: *Annales de Physique* 10.3 (1925), S. 22–128 (siehe S. 19).
- [56] R. Gross und A. Marx. *Festkörperphysik*.
2. Aufl. Berlin, Boston: De Gruyter, 2014 (siehe S. 19, 22).
- [57] M. P. Seah und W. A. Dench. *Quantitative Electron Spectroscopy of Surfaces: A Standard Data Base for Electron Inelastic Mean Free Paths in Solids*.
In: *Surface and Interface Analysis* 1.1 (1979), S. 2–11 (siehe S. 19).
- [58] C. J. Powell.
Practical guide for inelastic mean free paths, effective attenuation lengths, mean escape depths, and information depths in x-ray photoelectron spectroscopy.
In: *Journal of Vacuum Science & Technology A* 38.2 (2020), S. 023209
(siehe S. 19).

-
- [59] S. Hüfner. *Photoelectron Spectroscopy. Principles and Applications*. Springer Series in Solid-State Sciences Bd. 82. Berlin, Heidelberg: Springer, 1995 (siehe S. 20).
- [60] H. Shinotsuka, S. Tanuma, C. J. Powell und D. R. Penn. *Calculations of electron inelastic mean free paths. XII. Data for 42 inorganic compounds over the 50 eV to 200 keV range with the full Penn algorithm*. In: *Surface and Interface Analysis* 51.4 (2019), S. 427–457 (siehe S. 20 f., 112).
- [61] S. Tanuma, C. J. Powell und D. R. Penn. *Calculations of electron inelastic mean free paths. v. Data for 14 organic compounds over the 50–2000 eV range range*. In: *Surface and Interface Analysis* 21.3 (1994), S. 165–176 (siehe S. 20, 111).
- [62] T. Fauster, L. Hammer, K. Heinz und M. A. Schneider. *Oberflächenphysik. Grundlagen und Methoden*. 2. Aufl. Berlin, Boston: De Gruyter, 2019 (siehe S. 21, 26).
- [63] W. S. M. Werner. *Electron transport in solids for quantitative surface analysis*. In: *Surface and Interface Analysis* 31.3 (2001), S. 141–176 (siehe S. 21).
- [64] S. A. Chambers. *The Relative Importance of Elastic and Inelastic Scattering in Angle-Resolved Auger Electron Spectroscopy—A Response to “ADAM”*. In: *Langmuir* 6.8 (1990), S. 1427–1430 (siehe S. 21, 97).
- [65] C. Westphal, S. Dreiner, M. Schürmann und H. Zacharias. *Holographic reconstruction of Si(111) atom positions from energy- and angle-resolved photoelectron diffraction*. In: *Surface Science* 462.1-3 (2000), S. 103–122 (siehe S. 21, 97).
- [66] L.-M. Peng. *Electron atomic scattering factors and scattering potentials of crystals*. In: *Micron* 30.6 (1999), S. 625–648 (siehe S. 21).
- [67] R. A. Schwarzer. *Automated Crystal Lattice Orientation Mapping Using a Computer-controlled SEM*. In: *Micron* 28.3 (1997), S. 249–265 (siehe S. 22–24, 27, 97).
- [68] S. Nishikawa und S. Kikuchi. *Diffraction of Cathode Rays by Calcite*. In: *Nature* 122.3080 (1928), S. 726 (siehe S. 22).
- [69] W. Kossel. *Bemerkung zur scheinbaren selektiven Reflexion von Röntgenstrahlen an Kristallen*. In: *Zeitschrift für Physik* 23 (1924), S. 278–285 (siehe S. 22).
- [70] W. Kossel. *Zur Systematik der Röntgenreflexe eines Raumgitters*. In: *Annalen der Physik* 417.6 (1936), S. 512–526 (siehe S. 22).
- [71] M. von Laue. *Materiewellen und ihre Interferenzen*. 2. Aufl. Physik und Chemie und ihre Anwendungen in Einzeldarstellungen Bd. 7. Leipzig: Akademische Verlagsgesellschaft Geest & Portig, 1948 (siehe S. 22 f.).
- [72] M. Gajdardziska-Josifovska und J. M. Cowley. *Brillouin Zones and Kikuchi Lines for Crystals Under Electron Channeling Conditions*. In: *Acta Crystallographica Section A Foundations of Crystallography* 47.2 (1991), S. 74–82 (siehe S. 22).

- [73] G. I. Watson. *Neutron Compton scattering*.
In: *Journal of Physics: Condensed Matter* 8.33 (1996), S. 5955–5975
(siehe S. 23, 25).
- [74] H. Hashimoto, A. Howie und M. J. Whelan. *Anomalous electron absorption effects in metal foils: theory and comparison with experiment*.
In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 269.1336 (1962), S. 80–103 (siehe S. 23).
- [75] L. Reimer, U. Heilers und G. Saliger. *Kikuchi Band Contrast in Diffraction Patterns Recorded by Transmitted and Backscattered Electrons*.
In: *Scanning* 8.3 (1986), S. 101–118 (siehe S. 23).
- [76] A. Winkelmann, G. Nolze, M. Vos, F. Salvat-Pujol und W. S. M. Werner. *Physics-based simulation models for EBSD: advances and challenges*.
In: *IOP Conference Series: Materials Science and Engineering* 109 (2016), S. 012018 (siehe S. 23 f., 97).
- [77] M. N. Alam, M. Blackman und D. W. Pashley. *High-angle Kikuchi patterns*.
In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 221.1145 (1954), S. 224–242 (siehe S. 23).
- [78] W. Nolting.
Grundkurs Theoretische Physik 5/1. Quantenmechanik – Grundlagen.
8. Aufl. Berlin, Heidelberg: Springer Spektrum, 2013 (siehe S. 23).
- [79] J. C. H. Spence, X. Zhang, U. Weierstall, J. M. Zuo, E. Munro und J. Rouse. *LOW ENERGY POINT REFLECTION ELECTRON MICROSCOPY*.
In: *Surface Review and Letters* 4.3 (1997), S. 577–587 (siehe S. 23, 27).
- [80] B. L. Adams, S. I. Wright und K. Kunze.
Orientation Imaging: The Emergence of a New Microscopy.
In: *Metallurgical Transactions A* 24 (1993), S. 819–831 (siehe S. 24, 97).
- [81] D. Juul Jensen und N. H. Schmidt. *An automatic on-line technique for determination of crystallographic orientations by EBSP*.
In: *Recrystallization '90. Conference Proceedings*. Hrsg. von T. Chandra.
Warrendale: The Minerals, Metals and Materials Society, 1990, S. 219–224
(siehe S. 24).
- [82] J. B. Burns, A. R. Hanson und E. M. Riseman. *Extracting Straight Lines*.
In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-8.4 (1986), S. 425–455 (siehe S. 24).
- [83] S. I. Wright und B. L. Adams.
Automatic Analysis of Electron Backscatter Diffraction Patterns.
In: *Metallurgical Transactions A* 23 (1992), S. 759–767 (siehe S. 24).
- [84] N. C. Krieger Lassen, D. Juul Jensen und K. Conradsen.
Image Processing Procedures for Analysis of Electron Back Scattering Patterns.
In: *Scanning Microscopy* 6.1 (1992), S. 115–121 (siehe S. 24).
- [85] N. C. Krieger Lassen. *Automated Determination of Crystal Orientations from Electron Backscattering Patterns*.
Dissertation. Technical University of Denmark, 1994 (siehe S. 24).

-
- [86] N. C. Krieger Lassen, D. Juul Jensen und K. Conradsen. *Automatic Recognition of Deformed and Recrystallized Regions in Partly Recrystallized Samples Using Electron Back Scattering Patterns*. In: *Materials Science Forum* 157-162 (1994), S. 149–158 (siehe S. 24).
- [87] R. A. Schwarzer und J. Sukkau. *Automated Evaluation of Kikuchi Patterns by Means of Radon and Fast Fourier Transformation, and Verification by an Artificial Neural Network*. In: *Advanced Engineering Materials* 5.8 (2003), S. 601–606 (siehe S. 24).
- [88] P. V. C. Hough. *Method and means for recognizing complex patterns*. US-Patent 3069654. 1962 (siehe S. 24).
- [89] J. Illingworth und J. Kittler. *A Survey of the Hough Transform*. In: *Computer Vision, Graphics, and Image Processing* 44.1 (1988), S. 87–116 (siehe S. 24).
- [90] G. C. Stockman und A. K. Agrawala. *Equivalence of Hough Curve Detection to Template Matching*. In: *Communications of the ACM* 20.11 (1977), S. 820–822 (siehe S. 24).
- [91] J. Sklansky. *On the Hough Technique for Curve Detection*. In: *IEEE Transactions on Computers* C-27.10 (1978), S. 923–926 (siehe S. 24).
- [92] J. Radon. *Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten*. In: *Berichte über die Verhandlungen der Königlich-Sächsischen Gesellschaft der Wissenschaften zu Leipzig, Mathematisch-Physische Klasse* 69 (1917), S. 262–277 (siehe S. 24).
- [93] S. R. Deans. *Hough Transform from the Radon Transform*. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-3.2 (1981), S. 185–188 (siehe S. 24).
- [94] S. F. Bordín, S. Limandri, J. Ranalli und G. Castellano. *EBSA spatial resolution for detecting sigma phase in steels*. In: *Ultramicroscopy* 171 (2016), S. 177–185 (siehe S. 25).
- [95] S. Ren, E. Kenik, K. Alexander und A. Goyal. *Exploring Spatial Resolution in Electron Back-Scattered Diffraction Experiments via Monte Carlo Simulation*. In: *Microscopy and Microanalysis* 4.1 (1998), S. 15–22 (siehe S. 25).
- [96] D. P. Field. *Improving the Spatial Resolution of EBSD*. In: *Microscopy and Microanalysis* 11.S02 (2005), S. 52–53 (siehe S. 25).
- [97] A. Deal, T. Hooghan und A. Eades. *Energy-filtered electron backscatter diffraction*. In: *Ultramicroscopy* 108.2 (2008), S. 116–125 (siehe S. 25, 27, 67).
- [98] N. Brodusch, H. Demers und R. Gauvin. *Nanometres-resolution Kikuchi patterns from materials science specimens with transmission electron forward scatter diffraction in the scanning electron microscope*. In: *Journal of Microscopy* 250.1 (2013), S. 1–14 (siehe S. 25, 27).

- [99] R. Keller und R. Geiss.
Transmission EBSD from 10 nm domains in a scanning electron microscope.
In: *Journal of Microscopy* 245.3 (2012), S. 245–251 (siehe S. 25).
- [100] P. W. Trimby. *Orientation mapping of nanostructured materials using transmission Kikuchi diffraction in the scanning electron microscope.*
In: *Ultramicroscopy* 120 (2012), S. 16–24 (siehe S. 25).
- [101] J. A. Small, J. R. Michael und D. S. Bright. *Improving the quality of electron backscatter diffraction (EBSD) patterns from nanoparticles.*
In: *Journal of Microscopy* 206.2 (2002), S. 170–178 (siehe S. 25).
- [102] H.-W. Fink, H. Schmid, H. J. Kreuzer und A. Wierzbicki.
Atomic resolution in lensless low-energy electron holography.
In: *Physical Review Letters* 67.12 (1991), S. 1543–1546 (siehe S. 25).
- [103] H. Lichte. *Electron holography approaching atomic resolution.*
In: *Ultramicroscopy* 20.3 (1986), S. 293–304 (siehe S. 25).
- [104] K. Heinz, U. Starke und J. Bernhardt. *Surface holography with LEED electrons.*
In: *Progress in Surface Science* 64.3-8 (2000), S. 163–178 (siehe S. 25, 91, 97).
- [105] E. Rudberg. *Inelastic Scattering of Electrons from Solids.*
In: *Physical Review* 50.2 (1936), S. 138–150 (siehe S. 25).
- [106] G. Ruthemann. *Diskrete Energieverluste schneller Elektronen in Festkörpern.*
In: *Naturwissenschaften* 29.42 (1941), S. 648 (siehe S. 25).
- [107] A. Winkelmann und M. Vos.
Site-Specific Recoil Diffraction of Backscattered Electrons in Crystals.
In: *Physical Review Letters* 106.8 (2011), S. 085503 (siehe S. 25).
- [108] T. Lühr, A. Winkelmann, G. Nolze, D. Krull und C. Westphal.
Direct Atom Imaging by Chemical-Sensitive Holography.
In: *Nano Letters* 16.5 (2016), S. 3195–3201 (siehe S. 26 f., 97, 99).
- [109] A. Szöke. *Theory of multiphoton ionization of atoms.*
In: *AIP Conference Proceedings* 147.1 (1986), S. 169–173 (siehe S. 26).
- [110] B. P. Tonner, Z.-L. Han, G. R. Harp und D. K. Saldin.
Scattered-wave integral-transform method of holographic-image reconstruction from forward-scattering diffraction patterns.
In: *Physical Review B* 43.18 (1991), S. 14423–14433 (siehe S. 27).
- [111] J. Osterwalder, R. Fasel, A. Stuck, P. Aebi und L. Schlapbach.
Holographic interpretation of photoelectron diffraction.
In: *Journal of Electron Spectroscopy and Related Phenomena* 68 (1994), S. 1–18 (siehe S. 27).
- [112] J. Wider, F. Baumberger, M. Sami, R. Gotter, A. Verdini, F. Bruno, D. Cvetko, A. Morgante, T. Greber und J. Osterwalder. *Atomically Resolved Images from Near Node Photoelectron Holography Experiments on Al(111).*
In: *Physical Review Letters* 86.11 (2001), S. 2337–2340 (siehe S. 27).
- [113] J. J. Barton.
Removing Multiple Scattering and Twin Images from Holographic Images.
In: *Physical Review Letters* 67.22 (1991), S. 3106–3109 (siehe S. 27).

-
- [114] Y. Nishino, T. Ishikawa, K. Hayashi, Y. Takahashi und E. Matsubara. *Two-energy twin image removal in atomic-resolution x-ray holography*. In: *Physical Review B* 66.9 (2002), S. 092105 (siehe S. 27).
- [115] E. Hering, R. Martin und M. Stohrer. *Physik für Ingenieure*. 13. Aufl. Berlin, Heidelberg: Springer Vieweg, 2021 (siehe S. 27).
- [116] S. Omori, Y. Nihei, E. Rotenberg, J. D. Denlinger, S. Marchesini, S. D. Kevan, B. P. Tonner, M. A. Van Hove und C. S. Fadley. *Differential Photoelectron Holography: A New Approach for Three-Dimensional Atomic Imaging*. In: *Physical Review Letters* 88.5 (2002), S. 055504 (siehe S. 29, 97).
- [117] A. Maimone, A. Georgiou und J. S. Kollin. *Holographic Near-Eye Displays for Virtual and Augmented Reality*. In: *ACM Transactions on Graphics* 36.4 (2017), S. 85 (siehe S. 31).
- [118] J.-R. Ohm und H. D. Lüke. *Signalübertragung. Grundlagen der digitalen und analogen Nachrichtenübertragungssysteme*. 12. Aufl. Berlin, Heidelberg: Springer Vieweg, 2014 (siehe S. 31 f., 34 f., 38, 41).
- [119] L. Carleson. *On convergence and growth of partial sums of Fourier series*. In: *Acta Mathematica* 116.1 (1966), S. 135–157 (siehe S. 32).
- [120] R. Remmert. *Funktionentheorie*. Berlin, Heidelberg: Springer, 1991 (siehe S. 32).
- [121] E. Hewitt und R. E. Hewitt. *The Gibbs-Wilbraham phenomenon: An episode in fourier analysis*. In: *Archive for History of Exact Sciences* 21.2 (1979), S. 129–160 (siehe S. 32).
- [122] P. Bleicker, J. Stolze und G. S. Uhrig. *Probing thermalization in quenched integrable and nonintegrable Fermi-Hubbard models*. In: *Physical Review A* 102.1 (2020), S. 013321 (siehe S. 33).
- [123] R. Unbehauen. *Systemtheorie 1. Allgemeine Grundlagen, Signale und lineare Systeme im Zeit- und Frequenzbereich*. 8. Aufl. München: Oldenbourg Wissenschaftsverlag, 2002 (siehe S. 33 f.).
- [124] A. Weiße, G. Wellein, A. Alvermann und H. Fehske. *The kernel polynomial method*. In: *Reviews of Modern Physics* 78.1 (2006), S. 275–306 (siehe S. 33).
- [125] I. N. Bronstein, K. A. Semendjajew, G. Musiol und H. Mühlig. *Taschenbuch der Mathematik*. 7. Aufl. Frankfurt am Main: Verlag Harri Deutsch, 2008 (siehe S. 34, 38, 62, 90).
- [126] T. Butz. *Fouriertransformation für Fußgänger*. 7. Aufl. Wiesbaden: Vieweg+Teubner, 2011 (siehe S. 34).
- [127] J. W. Goodman. *Introduction to Fourier Optics*. 2. Aufl. McGraw-Hill Series in Electrical and Computer Engineering. Electromagnetics. New York [u. A.]: McGraw-Hill, 1996 (siehe S. 34).
- [128] R. N. Bracewell. *The Fourier Transform And Its Applications*. 3. Aufl. Singapur: McGraw-Hill, 2000 (siehe S. 35–38).

- [129] D. E. Knuth. *The Art of Computer Programming. Volume 2: Seminumerical Algorithms*. 3. Aufl. Reading, Massachusetts [u. A.]: Addison-Wesley, 1998 (siehe S. 35 f.).
- [130] S. W. Smith. *The Scientist and Engineer's Guide to Digital Signal Processing*. San Diego: California Technical Publishing, 1997 (siehe S. 36).
- [131] T. Görne. *Tontechnik*. 4. Aufl. München: Carl Hanser, 2015 (siehe S. 39).
- [132] C. E. Shannon. *Communication in the Presence of Noise*. In: *Proceedings of the IRE* 37.1 (1949), S. 10–21 (siehe S. 39).
- [133] J. M. Whittaker. *The "Fourier" Theory of the Cardinal Function*. In: *Proceedings of the Edinburgh Mathematical Society* 1.3 (1928), S. 169–176 (siehe S. 39).
- [134] V. A. Kotel'nikov. *On the transmission capacity of 'ether' and wire in electric communications*. In: *Physics-USpekhi* 49.7 (2006), S. 736–744 (siehe S. 39).
- [135] H. D. Lüke. *The Origins of the Sampling Theorem*. In: *IEEE Communications Magazine* 37.4 (1999), S. 106–108 (siehe S. 39).
- [136] M. Unser. *Sampling—50 Years After Shannon*. In: *Proceedings of the IEEE* 88.4 (2000), S. 569–587 (siehe S. 39).
- [137] P. Bachmann. *Zahlentheorie. Zweiter Teil: Die analytische Zahlentheorie*. Leipzig: B. G. Teubner, 1894 (siehe S. 41).
- [138] E. Landau. *Handbuch der Lehre von der Verteilung der Primzahlen*. 2. Aufl. New York: Chelsea Publishing Company, 1953 (siehe S. 41).
- [139] D. E. Knuth. *Big Omicron and big Omega and big Theta*. In: *ACM SIGACT News* 8.2 (1976), S. 18–24 (siehe S. 41).
- [140] E. O. Brigham. *The fast Fourier transform*. Englewood Cliffs: Prentice Hall, 1974 (siehe S. 41).
- [141] B. Fröhlich. *Ray Tracing mit Strahlenbündeln*. Dissertation. Technische Universität Carolo-Wilhelmina zu Braunschweig, 1993 (siehe S. 43).
- [142] A. S. Glassner. *An Overview of Ray Tracing*. In: *An Introduction to Ray Tracing*. Hrsg. von A. S. Glassner. Neudruck. London: Academic Press, 1991. Kap. 1, S. 1–31 (siehe S. 43).
- [143] P. Shirley und S. Marschner. *Fundamentals of Computer Graphics*. 3. Aufl. Natick: A K Peters, 2009 (siehe S. 43).
- [144] J. Miao, D. Sayre und H. N. Chapman. *Phase retrieval from the magnitude of the Fourier transforms of nonperiodic objects*. In: *Journal of the Optical Society of America A* 15.6 (1998), S. 1662–1669 (siehe S. 46, 87, 91).
- [145] T. Lатычевская. *Iterative phase retrieval in coherent diffractive imaging: practical issues*. In: *Applied Optics* 57.25 (2018), S. 7187–7197 (siehe S. 46, 87).

-
- [146] T. Latychevskaia. *Reconstruction of missing information in diffraction patterns and holograms by iterative phase retrieval*.
In: *Optics Communications* 452 (2019), S. 56–67 (siehe S. 47).
- [147] J. P. Waters.
HOLOGRAPHIC IMAGE SYNTHESIS UTILIZING THEORETICAL METHODS.
In: *Applied Physics Letters* 9.11 (1966), S. 405–407 (siehe S. 49).
- [148] B. R. Brown und A. W. Lohmann. *Complex Spatial Filtering with Binary Masks*.
In: *Applied Optics* 5.6 (1966), S. 967–969 (siehe S. 49).
- [149] A. W. Lohmann und D. P. Paris.
Binary Fraunhofer Holograms, Generated by Computer.
In: *Applied Optics* 6.10 (1967), S. 1739–1748 (siehe S. 49).
- [150] B. R. Brown und A. W. Lohmann. *Computer-generated Binary Holograms*.
In: *IBM Journal of Research and Development* 13.2 (1969), S. 160–168 (siehe S. 49).
- [151] J. J. Burch.
A Computer Algorithm for the Synthesis of Spatial Frequency Filters.
In: *Proceedings of the IEEE* 55.4 (1967), S. 599–601 (siehe S. 49).
- [152] P. W. M. Tsang, T.-C. Poon und Y. M. Wu.
Review of fast methods for point-based computer-generated holography [Invited].
In: *Photonics Research* 6.9 (2018), S. 837–846 (siehe S. 49, 53, 56).
- [153] M. Bayraktar und M. Özcan. *Method to calculate the far field of three-dimensional objects for computer-generated holography*.
In: *Applied Optics* 49.24 (2010), S. 4647–4654 (siehe S. 51).
- [154] T. Sonar. *3000 Years of Analysis. Mathematics in History and Culture*.
Übers. von M. Patricia und K. W. Morton. Cham: Birkhäuser, 2021 (siehe S. 52).
- [155] M. Lucente. *Interactive computation of holograms using a look-up table*.
In: *Journal of Electronic Imaging* 2.1 (1993), S. 28–34 (siehe S. 52).
- [156] S.-C. Kim und E.-S. Kim. *Effective generation of digital holograms of three-dimensional objects using a novel look-up table method*.
In: *Applied Optics* 47.19 (2008), S. D55–D62 (siehe S. 54).
- [157] S.-C. Kim, J.-H. Kim und E.-S. Kim. *Effective reduction of the novel look-up table memory size based on a relationship between the pixel pitch and reconstruction distance of a computer-generated hologram*.
In: *Applied Optics* 50.19 (2011), S. 3375–3382 (siehe S. 54).
- [158] S.-C. Kim und E.-S. Kim. *Fast computation of hologram patterns of a 3D object using run-length encoding and novel look-up table methods*.
In: *Applied Optics* 48.6 (2009), S. 1030–1041 (siehe S. 54).
- [159] Z. Yang, Q. Fan, Y. Zhang, J. Liu und J. Zhou.
A new method for producing computer generated holograms.
In: *Journal of Optics* 14.9 (2012), S. 095702 (siehe S. 54).

- [160] J. Bresenham.
A Linear Algorithm for Incremental Digital Display of Circular Arcs.
In: *Communications of the ACM* 20.2 (1977), S. 100–106 (siehe S. 55).
- [161] T. Nishitsuji, T. Shimobaba, T. Kakue, N. Masuda und T. Ito. *Fast calculation of computer-generated hologram using the circular symmetry of zone plates.*
In: *Optics Express* 20.25 (2012), S. 27496–27502 (siehe S. 55).
- [162] S. Jiao, Z. Zhuang und W. Zou. *Fast computer generated hologram calculation with a mini look-up table incorporated with radial symmetric interpolation.*
In: *Optics Express* 25.1 (2017), S. 112–123 (siehe S. 55).
- [163] Y. Pan, X. Xu, S. Solanki, X. Liang, R. B. A. Tanjung, C. Tan und T.-C. Chong. *Fast CGH computation using S-LUT on GPU.*
In: *Optics Express* 17.21 (2009), S. 18543–18555 (siehe S. 55).
- [164] H. Q. Lin und J. E. Gubernatis.
Exact Diagonalization Methods for Quantum Systems.
In: *Computers in Physics* 7.4 (1993), S. 400–407 (siehe S. 55).
- [165] J. Jia, Y. Wang, J. Liu, X. Li, Y. Pan, Z. Sun, B. Zhang, Q. Zhao und W. Jiang. *Reducing the memory usage for effective computer-generated hologram calculation using compressed look-up table in full-color holographic display.*
In: *Applied Optics* 52.7 (2013), S. 1404–1412 (siehe S. 55).
- [166] T. Shimobaba, N. Masuda und T. Ito. *Simple and fast calculation algorithm for computer-generated hologram with wavefront recording plane.*
In: *Optics Letters* 34.20 (2009). Hrsg. von S. A. Benton, S. 3133–3135 (siehe S. 55).
- [167] A.-H. Phan, M.-l. Piao, S.-K. Gil und N. Kim.
Generation speed and reconstructed image quality enhancement of a long-depth object using double wavefront recording planes and a GPU.
In: *Applied Optics* 53.22 (2014), S. 4817–4824 (siehe S. 56).
- [168] A.-H. Phan, M. A. Alam, S.-H. Jeon, J.-H. Lee und N. Kim. *Fast hologram generation of long depth object using multiple wavefront recording planes.*
In: *Proceedings of SPIE* 9006 (2014), S. 900612 (siehe S. 56).
- [169] P. Tsang, W.-K. Cheung, T.-C. Poon und C. Zhou.
Holographic video at 40 frames per second for 4-million object points.
In: *Optics Express* 19.16 (2011), S. 15205–15211 (siehe S. 56).
- [170] Y. Pan, Y. Wang, J. Liu, X. Li und J. Jia. *Fast polygon-based method for calculating computer-generated holograms in three-dimensional display.*
In: *Applied Optics* 52.1 (2013), A290–A299 (siehe S. 56).
- [171] T. Shimobaba und T. Ito.
Fast generation of computer-generated holograms using wavelet shrinkage.
In: *Optics Express* 25.1 (2017), S. 77–86 (siehe S. 57).
- [172] M. Yamaguchi, H. Hoshino, T. Honda und N. Ohyama. *Phase added stereogram: calculation of hologram using computer graphics technique.*
In: *Proceedings of SPIE* 1914 (1993), S. 25–31 (siehe S. 57).

-
- [173] H. Kang, T. Yamaguchi und H. Yoshikawa.
Accurate phase-added stereogram to improve the coherent stereogram.
In: *Applied Optics* 47.19 (2008), S. D44–D54 (siehe S. 57).
- [174] D. Blinder und P. Schelkens.
Phase added sub-stereograms for accelerating computer generated holography.
In: *Optics Express* 28.11 (2020), S. 16924–16934 (siehe S. 57).
- [175] A. Krizhevsky, I. Sutskever und G. E. Hinton.
ImageNet classification with deep convolutional neural networks.
In: *Communications of the ACM* 60.6 (2017), S. 84–90 (siehe S. 57).
- [176] O. Ronneberger, P. Fischer und T. Brox.
U-Net: Convolutional Networks for Biomedical Image Segmentation.
In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015.* Hrsg. von N. Navab, J. Hornegger, W. Wells und A. Frangi.
Lecture Notes in Computer Science Bd. 9351. Cham: Springer, 2015, S. 234–241 (siehe S. 57).
- [177] K. He, X. Zhang, S. Ren und J. Sun.
Deep Residual Learning for Image Recognition.
In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (2016),* S. 770–778 (siehe S. 57).
- [178] R. Horisaki, R. Takagi und J. Tanida. *Deep-learning-generated holography.*
In: *Applied Optics* 57.14 (2018), S. 3859–3863 (siehe S. 57).
- [179] J. Wu, K. Liu, X. Sui und L. Cao. *High-speed computer-generated holography using an autoencoder-based deep neural network.*
In: *Optics Letters* 46.12 (2021), S. 2908–2911 (siehe S. 57).
- [180] L. Shi, B. Li, C. Kim, P. Kellnhofer und W. Matusik.
Towards real-time photorealistic 3D holography with deep neural networks.
In: *Nature* 591.7849 (2021), S. 234–239 (siehe S. 57).
- [181] Y. Wang, D. Dong, P. J. Christopher, A. Kadis, R. Mouthaan, F. Yang und T. D. Wilkinson.
Hardware implementations of computer-generated holography: a review.
In: *Optical Engineering* 59.10 (2020), S. 102413 (siehe S. 57).
- [182] H.-P. Gumm und M. Sommer. *Einführung in die Informatik.*
10. Aufl. München: Oldenbourg Wissenschaftsverlag, 2013 (siehe S. 58).
- [183] G. Barlas. *Multicore and GPU Programming. An Integrated Approach.*
Waltham: Morgan Kaufmann, 2015 (siehe S. 58).
- [184] C. M. R. Thimmannagari. *CPU Design. Answers to Frequently Asked Questions.*
New York: Springer, 2005 (siehe S. 58).
- [185] T. M. Aamodt, W. W. Lun Fung und T. G. Rogers.
General-Purpose Graphics Processor Architectures.
Synthesis Lectures on Computer Architecture Bd. 44.
San Rafael: Morgan & Claypool, 2018 (siehe S. 58).
- [186] R. Farber. *CUDA Application Design and Development.*
Waltham: Morgan Kaufmann, 2011 (siehe S. 58).

- [187] T. Ito, T. Yabe, M. Okazaki und M. Yanagi. *Special-purpose computer HORN-1 for reconstruction of virtual image in three dimensions*.
In: *Computer Physics Communications* 82.2-3 (1994), S. 104–110 (siehe S. 59).
- [188] T. Nishitsuji, Y. Yamamoto, T. Sugie, T. Akamatsu, R. Hirayama, H. Nakayama, T. Kakue, T. Shimobaba und T. Ito.
Special-purpose computer HORN-8 for phase-type electro-holography.
In: *Optics Express* 26.20 (2018), S. 26722–26733 (siehe S. 59).
- [189] O. Nishikawa, T. Okada, H. Yoshikawa, K. Sato und T. Honda.
High-speed Holographic-Stereogram Calculation Method Using 2D FFT.
In: *Proceedings of SPIE* 3010 (1997), S. 49–57 (siehe S. 59).
- [190] Y. Yamamoto, N. Masuda, R. Hirayama, H. Nakayama, T. Kakue, T. Shimobaba und T. Ito. *Special-purpose computer for electroholography in embedded systems*.
In: *OSA Continuum* 2.4 (2019), S. 1166–1173 (siehe S. 59).
- [191] J. Vince. *Quaternions for Computer Graphics*.
2. Aufl. London: Springer, 2021 (siehe S. 62).
- [192] P. Flach.
Machine Learning. The Art and Science of Algorithms that Make Sense of Data.
Cambridge: Cambridge University Press, 2012 (siehe S. 65, 70 f., 84 f., 113).
- [193] A. F. Holleman, N. Wiberg und E. Wiberg. *Lehrbuch der Anorganischen Chemie*.
102. Aufl. Berlin: De Gruyter, 2007 (siehe S. 65–67).
- [194] K. Becker und H. Ewest.
Die physikalischen und strahlungstechnischen Eigenschaften des Tantalkarbid.
In: *Zeitschrift für technische Physik* 5-6 (1930), 148–150 u. 216–220 (siehe S. 65).
- [195] C. Agte und H. Alterthum. *Untersuchungen über Systeme hochschmelzender Carbide nebst Beiträgen zum Problem der Kohlenstoffschmelzung*.
In: *Zeitschrift für technische Physik* 6 (1930), S. 182–191 (siehe S. 65).
- [196] C. Agte und J. Vacek. *Wolfram und Molybdän*.
Übers. von H. Herklotz. Berlin: Akademie-Verlag, 1959 (siehe S. 65 f.).
- [197] F. W. Clarke und H. S. Washington. *The composition of the Earth's Crust*.
Bericht. United States Geological Survey, 1924 (siehe S. 65).
- [198] A. A. Beus. *WALL ROCK ALTERATIONS OF HYDROTHERMAL PNEUMATOLYTIC DEPOSITS OF RARE ELEMENTS*.
In: *International Geology Review* 4.10 (1962), S. 1144–1153 (siehe S. 65).
- [199] H. Herrmann und H. Bucksch.
Wörterbuch GeoTechnik/Dictionary Geotechnical Engineering.
2. Aufl. Berlin, Heidelberg: Springer, 2013 (siehe S. 65).
- [200] M. Okrusch und H. E. Frimmel. *Mineralogie. Eine Einführung in die spezielle Mineralogie, Petrologie und Lagerstättenkunde*.
10. Aufl. Berlin, Heidelberg: Springer Spektrum, 2022 (siehe S. 65).
- [201] I. Noddack und W. Noddack. *Die Häufigkeit der chemischen Elemente*.
In: *Die Naturwissenschaften* 18.35 (1930), S. 757–764 (siehe S. 65).

-
- [202] E. W. Mayer und H. Schranz. *Flotation*. Chemie und Technik der Gegenwart Bd. 13. Leipzig: Hirzel, 1931 (siehe S. 66).
- [203] H. Schubert. *Aufbereitung fester Stoffe. Band 2: Sortierprozesse*. 4. Aufl. Stuttgart: Deutscher Verlag für Grundstoffindustrie, 1996 (siehe S. 66).
- [204] W. Eitel. *Kristallzüchtung*. In: *Ausgewählte Kapitel der präparativen Chemie*. Hrsg. von A. Stähler, E. Tiede und F. Richter. Handbuch der Arbeitsmethoden in der anorganischen Chemie Bd. 4. Berlin, Leipzig: De Gruyter, 1926, S. 448–473 (siehe S. 66).
- [205] H. W. King. *Crystal Structures of the Elements at 25°C*. In: *Bulletin of Alloy Phase Diagrams* 2.3 (1981), S. 401–402 (siehe S. 66).
- [206] J. Meija, T. B. Coplen, M. Berglund, W. A. Brand, P. De Bièvre, M. Gröning, N. E. Holden, J. Irrgeher, R. D. Loss, T. Walczyk und T. Prohaska. *Atomic weights of the elements 2013 (IUPAC Technical Report)*. In: *Pure and Applied Chemistry* 88.3 (2016), S. 265–291 (siehe S. 66).
- [207] H. Royer. *An application of high-speed microholography: the mertology of fogs*. In: *Nouvelle Revue d'Optique* 5.2 (1974), S. 87–93 (siehe S. 68).
- [208] H. Meng, W. L. Anderson, F. Hussain und D. D. Liu. *Intrinsic speckle noise in in-line particle holography*. In: *Journal of the Optical Society of America A* 10.9 (1993), S. 2046–2058 (siehe S. 69).
- [209] M. Malek, D. Allano, S. Coëtmelec und D. Lebrun. *Digital in-line holography: influence of the shadow density on particle field extraction*. In: *Optics Express* 12.10 (2004), S. 2270–2279 (siehe S. 69).
- [210] S. Valeri, G. C. Gazzadi und A. di Bona. *Scattering interference of energetic electrons along atomic chains: The effect of the atomic environment*. In: *Physical Review B* 52.19 (1995), S. 14048–14057 (siehe S. 69).
- [211] M. Kohl. *Performance Measures in Binary Classification*. In: *International Journal of Statistics in Medical Research* 1.1 (2012), S. 79–81 (siehe S. 73 f., 113).
- [212] A. Tharwat. *Classification assessment methods*. In: *Applied Computing and Informatics* 17.1 (2021), S. 168–192 (siehe S. 84, 113).
- [213] S. Barocas und A. D. Selbst. *Big Data's Disparate Impact*. In: *California Law Review* 104.3 (2016), S. 671–732 (siehe S. 84).
- [214] A. K. Menon und R. C. Williamson. *The Cost of Fairness in Binary Classification*. In: *Proceedings of the 1st Conference on Fairness, Accountability and Transparency*. Proceedings of Machine Learning Research 81 (2018), S. 107–118 (siehe S. 84).

- [215] C. Goutte und E. Gaussier. *A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation*.
In: *ECIR 2005: Advances in Information Retrieval*.
Hrsg. von D. E. Losada und J. M. Fernández-Luna.
Lecture Notes in Computer Science Bd. 3408. Berlin, Heidelberg: Springer, 2005, S. 345–359 (siehe S. 84).
- [216] E. B. Fowlkes und C. L. Mallows.
A Method for Comparing Two Hierarchical Clusterings.
In: *Journal of the American Statistical Association* 78.383 (1983), S. 553–569 (siehe S. 84, 113).
- [217] G. W. Milligan und M. C. Cooper. *A Study of the Comparability of External Criteria for Hierarchical Cluster Analysis*.
In: *Multivariate Behavioral Research* 21.4 (1986), S. 441–458 (siehe S. 84, 113).
- [218] J. T. Schaefer. *The Critical Success Index as an Indicator of Warning Skill*.
In: *Weather and Forecasting* 5.4 (1990), S. 570–575 (siehe S. 84, 113).
- [219] W. Demtröder. *Experimentalphysik 2. Elektrizität und Optik*.
5. Aufl. Berlin, Heidelberg: Springer, 2009 (siehe S. 88).
- [220] K. A. Grebenyuk. *Classical sampling theorem in digital holography*.
In: *Journal of Physics: Conference Series* 536 (2014), S. 012005 (siehe S. 88).
- [221] I. I. Hirschman und D. V. Widder. *The Convolution Transform*.
Princeton Mathematical Series Bd. 20.
Princeton: Princeton University Press, 1955 (siehe S. 90).
- [222] B. Jähne. *Digitale Bildverarbeitung*.
6. Aufl. Berlin, Heidelberg: Springer, 2005 (siehe S. 90).
- [223] C. Kohlmann, L. Kesper, R. Hönig, T. Muro, Y. Takagi, A. Yasui und C. Westphal. *Hard X-ray Photoelectron Diffraction Applied to the System Pyrite (FeS₂) for Direct Atom Imaging by means of Chemical Sensitive Photoelectron Holography: proof of principle study*.
In: *Spring-8/SACLA Research Report* 9.4 (2021), S. 195–199 (siehe S. 97).
- [224] R. Torres, P. Pellat-Finet und Y. Torres. *Sampling Theorem for Fractional Bandlimited Signals: A Self-Contained Proof. Application to Digital Holography*.
In: *IEEE Signal Processing Letters* 13.11 (2006), S. 676–679 (siehe S. 98).
- [225] G. Van Rossum und F. L. Drake. *Python 3 Reference Manual*.
Scotts Valley: CreateSpace, 2009 (siehe S. 107).
- [226] C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith et al.
Array programming with NumPy.
In: *Nature* 585.7825 (2020), S. 357–362 (siehe S. 107).
- [227] S. Hoyer und J. J. Hamman. *xarray: N-D labeled Arrays and Datasets in Python*.
In: *Journal of Open Research Software* 5.1 (2017), S. 10 (siehe S. 107).
- [228] R. Rew und G. Davis. *NetCDF: An Interface for Scientific Data Access*.
In: *IEEE Computer Graphics and Applications* 10.4 (1990), S. 76–82 (siehe S. 107).

-
- [229] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright et al. *SciPy 1.0: fundamental algorithms for scientific computing in Python*. In: *Nature Methods* 17.3 (2020), S. 261–272 (siehe S. 107).
- [230] S. K. Lam, A. Pitrou und S. Seibert. *Numba: A LLVM-based Python JIT Compiler*. In: *LLVM '15: Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC* (2015), S. 7 (siehe S. 107).
- [231] L. P. René de Cotret, M. R. Otto, M. J. Stern und B. J. Siwick. *An open-source software ecosystem for the interactive exploration of ultrafast electron scattering data*. In: *Advanced Structural and Chemical Imaging* 4.1 (2018), S. 11 (siehe S. 107, 111).
- [232] J. B. Pendry. *Reliability factors for LEED calculations*. In: *Journal of Physics C: Solid State Physics* 13.5 (1980), S. 937–944 (siehe S. 112).

Danksagungen

Annähernd fünf Jahre des Forschens und Experimentierens finden mit dieser Arbeit ihr Ende. Dabei verlief diese Zeit keineswegs so geradlinig, wie es der vorgehende Text suggerieren könnte. Vielmehr war der Weg genauso gezeichnet von Hindernissen und Herausforderungen, Abzweigungen und Sackgasen. Es bedurfte vielerlei Unterstützung, die in den Abschluss dieser Promotion mündete. Anknüpfend an das dieser Arbeit vorangestellte Zitat stelle ich fest: inzwischen scheint das fern und vergangen. Es bleibt die vorliegende Ausarbeitung. An dieser abschließenden Stelle möchte ich die Gelegenheit nutzen, um allen Beteiligten ganz herzlich dafür zu danken.

Der Konrad-Adenauer-Stiftung danke ich für die mir zuteil gewordene ideelle und finanzielle Unterstützung, die mir viele Freiheiten in meiner Arbeit erlaubte und mir Gelegenheit gab, meinen Horizont im Austausch mit Konstipendiaten zu erweitern. Insbesondere Herrn Dr. Johannes Gold für die Betreuung und so manches aufmunternde Wort in schwieriger Zeit.

Für die durchgehende Betreuung auch über einen Themenwechsel hinweg bin ich Herrn Prof. Dr. Carsten Westphal sehr dankbar, der mir viele Möglichkeiten in der Ausgestaltung des vorgeschlagenen Themas einräumte, die Arbeit aufmerksam begleitete und zeitgleich ein steter Berater in fachlichen aber auch persönlichen Fragen war. Herrn Prof. Dr. Dr. Rhode danke ich für die Übernahme des Zweitgutachtens dieser Arbeit.

Dem Team um den für sämtliche Simulationen verwendeten HPC-Cluster LiDO3 danke ich für die zur Verfügung gestellten Ressourcen und Hilfestellung bei der Optimierung von Laufzeiten.

Ich danke der Arbeitsgruppe Westphal des Lehrstuhls E1 für hilfsbereite Unterstützung

und ein immer einladendes Arbeitsklima. Insbesondere Herrn Julian Hochhaus für seine Expertise im Fitten von Daten und Herrn Dr. Malte Schulte für sein gründliches Lektorat ausgewählter Kapitel und manchen Austausch zu typographischen Fragen.

Herrn Dr. Thomas Fuchs danke ich ebenfalls für sein sorgfältiges Lektorat. Lieber Thomas, herzlichen Dank für so manches aufgewertete Seminar, regelmäßiges Aufbauen und eine umfangreiche Stadtführung.

Ich danke Herrn Dr. Philip Bleicker für seine Begleitung in der Anfertigung dieser Dissertation. Für etliche Diskussionen über fachliche und völlig fachfremde Problemstellungen, stetige Verfügbarkeit bei Fragen und gründliche Lektorate. In neun Jahren des gemeinsamen Studiums hat viel zusammengeschweißt: so manche Übungsgruppe, das Praktikum, die Klausurvorbereitungen und so manches angefangene und wieder aufgegebene Projekt. Dafür danke ich Dir sehr herzlich.

Meiner Familie danke ich für die dauerhafte Unterstützung in allen Lebenslagen, das Ertragen so mancher Laune, das stete Mittragen jeder Unternehmung und das Auffangen in schweren Zeiten. Insbesondere auch Euch, Tobias, Missy und Mama für Eure Lektorate dieser Arbeit. Liebe Katharina, herzlichen Dank für Deine Begleitung auf diesem Weg. Für die Nachsicht gegenüber so mancher Eigenheit und das Stützen, wenn es scheinbar nicht weiter ging. Euch allen danke ich für den Rückhalt und das Ermöglichen dieser Arbeit.

Gevelsberg, im September 2022

Matthias Gianfelice

Eidesstattliche Versicherung

Ich versichere hiermit an Eides statt, dass ich die vorliegende Dissertation mit dem Titel „Elektronenholographie mittels Elektronenrückstreubeugung“ selbstständig und ohne unzulässige fremde Hilfe erbracht habe. Ich habe keine anderen als die angegebenen Quellen und Hilfsmittel benutzt sowie wörtliche und sinngemäße Zitate kenntlich gemacht. Die Arbeit hat in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegen.

Dortmund, den 09.09.2022

Ort, Datum

Matthias Gianfelice

Unterschrift

Belehrung

Wer vorsätzlich gegen eine die Täuschung über Prüfungsleistungen betreffende Regelung einer Hochschulprüfungsordnung verstößt, handelt ordnungswidrig. Die Ordnungswidrigkeit kann mit einer Geldbuße von bis zu 50000 € geahndet werden. Zuständige Verwaltungsbehörde für die Verfolgung und Ahndung von Ordnungswidrigkeiten ist der Kanzler/die Kanzlerin der Technischen Universität Dortmund. Im Falle eines mehrfachen oder sonstigen schwerwiegenden Täuschungsversuches kann der Prüfling zudem exmatrikuliert werden (§ 63 Abs. 5 Hochschulgesetz – HG).

Die Abgabe einer falschen Versicherung an Eides statt wird mit Freiheitsstrafe bis zu drei Jahren oder mit Geldstrafe bestraft.

Die Technische Universität Dortmund wird ggf. elektronische Vergleichswerkzeuge zur Überprüfung von Ordnungswidrigkeiten in Prüfungsverfahren nutzen.

Die oben stehende Belehrung habe ich zur Kenntnis genommen.

Dortmund, den 09.09.2022

Ort, Datum

Matthias Gianfelice

Unterschrift