PHILIP F. CHIMENTO

TWO-DIMENSIONAL OPTICS

DIFFRACTION AND DISPERSION OF SURFACE PLASMONS

Published by Casimir Research School, Delft, The Netherlands

Cover design: Philip F. Chimento

Printed by Gildeprint Drukkerijen, The Netherlands

ISBN: 978-90-8593-155-3

TWO-DIMENSIONAL OPTICS
DIFFRACTION AND DISPERSION OF SURFACE PLASMONS

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 22 mei 2013
klokke 11:15 uur

door

Philip Chimento
geboren te Raleigh, North Carolina, Verenigde Staten
in 1981

PROMOTIECOMMISSIE

Promotores:	Prof. dr. E. R. Eliel	Universiteit Leiden
	Prof. dr. G. W. 't Hooft	Philips Research & Universiteit Leiden
Leden:	Dr. M. P. van Exter	Universiteit Leiden
	Dr. C. Genet	Université de Strasbourg
	Prof. dr. J. Gómez Rivas	Philips Research & TU Eindhoven
	Dr. H. L. Offerhaus	Universiteit Twente
	Prof. dr. T. D. Visser	Vrije Universiteit & TU Delft
	Prof. dr. J. P. Woerdman	Universiteit Leiden

This work is part of the research program of the Foundation for Fundamental Research on Matter (FOM), which is part of the Netherlands Organization for Scientific Research (NWO).

An electronic version of this disseration is available at the Leiden University Repository (https://openaccess.leidenuniv.nl).

Casimir PhD series, Delft-Leiden, 2013-14

To caffeine
To the future, let's hope it's worth saving

Contents

Preface (Introduction for non-scientists)

A *surface plasmon* is a light wave that is trapped on a flat two-dimensional surface, henceforth called Flatland, as in a famous novel.[1] Many devices and effects that are familiar from normal, three-dimensional optics also exist in Flatland, usually created by applying some sort of material or structure to the surface. For example, there are mirrors (Figure 1) and lenses (Figure 2).

Surface plasmons are special because they can only exist on the boundary surface between a metal that conducts electricity very well, like silver or gold, and a non-metal substance, such as glass, plastic, or air. (As always in science, there are exceptions to this rule: semiconductors can also work instead of metals,[2] and just recently a layer of graphene was proposed.[3]) The metal has to be a good conductor, so that some of the electrons belonging to its atoms, called *free electrons*, can move more or less unhindered through the metal, from atom to atom. Surface plasmons cannot exist without free electrons.

In an ocean wave, the water level rises and dips, but nothing like that happens in Flatland. In a surface plasmon, the wave is connected to back-and-forth movements of the free electrons in the metal. Moving electrons create a wave on the outside of the metal, and the wave moves other electrons inside the metal, which is how surface plasmons move.

Surface plasmons being trapped on the metal's surface, or *confined*, is actually one of the desirable properties of surface plasmons, and partly explains why they are such a popular research subject. Transforming light into surface plasmons allows light to squeeze into tiny spaces, smaller than it would otherwise be able to fit into. The more you try to confine regular light, the more it tends to spread out, and if you try to cram light into a channel that is too small to contain it (less than half the light's wavelength), then it simply won't fit.[4] Surface plasmons, however, can be stuffed into tiny strip-shaped metal channels[5] or grooves.[6] Researchers can use them to develop ultra-small components for circuits that carry

[1] Abbott, 1884.

Figure 1: A Flatland mirror, seen from above; it is tilted at 45°. The surface plasmon enters from the right side of the figure and is partially reflected downwards. (Reprinted from González et al. (2006), with kind permission of the author. Copyright 2006, the American Physical Society.)

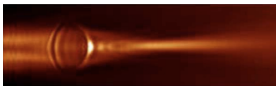


Figure 2: A Flatland lens, seen from above. The surface plasmon enters from the left side of the figure and is focused to a small spot about halfway through the figure. (Reprinted from Devaux et al. (2010), with kind permission of the Optical Society of America.)

[2] Gómez Rivas, Kuttge, Kurz, Haring Bolivar, and Sánchez-Gil, 2006.

[3] Gorbach, 2013.

[4] Although that is a simplification; more about that in chapter 2.

[5] Maier and Atwater, 2005.

[6] García-Vidal, 2006.

light instead of electricity: this is the field of *nanophotonics*.[7] Even though surface plasmons were discovered over fifty years ago, nanotechnology has only caught up in the last ten to fifteen years and made nanophotonics possible.

[7] Nano is a word meaning, roughly, smaller than one micron. It refers to the realm of objects the size of a cell membrane, a virus, or one one-thousandth of a human hair.

Another important property of surface plasmons is that confining the light into a small space squeezes all the energy it carries into a small space too — this effect is known as *field enhancement*. Researchers can then do processes that require a lot of energy without needing a lot of light, because all the light's energy is concentrated in one tiny place. This is also important for antennas in nanophotonics.[8] An antenna is nothing more than a device that converts free radiation (cell tower signals) into localized energy (the electronics in your mobile phone) and vice versa. We can engineer tiny metal antennas in such a way that they have a single spot where the field enhancement is very large. If we position a molecule at that spot, the molecule can broadcast its energy very efficiently through the antenna. So a good plasmonic antenna is an efficient bridge between molecule-sized phenomena and human-scale signals in the laboratory.

[8] Novotny and van Hulst, 2011.

One problem with engineering photonic devices is that light is damped when it interacts with metals. The light's energy is simply converted into heat. Obviously, that is a nuisance, but as is usual in science, someone has figured out a way to turn it into an advantage. Researchers are working on an experimental cancer treatment that works with tiny particles called "nanoshells" coated with a thin layer of metal.[9] These nanoshells can be attached to antibodies that seek out cancerous tissue and congregate in the tumor cells. Infrared light normally passes harmlessly through body tissue; however, the nanoshells are engineered to act as receiver antennas for the light, absorbing it and concentrating all the energy in a small space. The resulting release of heat kills the cancer cells.

[9] Loo, Lowery, Halas, West, and Drezek, 2005.

For further reading, I recommend a 2007 *Scientific American* article about surface plasmons and their applications.[10] For a more technical review of the latest developments, there is an open-access article in *Journal of Physics D*,[11] only a few months old at the time of writing.

[10] Atwater, 2007.

[11] Hayashi and Okamoto, 2012.

Polarization and holes in metal sheets

THE CURRENT WAVE of surface plasmon research was unleashed when Thomas Ebbesen at NEC Corporation asked a technician to make a grid of tiny holes in a metal sheet. When Ebbesen picked up the sheet he was surprised that he could partially see through it even though the holes were supposedly tiny enough that hardly any light should have been able to

get through. Moreover, the transmitted light was colored, and the color changed when he turned the sheet and viewed it at an angle. He first thought the technician had made a mistake and drilled the holes too large, but when it became clear that there was nothing wrong with the holes, he and his co-workers hit upon surface plasmons as an explanation. This resulted in a landmark paper[12] and the discovery of an effect called *extraordinary optical transmission*.

[12] Ebbesen, Lezec, Ghaemi, Thio, and Wolff, 1998.

Simply speaking, light falling on a small enough hole in a thin metal sheet launches surface plasmons into Flatland from the edge of the hole. These surface plasmons travel across the metal sheet to the next hole, where they turn into light again and pass through the hole. This extra light augments the small amount of light that was already traveling through the hole; combined, enough light travels through the metal sheet for it to be translucent.

Chapter 2 describes research where the opposite result occurs; we studied rectangular-shaped slits in a metal sheet, but only one slit at a time. Studying the slits in isolation means that they still launch surface plasmons from their edges, but since the surface plasmons have no other slits to go to, they just travel through Flatland to nowhere and eventually die out. This actually causes *less* light to make it through the slit than otherwise would.

We use this effect to create a tiny version of a device called a *quarter-wave plate*. It takes light with *linear polarization* and converts it into *circular polarization*. Polarization is best thought of as two people, Alice and Bob, holding opposite ends of a long rope (Figure 3). If Alice wishes to send a wave to Bob over the rope, she shakes her end back and forth, and a wave travels down the rope to Bob. The rope oscillates back and forth in one plane, and we call this *linear polarization*. However, Alice can also spin her end of the rope in a circle, in which case a circular wave travels down the rope to Bob; we call this *circular polarization*. There are even two variations, depending on whether Alice spins clockwise or counterclockwise, called *left-handed* and *right-handed*. Light can do the exact same thing.

Since a quarter-wave plate converts linear polarization to circular, it is as if Alice sends a linearly polarized wave to Bob, but by the time it reaches Bob it has become circularly polarized. This polarization change doesn't happen in a rope, but it does happen in light. Quarter-wave plates are present in 3D projection systems: the image meant for your left eye is projected with left-handed circular polarization, and the one for your right eye is right-handed. Filters in your 3D glasses make sure that each

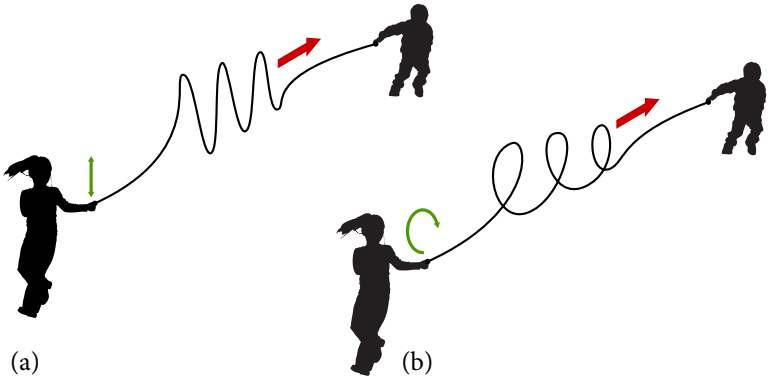


Figure 3: (a) Alice is sending a linearly polarized wave to Bob. (b) Alice is sending a left-handed circularly polarized wave to Bob.

eye only sees the correct image.

Just as in Ebbesen's extraordinary optical transmission, this miniaturized quarter-wave plate also stems from an accidental discovery: in 2005 my co-worker Nikolay Kuzmin discovered by chance that circularly polarized light was coming out of the back sides of slits in a metal sheet.[13]

[13] Kuzmin, 2008, pp. 73–87.

We continued this line of research in chapter 3, using the effects explored in chapter 2 to build a device that converts light with *spin angular momentum* into light with *orbital angular momentum*. Consider the Earth: our planet has spin angular momentum because it revolves around its own axis, causing day and night. The Earth also has orbital angular momentum because it orbits around the Sun, which causes the seasons. Particles of light also have both of these kinds of angular momentum. In fact, a light particle spinning around like a top is just another way of looking at circular polarization. Light is both a particle and a wave; spin angular momentum is to a light particle as circular polarization is to a light wave.

Phase vortices

Similarly, orbital angular momentum is to a light particle as *phase vortex* is to a light wave. "Phase vortex" sounds as if it came straight out of *Star Trek*, but it can be easily illustrated with a nice piece of oceanographical research from NASA[14] (Figure 4.)

[14] Ray, 2006.

The ocean tides exhibit phase vortices. For example, we can see one in the middle of the Pacific Ocean, at about 15° south and 150° west. It is not a whirlpool; in fact, if you were to travel there, you would notice nothing special about that spot. The white lines indicate places where it is high tide at the same time; each line is an hour earlier or later than its neighbor. We could say that the instant of high tide "rotates" around the

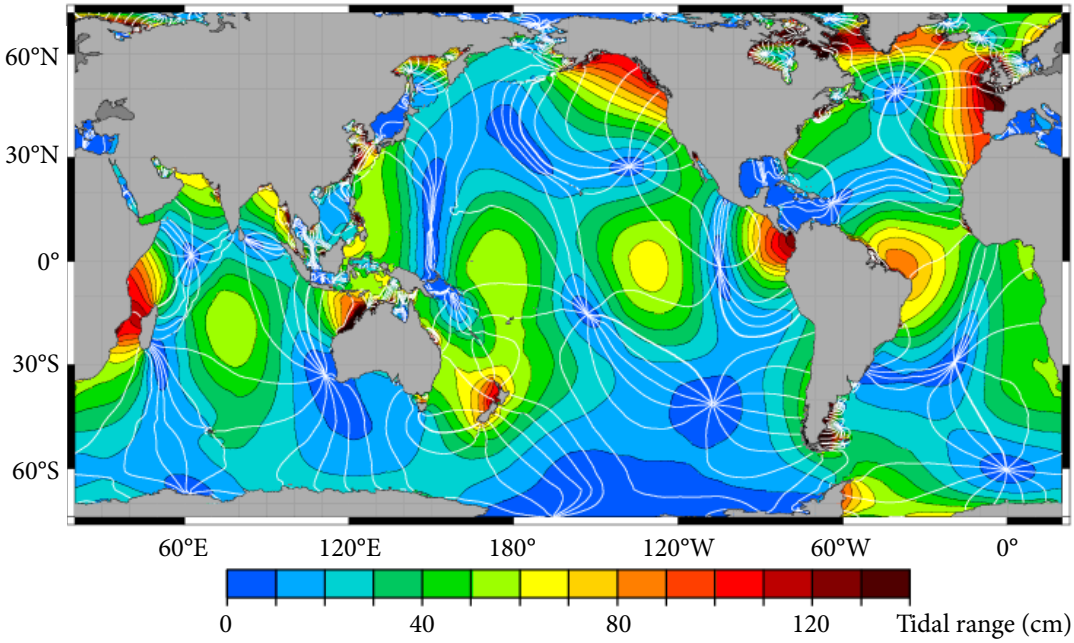


Figure 4: Tides in the world's oceans. Along each white line, it is high tide at exactly the same time, and neighboring white lines' high tides are separated by one hour. The colored spaces represent the tide strength (blue is weaker and red is stronger), and the amplitude is indicated in centimeters. (Public domain image. Credit to NASA Goddard Space Flight Center; NASA Jet Propulsion Laboratory; Scientific Visualization Studio; Television Production NASA-TV/GSFC. Special thanks to Dr. Richard Ray, Space Geodesy branch, NASA/GSFC)

phase vortex. So when the lines all meet in the center of the phase vortex, then when is it high tide? Always? Never? The answer is that the phase vortices are always in the blue regions where the tide is weakest: there is no tide there!

Light can do the same thing: we can create a laser beam where the "high tide" of the light wave rotates around the center of the laser beam. Since there is no tide in the center, the laser beam is dark there. We can also have two, three, or more high tides rotating around one laser beam, and they can go clockwise or counterclockwise: the number of high tides is called *topological charge*. These laser beams with dark holes in the middle, also called *donut beams*, can be used to encode information densely.[15]

So, converting spin angular momentum to orbital angular momentum, as we describe in chapter 3, means that we create a donut beam out of a circularly polarized laser beam with no donut. Returning to the example of the Earth, it is as if we could make it spin slower but orbit faster, thereby lengthening the day and shortening the year. Hitting the Earth with a giant comet would do that, but that requires the comet to contribute its angular momentum. With our device, we can create donut beams nondestructively, without adding or losing angular momentum.

One problem with donut beams is that they are hard to identify. The donut part is easy to see, but light waves are so fast (trillions of cycles per

[15] Padgett, Courtial, and Allen, 2004; Molina-Terriza, Torres, and Torner, 2007.

second) that it's completely impossible to measure the number of high tides directly, or the direction in which they are rotating. However, we need to know this information in order to use the beams. The usual way of determination involves using a second laser beam to probe the first one. However, with the research in chapter 4 we have created a device where the second laser beam is not necessary. We used another metal sheet with two slits: one to launch surface plasmons into Flatland and one to catch them and take them out of Flatland. The slit takes a "slice" of the phase vortex, similar to the way that a longitude line in Fig. 4 represents a slice of the ocean. By examining the surface plasmons caught at the second slit, we can recover all the information we need about the phase vortex: topological charge, and whether it is rotating clockwise or counterclockwise.

Aluminum and solitons

THE SECOND PART of this dissertation examines surface plasmons traveling on an aluminum surface. Most surface plasmon research uses gold or silver because these materials absorb less light. Aluminum absorbs light in a certain frequency range, which makes it less desirable for some uses. However, we are interested in aluminum specifically because of this feature!

The familiar shiny gray color of a polished metal surface arises when the metal reflects all the colors of visible light approximately equally. Figure 5 shows what percentage of light of each color aluminum reflects (as well as light that is invisible to the human eye and therefore has no color: ultraviolet to the left and infrared to the right.) We see that this is true for aluminum as well, but something else happens in the infrared, to the right of the visible part of the spectrum: aluminum reflects less light and absorbs more.

In order to see why this is interesting we need to think about a *soliton*: a short wave that remains unchanged as it travels along. Ocean waves, for example, break and disperse, while short light pulses get longer and spread out more as they travel. Solitons, on the other hand, do not. John Scott Russell discovered them by chance in 1834 when he observed one in Scotland's Union Canal. Figure 6 shows a modern-day reconstruction of the discovery. Solitons remained an oddball curiosity for over a century until they found an application in fiber-optic communications: the ability to send a pulse of light through a fiber without any distortion proved invaluable.

Since pulses disperse in any material, the existence of a soliton depends

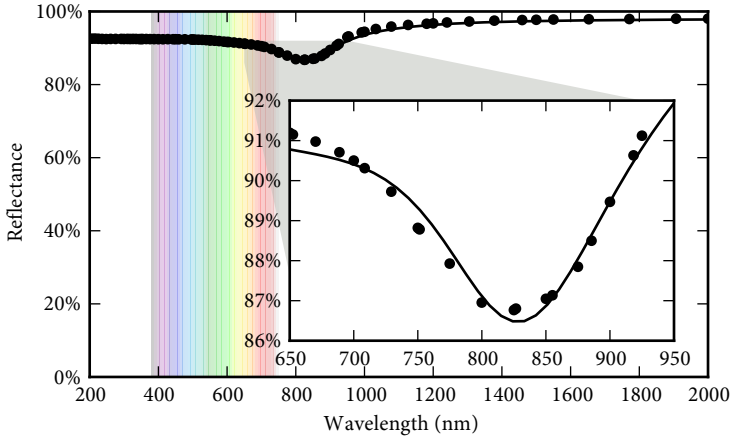


Figure 5: Percentage of light reflected by aluminum, versus wavelength of the light (color). To the left is ultraviolet light (invisible to human eyes), then visible light (indicated by a rainbow of colors), and to the right of that is infrared light (again invisible.) About 93% of visible light is reflected, but there is a downward dip in the near infrared, which is what we are interested in here. The dots and the solid line indicate measurements from two different sources (Smith, Shiles, and Inokuti, 1985; Rakić, Djurišić, Elazar, and Majewski, 1998.) The inset shows the interesting region in detail.

on two opposite effects that counterbalance each other. One of these effects, where blue light travels faster through the material than red light, is called *anomalous dispersion*.[16] It is often found paired with an absorption such as that of aluminum, shown in Fig. 5. Anomalous dispersion can be counterbalanced by something called the *Kerr effect* in order to create a soliton, whereas normal dispersion can't. This is why anomalous dispersion is an interesting subject of research.

[16] This is a terribly inappropriate name, because anomalous dispersion is not an anomaly at all: it is present in most materials. 'Anomalous' implies that it's rare or not understood, but in this case it's simply the opposite of normal dispersion, where red light travels faster than blue light in a material.

Aluminum itself is not a good material for transporting solitons, for the simple reason that it is opaque and therefore not very good at transporting light pulses. However, aluminum paired with another material could create some of the conditions for a surface plasmon soliton! It is with this in mind that we conducted the research described in chapters 5, 6, and 7.

Chapter 5 is a discussion of how to conduct and interpret surface plasmon measurements. We discovered in the course of our aluminum research that one of the usual methods for probing surface plasmons, called the *Kretschmann configuration*, does not work as well for surface plasmons on aluminum as it does for gold and silver. The aluminum measurements are more difficult to interpret, so one of the new findings of chapter 5 is an effective method of interpreting them. We also find that a different arrangement of the experiment, called the *Otto configuration*, is actually quite useful under these circumstances, even though it is usually considered less effective than the Kretschmann configuration.

In chapter 6, we demonstrate the first measurements of anomalous dispersion for surface plasmons, and in chapter 7, we show how cooling down the aluminum with liquid nitrogen enhances the effect quite a bit.

Figure 6: Modern-day reenactment of Scott Russell's discovery of solitons in the Union Canal. (Figure reprinted from "Soliton wave receives crowd of admirers" (1995), with license.) The soliton is visible as the "mountain" of water behind the boat.

With this discovery, we are one step closer to creating a surface plasmon soliton and the Flatland analog to fiber-optic technology.

THE COMMON THEME through all of this research is finding and exploring the Flatland equivalents of phenomena from optics in the normal, three-dimensional world, such as solitons. We also exploit Flatland effects to bring about other phenomena in three-dimensional optics, such as spin-to-orbit conversion. In the future, as surface plasmons become more and more important, from antennas to sensors to curing cancer, technology will move more into Flatland. For this, we need to have a Flatland 'engineering toolbox' that is as complete as our three-dimensional engineering toolbox that has been filled gradually over the past few centuries.

1
Introduction

A SURFACE PLASMON is a light wave bound to a metal surface, first predicted in 1957 as a side-effect of bombarding metal films with fast electrons[1] and observed two years later.[2] Surface plasmons occur in many different geometries of metal, from nanoparticles[3] to flat metal surfaces. This dissertation is about the latter type, which propagates along the two-dimensional metal surface, as opposed to 'normal' light which travels through three-dimensional space, as a sort of two-dimensional light wave.[4] The surface plasmon's restriction to the metal surface allows us to send optical signals through channels of extremely small size.[5]

This chapter is a short scientific introduction to the work described in this thesis. Readers interested in an introduction accessible to non-scientists should turn to page 1.

[1] Ritchie, 1957.

[2] Powell and Swan, 1959.

[3] Moskovits, 1985.

[4] Bell et al., 1975; Bozhevolnyi and Pudonin, 1997; Ditlbacher, Krenn, Schider, Leitner, and Aussenegg, 2002.

[5] Maier and Atwater, 2005; García-Vidal, 2006.

1.1 Devices using subwavelength slits in metal films

IT HAS BEEN KNOWN FOR SOME TIME[6] that a very narrow slit or scratch in a thin metal film, under certain circumstances, can convert incident light with the correct polarization into surface plasmons, and vice versa. Which metal is used makes an important difference.

Such conversions take a three-dimensional optical mode and change it to a corresponding two-dimensional one, and back again.[7] This conversion is sensitive to the mode's local phase front. This influence of phase, and the aforementioned sensitivity to polarization, allow all sorts of surface plasmon effects having to do with polarization, phase, or both.

In chapter 2, taking a metal film made of gold, 200 nm thick, we examine how much light is transmitted through slits of varying thicknesses from 50 to 500 nm, and how this transmitted light is polarized. Surface plasmons are excited at one particular polarization, and we exploit this to control the polarization of the transmitted light. At a certain slit width and film thickness, the slit turns out to be able to convert linearly polarized light into circularly polarized, and vice versa. We also developed a sim-

[6] Jasperson and Schnatterly, 1969; Sánchez-Gil and Maradudin, 1999; Lalanne, Hugonin, and Rodier, 2005; Schouten, Kuzmin, et al., 2005.

[7] Altewischer, van Exter, and Woerdman, 2002.

ple model that explains this phenomenon in an intuitive way, by viewing the slit as a waveguide with imperfectly conducting walls, albeit a very short one. Our approach provides a convenient way to implement the functionality of a quarter-wave plate at a very small scale.

We exploit this phenomenon again in chapter 3; this time with circular slits, where the slit creates an optical vortex from circularly polarized light, thereby converting optical spin angular momentum[8] to optical orbital angular momentum.[9] This curious interaction due to symmetry has been studied in birefringent materials,[10] space-variant gratings,[11] and even in electron beams.[12]

Chapter 4 describes an experiment with two very narrow slits milled parallel to each other in a very thin gold film. One slit is illuminated with light, and at the slit it is partly converted to surface plasmons. The surface plasmons travel across the film to the other slit, where they are converted once again into light, and we record the light intensity distribution; during transit, the shape of the plasmon wavefront changes due to diffraction. We use this diffraction to retrieve information about the incident light's phase; the phase cannot be measured directly, a well-known problem in physics,[13] and is usually probed using interference with a second light beam.[14] In order to demonstrate this technique, we measure the phase of beams containing various optical vortices. This technique could produce a wavefront sensor with a much higher spatial resolution than achievable with the usual techniques, which could be interesting for astronomy and UV lithography.

[8] Beth, 1936.

[9] Allen, Beijersbergen, Spreeuw, and Woerdman, 1992.

[10] Ciattoni, Cincotti, and Palma, 2003; Marrucci, Manzo, and Paparo, 2006.

[11] Bomzon, Kleiner, and Hasman, 2001; Lombard, Drezet, Genet, and Ebbesen, 2010.

[12] Karimi, Marrucci, Grillo, and Santamato, 2012.

[13] Fienup, 1982.

[14] Baranova, Zel'dovich, Mamaev, Pilipetskii, and Shkunov, 1981.

1.2 *Anomalous dispersion of surface plasmons*

DISPERSION IS THE PHENOMENON of the velocity of light in a material depending on the light's wavelength. For example, if we send a pulse of red light and a pulse of blue light into a glass brick at the same instant, they will emerge from the other side at different times. Usually the red light arrives earlier than the blue light (which is called "normal dispersion"), but sometimes the reverse is true: "anomalous" dispersion. Anomalous dispersion is a prerequisite for solitons, light pulses that can travel a long distance without changing their shape. Anomalous dispersion is needed to balance the normal dispersion caused by the other prerequisite for solitons, the Kerr effect. When the two occur together, they can cancel each other out, allowing a pulse that propagates without changing. There have been several, mainly mathematical, proposals for surface plasmon soliton pulses in recent years.[15]

[15] Feigenbaum and Orenstein, 2007; Bliokh, Bliokh, and Ferrando, 2009; Davoyan, Shadrivov, and Kivshar, 2009; Sámson, Horak, MacDonald, and Zheludev, 2011; Marini, Skryabin, and Malomed, 2011; Walasik, Nazabal, Chauvet, Kartashov, and Renversez, 2012.

Anomalous dispersion mostly occurs in the neighborhood of wavelengths that the material absorbs. An ideal metal behaves according to the free-electron model, or Drude model, where there are no absorptions, and the dispersion is always normal, in the frequency region of metallic behavior. However, the metal aluminum has an absorption in the near infrared,[16] a so-called parallel-band absorption,[17] which we aim to exploit. In the second part of this work, we try to answer the question of whether this absorption also causes surface plasmons on an aluminum surface to have anomalous dispersion. We probe this using a method where the surface plasmons are excited by incoming light from a prism. This technique has two variations, named after the German physicists Kretschmann[18] and Otto.[19] The Otto configuration is generally considered to be disadvantageous compared to the Kretschmann configuration. In chapter 5, we show that that is a misconception. In addition, we introduce a method of analysis with which we can properly interpret experimental results using lossy metals in both configurations, which is impossible with the usual approach.

Chapters 6 and 7 describe the results of measuring surface plasmons with anomalous dispersion. In chapter 6 we demonstrate the existence of surface plasmons with anomalous dispersion on an aluminum surface. Subsequently, in chapter 7, we increase the degree of anomalous dispersion (expressed in the second-order dispersion parameter) by a great deal, by cooling the metal to near the temperature of liquid nitrogen, approximately 86 K. However, there is a tradeoff between more anomalous dispersion and more surface plasmon loss, because the surface plasmons decay more quickly in the low-temperature metal: the parallel-band absorption also becomes stronger at low temperatures.[20]

[16] Strong, 1936; Bennett, Silver, and Ashley, 1963.

[17] Harrison, 1966.

[18] Kretschmann, 1971.

[19] Otto, 1968.

[20] Liljenvall, Mathewson, and Myers, 1971; Mathewson and Myers, 1972.

PART I

DEVICES USING SUBWAVELENGTH SLITS IN METAL LAYERS

2

A subwavelength slit as a quarter-wave retarder

We have experimentally studied the polarization-dependent transmission properties of a nanoslit in a gold film as a function of its width. The slit exhibits strong birefringence and dichroism. We find, surprisingly, that the transmission of the polarization parallel to the slit only disappears when the slit is much narrower than half a wavelength, while the transmission of the perpendicular component is reduced by the excitation of surface plasmons. We exploit the slit's dichroism and birefringence to realize a quarter-wave retarder.

2.1 Introduction

This chapter was previously published as: Chimento, Kuzmin, et al. (2011).

The study of the transmission of light through small perforations in metal films has a venerable history[1] and has important applications in the field of optical data storage.[2] It dates back to the middle of the nineteenth century when Fizeau described the polarizing effect of wedge-shaped scratches in such films.[3]

[1] Lord Rayleigh, 1897; Bethe, 1944; Bouwkamp, 1954; Jones and Richards, 1954.

[2] Bouwhuis et al., 1985.

[3] Fizeau, 1861.

This field has recently come back to center stage following the observation that, at a specific set of wavelengths, the transmission of a thin metal film containing a regular two-dimensional array of subwavelength apertures is much larger than elementary diffraction theory predicts.[4] This phenomenon of extraordinary optical transmission, which is commonly attributed to surface plasmons traveling along the corrugated interface, has spawned many studies of thin metal films carrying variously-shaped corrugations and perforations. These include holes with circular, cylindrical, or rectangular cross sections,[5] either individually or in arrays, and elongated slits.[6] The polarization of the incident light is an important parameter, in particular when the width of the hole or slit is subwavelength in one or both directions. The case of a slit which is long in one dimension

[4] Ebbesen et al., 1998.

[5] Astilean, Lalanne, and Palamaru, 2000.

[6] Takakura, 2001; Yang and Sambles, 2002; Suckling et al., 2004.

and subwavelength in the other seems particularly simple, as elementary waveguide theory predicts that it acts as a perfect polarizer when the slit width is less than about half the wavelength of the incident light.

TE: transverse electric
TM: transverse magnetic

For infinitely long slits, one can define TE and TM polarized modes. The TM mode's electric field vector is perpendicular to the long axis of the slit, and the TE mode has its electric field vector parallel to the long axis. In standard waveguide models, the metal is usually assumed to be perfect, so that the continuity equation for the electric field implies that its parallel component must be zero at the metallic boundaries. In a slit geometry, this implies that TE-polarized light incident on such a slit will not be transmitted by the structure if the wavelength λ of the incident light is larger than twice the slit width w. This width is commonly referred to as the cutoff width. The TM-polarized mode, on the other hand, can propagate unimpeded through the slit, the effective mode index increasing steadily as the width is reduced.[7] For this reason one expects very narrow slits in metal films to act as perfect polarizers.[8]

[7] Astilean et al., 2000; Takakura, 2001.
[8] Fizeau, 1861.

While the perfect metal model is an excellent approximation for wavelengths in the mid to far infrared or microwave domains, the model is too naïve when the wavelength of the incident light is smaller, because of the dispersion in the permittivity of metals. As a consequence, in the visible part of the spectrum the TE mode cutoff width of real metals like silver and aluminum is slightly smaller than $\lambda/2$,[9] and the cutoff is more gradual. Although the TM mode propagates through the slit, it couples to surface plasmon modes on the front and back surfaces of the slit,[10] which act as a loss channel. Since these losses are heavily dependent on the slit width,[11] the transmitted intensity of the TM mode is more dependent on this width than the perfectly conducting waveguide model predicts.

[9] Schouten, Visser, Lenstra, and Blok, 2003; Schouten, Visser, Gbur, Lenstra, and Blok, 2004.
[10] Schouten, Kuzmin, et al., 2005.
[11] Lalanne, Hugonin, and Rodier, 2006; Baudrion et al., 2008; Kihm, Lee, Kim, Kang, and Park, 2008.

Here we demonstrate that, for thin metal films, such a nanoslit also acts as an optical retarder, and that the TE/TM transmission ratio is around unity well below the cutoff width, approaching zero only when the slit is extremely narrow. We have employed these properties to turn such a slit into a quarter-wave retarder.

2.2 *Description of experiment*

In the experiment, shown schematically in Fig. 2.1, we illuminate an array of ten 10 μm by 50–500 nm slits with a laser beam at λ = 830 nm, at normal incidence (see Fig. 2.2.) For all practical purposes, each slit's length can be considered infinite compared to its width and the laser wavelength. The slits are milled through a 200 nm thick gold film

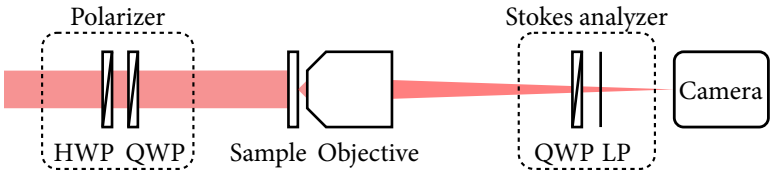


Figure 2.1: Sketch of the experimental setup. HWP: half-wave plate, QWP: quarter-wave plate, LP: linear polarizer. The sample (see Fig. 2.2) is illuminated on the gold side, using light with a controlled polarization. The transmitted light's polarization is analyzed for each pixel of a CCD camera. The Stokes analyzer consists of a quarter-wave plate and a linear polarizer, which can be rotated independently of each other under computer control to any desired orientation.

using a focused Ga^+ ion beam. The slits' widths increase stepwise from 50 nm, well below the cutoff width for TE-polarized light, to 500 nm, at which value the lowest TE mode can propagate through the slit. The film is deposited on a 0.5 mm thick Schott D263T borosilicate glass substrate, covered by a 10 nm titanium adhesion layer which damps surface plasmons, ensuring that their propagation length is negligibly short on the gold-glass interface. The laser beam width at the sample is approximately 4 mm so that, effectively, the structure is illuminated homogeneously with a flat wavefront. The light transmitted by the structure is imaged on a CCD camera (Apogee Alta U1) by means of a 0.65 NA microscope objective. The polarization of the light incident on the structure is controlled by a combination of half-wave and quarter-wave plates, enabling us to perform the experiment with a variety of input polarizations.

CCD: charge-coupled device
NA: numerical aperture

We analyze the polarization by measuring the Stokes parameters of the light transmitted through each slit, using a quarter-wave plate and a linear polarizer. We define the Stokes parameters according to the following standard convention: S_0 is the total intensity, S_1 is the intensity of the horizontal component (TE) minus the intensity of the vertical component (TM), S_2 is the intensity of the diagonal (45° clockwise) component minus the intensity of the anti-diagonal (45° counterclockwise) component, and S_3 is the intensity of the right-handed circular component minus the intensity of the left-handed circular component. Since the transmitted light is fully polarized, it is convenient to use the *normalized* Stokes parameters $s_1 = S_1/S_0$, $s_2 = S_2/S_0$, and $s_3 = S_3/S_0$, so that each ranges from −1 to +1.

$$
\begin{aligned}
S_0 &= I_{\text{total}} \\
S_1 &= I_H - I_V \\
S_2 &= I_D - I_A \\
S_3 &= I_R - I_L
\end{aligned}
$$

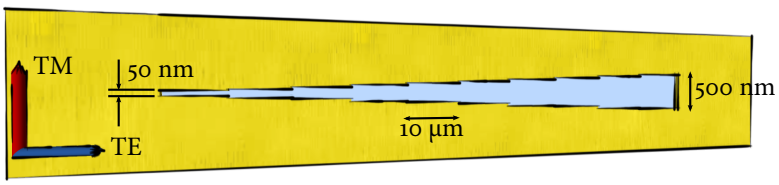


Figure 2.2: Sketch of the sample. It consists of a 200 nm gold film sputtered on top of a glass substrate. Note that the vertical scale is greatly exaggerated compared to the horizontal scale. Adapted from Kuzmin (2008, p. 78).

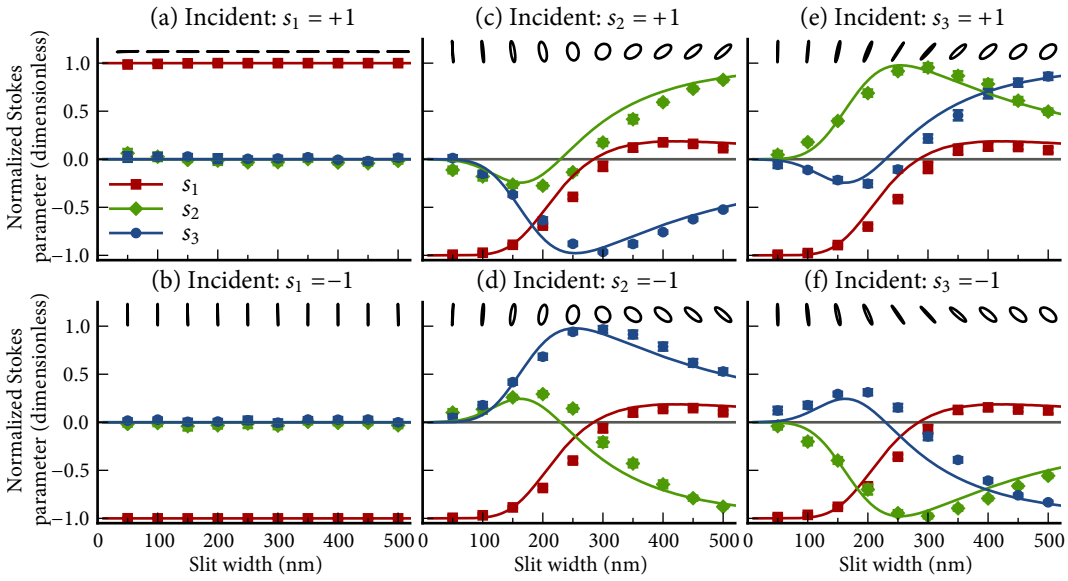


Figure 2.3: Normalized Stokes parameters of the light transmitted through the slit, for illumination with (a) horizontal linear polarization ($s_1 = +1$), (b) vertical linear polarization ($s_1 = -1$), (c) diagonal linear polarization ($s_2 = +1$), (d) antidiagonal linear polarization ($s_2 = -1$), (e) left-handed circular polarization ($s_3 = +1$), and (f) right-handed circular polarization ($s_3 = -1$). The polarization ellipses above each graph provide a quick visual indication of the polarization state of the transmitted light. The solid lines represent the results of our model, described later on in section 2.4, based on simple waveguide theory.

2.3 *Results and interpretation*

THE FULL STOKES ANALYSIS of the transmitted light, for each of the six basic Stokes input polarizations ($s_{1,2,3} = \pm 1$), is shown in Fig. 2.3. Figures 2.3a and 2.3b confirm that the TE and TM directions are the slit's eigenpolarizations. However, each has its own damping and propagation constant, as we will show. In the general case, a slit is therefore both dichroic and birefringent, both properties depending on the slit width w.

Figures 2.3c–f show the variation of the Stokes parameters of the transmitted light when the incident light is not polarized along one of the slits' eigenpolarizations. In all cases, s_1 is seen to go to -1 as the slit gets narrower, reflecting the fact that very narrow slits transmit only TM-polarized light.

Let us examine Figs. 2.3c–d more closely, where the incident wave is diagonally linearly polarized ($s_2 = \pm 1$). As the slit width w is reduced from 500 to 300 nm, the transmitted light gradually becomes more and more elliptically polarized, while the main axis of the polarization ellipse remains oriented along the polarization direction of the incident light; see the line of polarization ellipses in each frame. As w is reduced further to around 250 nm, the transmitted polarization assumes a more circular form. For narrower slits, the polarization ellipse orients itself essentially vertically, reflecting the fact that the polarization becomes more linear, ultimately

being purely TM-polarized at $w = 50$ nm. In Figs. 2.3e–f, with circular input polarization, a similar process happens as w is reduced, except that the transmitted polarization changes gradually from almost circular to linear, before becoming nearly TM-polarized at $w = 50$ nm.

We note that there is a point in Figs. 2.3e–f, around $w \approx 250$ nm, where circular polarization is transformed into linear polarization. This implies that the slit acts as a quarter-wave retarder, albeit with unequal losses for the fast and slow axes. Because of the inequality of these losses, the incident diagonal polarization in Figs. 2.3c–d is not transformed into a perfectly circular polarization. However, a properly oriented linear polarization incident on a $w \approx 250$ nm slit whose orientation compensates for the differential loss, *will* be transformed into circular polarization. Experiments on other slits have shown that the measured dichroism is highly dependent on the slit parameters, such as milling depth,[12] and the incident wavelength.[13] Realizing an ideal quarter-wave retarder therefore requires either careful design and manufacture of the slit, or serendipity.

[12] Bosman, 2011, p. 13.

[13] Bogers, 2012, p. 30.

As expected, the curves of s_2 and s_3 as a function of w flip their sign when the sign of the incident Stokes parameter is flipped. When the incident light's s_2 and s_3 are exchanged, on the other hand, so are s_2 and s_3 in the transmitted light. The curve of s_1 remains the same for all non-s_1 incident polarizations. The results shown in Fig. 2.3 can all be represented in one figure by plotting the measured Stokes parameters on the Poincaré sphere. Reducing the slit width then traces out a path of the transmitted polarization state over the Poincaré sphere's surface, as shown in Fig. 2.4.

In order to analyze our experimental data, we write the incident field as a Jones vector, preceded by an arbitrary complex amplitude such that the TE component is real and positive:

$$\mathbf{E}_{\text{in}} = \tilde{A} \begin{bmatrix} E_{\text{TE}} \\ E_{\text{TM}}\, e^{i\psi} \end{bmatrix}, \qquad \text{with } E_{\text{TE}}, E_{\text{TM}} \geq 0. \tag{2.1}$$

We express the transmission properties of the slit as a Jones matrix. Its off-diagonal elements are zero, because the TE and TM directions are the slit's eigenpolarizations, and the diagonal elements represent the complex amplitude transmission. The output field is then the Jones vector:

$$\mathbf{E}_{\text{out}} = \begin{bmatrix} t_{\text{TE}} & 0 \\ 0 & t_{\text{TM}} \end{bmatrix} \mathbf{E}_{\text{in}}. \tag{2.2}$$

First, it is instructive to calculate the transmission T_{TE} and T_{TM} in order to get an idea of the slit's dichroism. Here, we define the transmission $T = |t|^2$

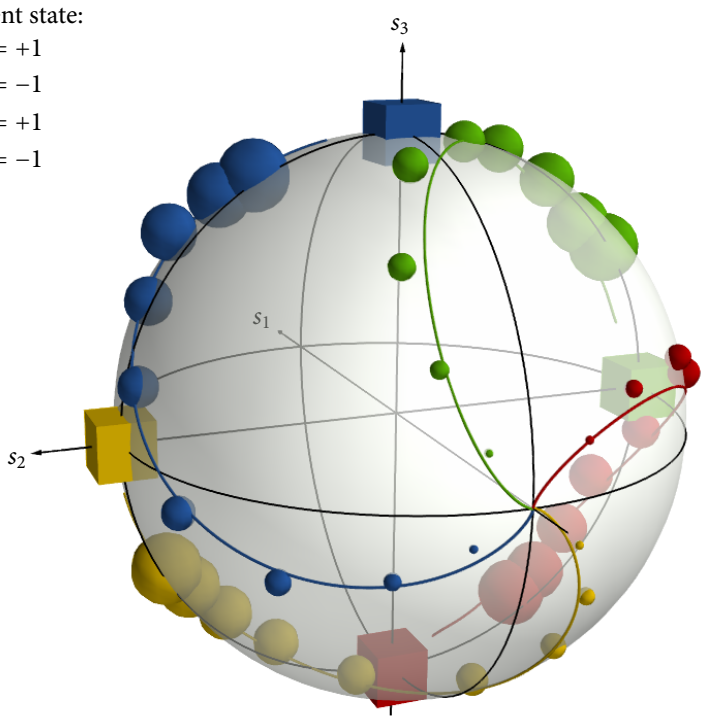


Figure 2.4: Path of the transmitted polarization state over the Poincaré sphere as the slit width decreases. The incident polarization state starts at one of the poles or equatorial points, represented by the boxlike markers. The spherical markers, with size proportional to the slit width, mark the transmitted polarization state as it travels over the sphere's surface. The solid lines are the predictions of our model.

as the ratio of power emerging from a slit to power incident on the area of the slit. It can be calculated from the unnormalized Stokes parameter S_1 for incident light with $s_1 = \pm 1$. T_{TE} and T_{TM} are plotted in Fig. 2.5, normalized so that $T_{\mathrm{TE}} = 1$ at $w = 500$ nm. As the slit width w is decreased, we see that the TE and TM transmission also decrease until $w \approx 350$ nm. When w is further reduced, the TM transmission goes through a minimum at $w \approx 150$ nm, where the light-surface plasmon coupling is maximum.[14] It increases again when the slit width gets even smaller, whereas the TE transmission goes through a gradual cutoff, becoming negligible only for the narrowest slits. Apparently, a narrow slit in a *thin* metal film is not such a good polarizer as often assumed.

[14] Lalanne et al., 2006.

In order to calculate the phase lag $\Delta\varphi$ between the TE and TM-polarized components of the transmitted field, we write the normalized Stokes

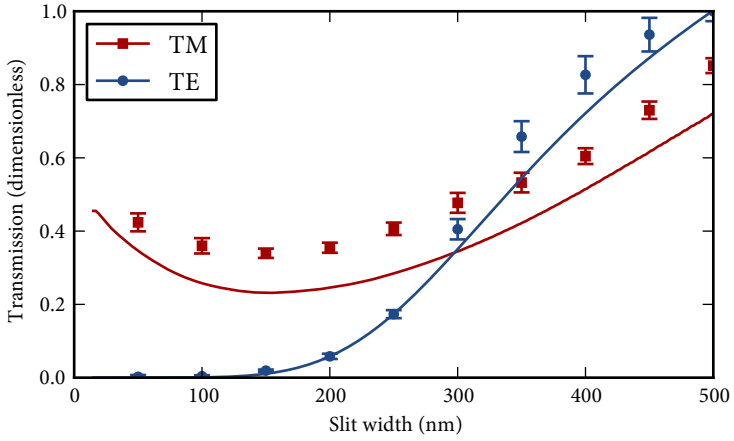


Figure 2.5: Dichroism of a subwavelength slit. The points show the measured transmission for TM and TE-polarized incident light as a function of the slit width w, normalized to the TE transmission at $w = 500$ nm. The solid lines show our model's result for the slit transmission according to (2.7) and (2.8).

parameters in terms of (2.2):

$$s_1 = -\frac{T_R E_{\mathrm{TM}}^2 - E_{\mathrm{TE}}^2}{T_R E_{\mathrm{TM}}^2 + E_{\mathrm{TE}}^2}, \tag{2.3}$$

$$s_2 = \frac{2\sqrt{T_R} E_{\mathrm{TM}} E_{\mathrm{TE}}}{T_R E_{\mathrm{TM}}^2 + E_{\mathrm{TE}}^2} \cos(\Delta\varphi - \psi), \tag{2.4}$$

$$s_3 = -\frac{2\sqrt{T_R} E_{\mathrm{TM}} E_{\mathrm{TE}}}{T_R E_{\mathrm{TM}}^2 + E_{\mathrm{TE}}^2} \sin(\Delta\varphi - \psi), \tag{2.5}$$

where $T_R = |t_{\mathrm{TM}}/t_{\mathrm{TE}}|^2$ is shorthand for the transmission ratio, E_{TM} and E_{TE} are the transmitted fields, and ψ is the TM–TE phase lag; see (2.1). We calculate $\Delta\varphi$ from our measured Stokes parameters using (2.3), (2.4), and (2.5). We see in Fig. 2.6 that $\Delta\varphi$ decreases almost linearly with increasing slit width. It passes through a value of $\pi/2$ at $w \approx 250$ nm. Although the retardation equals $\lambda/4$, the 250 nm slit does not act as an ideal quarter-wave retarder because the amplitudes of the TE and TM-polarized components of the transmitted light are not equal, as noted earlier.

Figure 2.5 illustrates the slit's dichroism and Fig. 2.6 its birefringence. The effect that we observe in Fig. 2.3 as the slit width is decreased from 500 to 300 nm can be explained in terms of increasing birefringence and small dichroism in that range. Below 300 nm, dichroism becomes more important, and consequently, the main axis of the polarization ellipse rotates. The dichroism observed here was also suggested by calculations by Nugrowati, Pereira, and van de Nes,[15] where ultrashort TE pulses were shown to experience lower propagation speeds than TM pulses through a slit in an aluminum layer.

If the slit width is further decreased past the surface plasmon-induced

[15] Nugrowati et al., 2008.

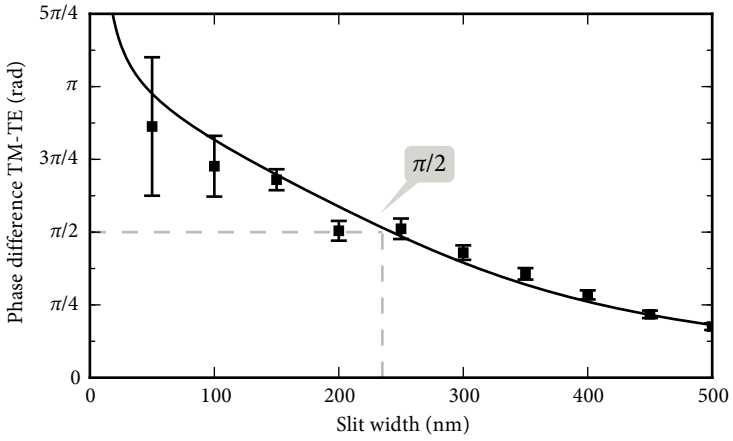


Figure 2.6: Birefringence of a subwavelength slit. The points represent the measured phase difference between the TM and TE modes as a function of the slit width. They are obtained from a fit of the various Stokes parameters of Fig. 2.3. The solid line shows the calculated phase difference according to (2.9). At a certain slit width, indicated by the arrow, the phase difference reaches $\pi/2$ and the slit acts as a quarter-wave retarder.

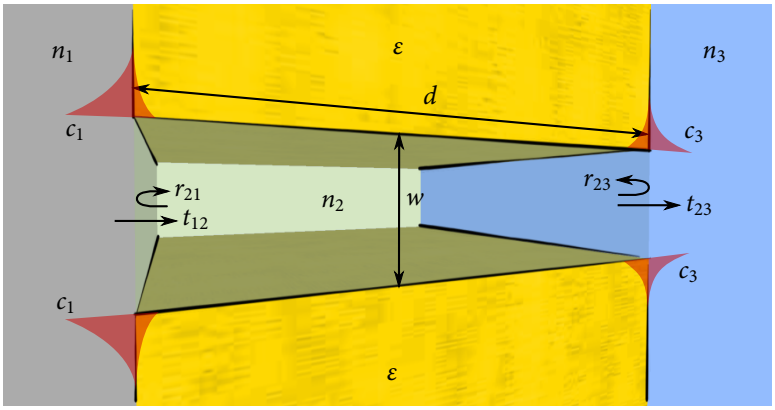


Figure 2.7: Cross-section of our model slit. The relevant physical quantities are illustrated. The indices of refraction $n_{1,2,3}$ are depicted using differently colored materials, although they could well be the same material in an experiment. The evanescent tails sketched in red represent surface plasmons. The light is transmitted through the slit from left to right.

minimum at $w \approx 150$ nm, the dichroic effect becomes even larger. The TE-polarized component of the transmitted light becomes weaker and weaker, while the TM component grows, causing the polarization ellipse to collapse to a vertical line. We see that the waveguide's TE cutoff does not resemble a sharp cutoff at $w = \lambda/2$ at all, but rather a gradual one.

2.4 *Waveguide model*

We have made our computer code for this model available (Chimento, 2013b).

WE WILL NOW PROCEED to explain these experimental results by modeling the slit as a simple lossy waveguide. Our metallic slit forms a rectangular waveguide with one dimension of the rectangle much larger than the other. For that reason we can effectively describe each slit as a step-index planar waveguide, with its walls made of a metal with relative per-

mittivity ε. Inside the waveguide, the solutions to Maxwell's equations separate into TE and TM modes, each with a complex propagation constant β. Although the equations are in closed form, we must calculate the propagation constants for each mode, β_{TE} and β_{TM}, numerically.[16]

[16] Snyder and Love, 1983.

For the TM and TE modes, we calculate complex reflection and transmission coefficients r_{21}, t_{12}, r_{23}, and t_{23} (see Fig. 2.7) using the Fresnel equations at normal incidence, substituting the effective mode index for the index of medium 2. The effective mode index is calculated by dividing the propagation constant by k_0. As shown in Fig. 2.7, the index 1 indicates the medium from which the light is incident (air), 2 the waveguide, and 3 the medium into which the transmitted light emerges (glass in our experiment). This simplification avoids calculating overlap integrals between the guided mode and the modes outside the waveguide, but still describes the observed phenomena quite well. We can then treat the waveguide as a Fabry-Pérot interferometer and calculate each mode's complex transmission through a waveguide of length d,

$$t_{123} = \frac{t_{12}t_{23}e^{i\beta d}}{1 - r_{21}r_{23}e^{2i\beta d}}, \tag{2.6}$$

which gives for the transmission

$$T_{\mathrm{TE}} = \frac{n_3}{n_1}|t_{123}^{\mathrm{TE}}|^2, \tag{2.7}$$

$$T_{\mathrm{TM}} = \frac{n_3}{n_1}|t_{123}^{\mathrm{TM}}|^2 - 2|c_1|^2 - 2|c_3|^2. \tag{2.8}$$

Here, c_1 and c_3 are the coupling constants of the slit system to a surface plasmon mode traveling in one direction away from the slit on the interface with medium 1 or 3, respectively. Numerical values for these parameters can be calculated[17] using Eq. (20) of Lalanne et al.,[18] which gives an approximate analytical model for the coupling of a slit mode to a surface plasmon mode. As an illustration of the important role these surface plasmon coupling constants play in the phenomenon described here, the TM transmission modelled with and without coupling to surface plasmons is shown in Fig. 2.8. The TE mode does not couple to surface plasmons.

[17] Calculating c_1 and c_3 requires evaluating an integral with poles close to the real axis. Common adaptive quadrature algorithms for numerical integration cannot handle it, yielding a garbage answer without obviously failing. Gaussian quadrature works for numerically evaluating the integral.

[18] Lalanne et al., 2006.

[19] Lalanne et al., 2006.

It is interesting to note in Fig. 2.8 that the surface plasmon coupling coefficients on both sides exhibit a maximum at $nw/\lambda \approx 0.23$ and a minimum at $nw/\lambda \approx 1$, as predicted by Lalanne et al.,[19] where n is the index of refraction of the medium outside the slit on each respective side. These two curves added together yield a maximum in the surface plasmon excitation, and therefore a dip in the TM transmission, at around $w \approx 150$ nm. Even though this dip is not at $nw/\lambda \approx 0.23$ as Lalanne et al. predict, it is caused by two plasmon excitation maxima that do follow the prediction.

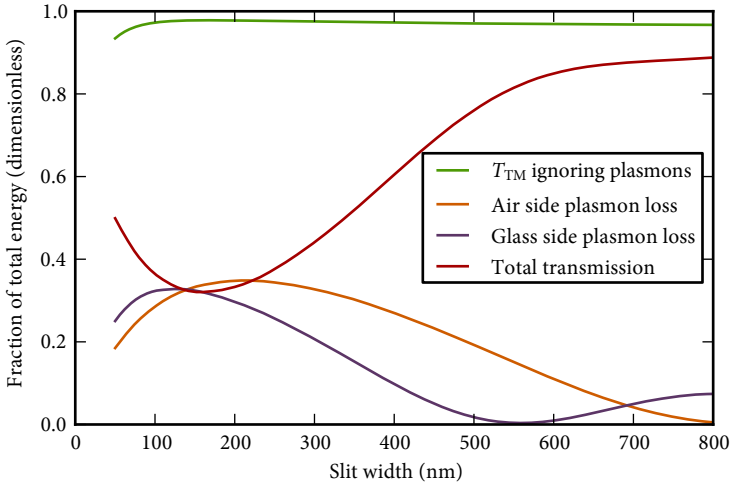


Figure 2.8: Calculated effect of surface plasmons on the transmission of TM-polarized light as a function of the slit width w. The green line shows the calculated TM transmission neglecting coupling to surface plasmons, based on waveguide theory alone, i.e. $(n_3/n_1)|t_{123}^{TM}|^2$. The orange line shows the total fraction of energy $2|c_1|^2$ converted to surface plasmons on the illuminated (air) side of the sample according to Lalanne et al. (2006). Likewise, the purple line shows the fraction $2|c_3|^2$ converted to surface plasmons on the unilluminated (glass) side. Finally, the red line shows the total TM transmission according to (2.8). In these calculations, we disregard the numerical aperture of the imaging system.

In our model we ignore the thin titanium adhesion layer present between the gold and the glass. According to the model, the $|c_3|$ coefficient for a thick titanium layer would be slightly higher than that of the gold layer. However, we expect that the layer is too thin to have any effect on the coupling between the slit TM mode and surface plasmons. It does not prevent the light from scattering into the surface plasmon mode, but only ensures that the surface plasmon mode is very lossy.

Our model exhibits good agreement with the measurements, despite the fact that it does not contain any fitting parameters. The slit's gradual TE cutoff is predicted well, and can be ascribed to gold not being a perfect conductor at this wavelength, and to the considerable dispersion of the reflection coefficients r_{12} and r_{23} around cutoff. The model also predicts a plasmon-related TM transmission dip at the right slit width. In Fig. 2.5, we compare these calculated values to our measurements. In our calculations, we took the finite NA and its influence on the TM and TE transmission into account, which is explained in section Appendix 2.A.

The complex transmission also gives us the relative phase delay between the TM and TE modes:

$$\Delta\varphi = \arg t_{123}^{TM} - \arg t_{123}^{TE} \pmod{2\pi}. \tag{2.9}$$

This phase difference is plotted in Fig. 2.6 and compared to the values calculated from our measurements using (2.3), (2.4), and (2.5). The values predicted by our simple model for the phase delay exhibit excellent agreement with the measurements.

The model presented here suggests exploring the parameter space in order to design slits that act as non-dichroic quarter-wave retarders. The requirements are that the TM and TE transmission are equal taking into account the TM loss to surface plasmons, and that the phase difference is $\pi/2$. All these requirements are influenced by the metal permittivity $\varepsilon(\lambda)$, the slit width w, and the film thickness d.

Our experimental results contradict a recently published proposal for a quarter-wave retarder using perpendicular metallic nanoslits,[20] in which the width of the slits is varied purely to control the TM transmission. Varying the width of the slit also changes the TE transmission of the incident light and the phase difference between the TM and TE components.

[20] Khoo, Li, and Crozier, 2011.

2.5 *Summary*

We have studied the transmission properties of a subwavelength slit milled in a 200 nm thick gold-metal film as a function of the slit width (50–500 nm), and of the polarization of the incident radiation (at $\lambda = 830$ nm). As the slit width is decreased, the transmission of the TE mode diminishes quite gradually until it becomes very small at a slit width of about $\lambda/8$, reminiscent of the phenomenon of waveguide cutoff. In contrast, the transmission of the TM mode does not vanish. Instead, it exhibits a minimum associated with the efficient excitation of surface plasmons.

Moreover, we have studied the birefringence of this subwavelength slit and found that the phase lag between the TM mode and TE mode passes through a value of $\pi/2$, so that a properly dimensioned slit can act as a quarter-wave retarder. We have successfully explained our experimental results with a simple waveguide model.

Appendix 2.A Reciprocity of the slit transmission

This section is an appendix that did not appear in the published paper.

WE STATED EARLIER IN THIS CHAPTER that we corrected our model for the finite NA of the detector in our experiment. The necessity of this correction was brought to our attention by an apparent violation of reciprocity in the experiment. In the experiments described in the foregoing sections, we illuminated the sample on the gold side (hereafter the 'forward' configuration), but when we turned the sample around and performed the experiment again while illuminating it from the glass side (the 'reverse' configuration), the results were different!

This is, of course, not really a violation of reciprocity, but it is caused by the detector's NA. The exit aperture of the slit is subwavelength, so it radiates in all directions, but not uniformly. The slit's scattering profile depends on the shape of the mode inside the slit, and also on the medium that the slit scatters into. Therefore, not all the radiation that actually exits the slit is emitted into the cone of angles that the detector can collect. Which fraction is collected by the detector depends on the circumstances, meaning that the two configurations cannot be compared directly without correcting for this effect.

We assume that the detector is situated in air, with an index of refraction n_4. In the forward configuration, $n_1 = n_4$ and n_3 is the index of the glass substrate. Conversely, in the reverse configuration, n_1 is the index of the glass substrate and $n_3 = n_4$.

THE SLIT SUPPORTS ONE TM AND ONE TE MODE. We treat the slit as a parallel-plate waveguide with metal walls in which we assume that higher-order modes do not propagate. The metal wall boundary is at $x = \pm w/2$. We call the complex amplitudes of the waveguide modes inside the slit $\tilde{\mathbf{E}}_{\mathrm{TM}}(x, z)$ and $\tilde{\mathbf{E}}_{\mathrm{TE}}(x, z)$. They depend on the permittivity ε, the index of the slit material n_2, and the slit width w.

To find the angular scattering profile of the modes, we take the Fourier transform of the mode profile at the exit aperture of the slit: $\tilde{\mathbf{E}}_{\mathrm{TM}}(x, d)$ and $\tilde{\mathbf{E}}_{\mathrm{TE}}(x, d)$. This gives us the scattered electric field amplitude $\tilde{\mathbf{F}}$ as a function of transverse wavenumber k_x. This is appropriate if the collection objective is in the far field of the slit. We estimate the Fresnel number N_F using typical values for our experiment,

$$N_F = \frac{a^2}{L\lambda} = \frac{w^2}{4L\lambda} \approx \frac{(5 \times 10^{-7})^2}{4 \cdot 1 \times 10^{-3} \cdot 8 \times 10^{-7}} \approx \frac{25 \times 10^{-14}}{32 \times 10^{-10}} \ll 1, \qquad (2.10)$$

which justifies the assumption of Fraunhofer diffraction at 1 mm distance from the slit. L is in this case the working distance of the objective.

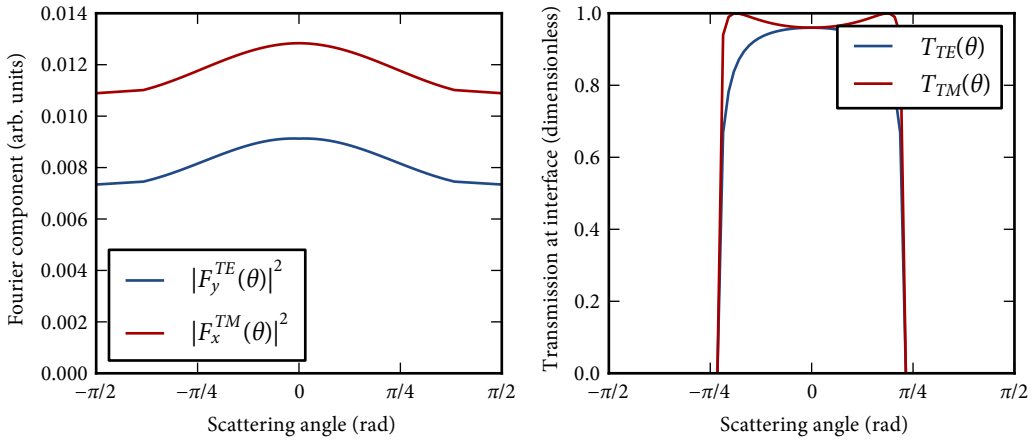


Figure 2.9: (a) Scattering profile of a 250 nm wide slit, as a function of angle. (b) Transmission of scattered light at the n_3–n_4 interface as a function of scattering angle. Note the Brewster angle at the point where the TM transmission reaches unity. Other parameters: $\lambda = 800$ nm, $n_2 = 1.0$, $n_3 = 1.5$, $n_4 = 1.0$.

The angle θ of the corresponding plane wave component is equal to

$$\theta = \arcsin(k_x/n_3 k_0), \tag{2.11}$$

where $k_0 = 2\pi/\lambda$ is the wavenumber in free space. We plot the angular scattering profiles for a 250 nm wide slit in Fig. 2.9a, calculated numerically by fast Fourier transform.

In the forward configuration (light incident on the air side) described above, we have to take into account the Fresnel losses at the n_3–n_4 (glass-air) interface. Part of the scattered energy never leaves the glass substrate, due to total internal reflection. The TE and TM components are also transmitted differently, since there is a Brewster angle for TM.

For the transmission $T = 1 - R$, we write:

$$T_{\mathrm{TE}}(\theta) = 1 - \left(\frac{n_3 \cos\theta - n_4 \sqrt{1 - (\frac{n_3}{n_4}\sin\theta)^2}}{n_3 \cos\theta + n_4 \sqrt{1 - (\frac{n_3}{n_4}\sin\theta)^2}} \right)^2 \tag{2.12}$$

$$T_{\mathrm{TM}}(\theta) = 1 - \left(\frac{n_3 \sqrt{1 - (\frac{n_3}{n_4}\sin\theta)^2} - n_4 \cos\theta}{n_3 \sqrt{1 - (\frac{n_3}{n_4}\sin\theta)^2} + n_4 \cos\theta} \right)^2 \tag{2.13}$$

We plot these transmission profiles in Fig. 2.9b. They are independent of the slit width, or indeed any of the slit parameters. Also note that for $n_3 = n_4$, $T = 1$, as it should be since there is no interface in that case.

The finite NA of the detector means that not all of the scattered light is collected. Light outside a maximum acceptance angle $\theta_{\max}$ misses the detector. Due to Snell's law, $NA = n_3 \sin\theta_{\max}$, and therefore

$$\theta_{\max} = \arcsin(NA/n_3), \tag{2.14}$$

no matter what medium n_4 the detector is actually in.

The detector signal, then, must be corrected by a factor

$$C_i = \frac{\displaystyle\int_{-\theta_{\max}}^{\theta_{\max}} T_i(\theta) \left|\tilde{\mathbf{F}}_i(n_3 k_0 \sin\theta)\right|^2 d\theta}{\displaystyle\int_{-\pi/2}^{\pi/2} \left|\tilde{\mathbf{F}}_i(n_3 k_0 \sin\theta)\right|^2 d\theta}, \tag{2.15}$$

where $i \in \{\text{TE}, \text{TM}\}$, which is the ratio of energy entering the detector to the total energy scattered from the slit. This factor depends on λ, ε, n_2, n_3, n_4, w, and NA. (It does not depend on d or n_1; the factor containing d can be taken outside the integrals and divided away, and n_1 only influences the coupling to surface plasmon modes on the front interface.)

In the experiment described in this chapter, we take λ, ε, n_2, and n_4 as design parameters, so we will not examine their influence on the correction factor. The index n_3 can take one of two values: for reverse illumination, $n_3 = n_4$, whereas for forward illumination $n_3 \neq n_4$. Since we are only interested in the ratio of TM to TE transmission, we will examine the ratio of the two correction factors $C_{\text{TM}}/C_{\text{TE}}$ in four situations: forward and reverse illumination, and for low and high NA.

We examine the ratio because the values of C_{TM} and C_{TE} individually are not particularly surprising; a high-NA detector will collect more of the light in both forward and reverse illumination, obviously. Reverse illumination also causes more light to enter the detector, no matter the NA because there is no total internal reflection loss at the n_3–n_4 interface, so $T_i(\theta) = 1$ in (2.15).

In Fig. 2.10a we plot the ratio $C_{\text{TM}}/C_{\text{TE}}$ for reverse illumination, calculated for two different numerical apertures. We see that the numerical aperture does not influence the shape of the curve very much.

Forward illumination, for which the ratio of the correction factors is plotted in Figs. 2.10b, is quite different. If the detector has low NA, it behaves much the same as in reverse illumination, with slightly more difference between TM and TE. In the high-NA case, on the other hand, the detector picks up more TM light than TE light for almost every slit width: there is about a 6% difference. A look at Fig. 2.9 suggests that this happens when the NA is large enough that $\theta_{\max}$ is large enough that it includes the Brewster angle at the n_3–n_4 interface, at which all the TM light is transmitted and not all the TE light.

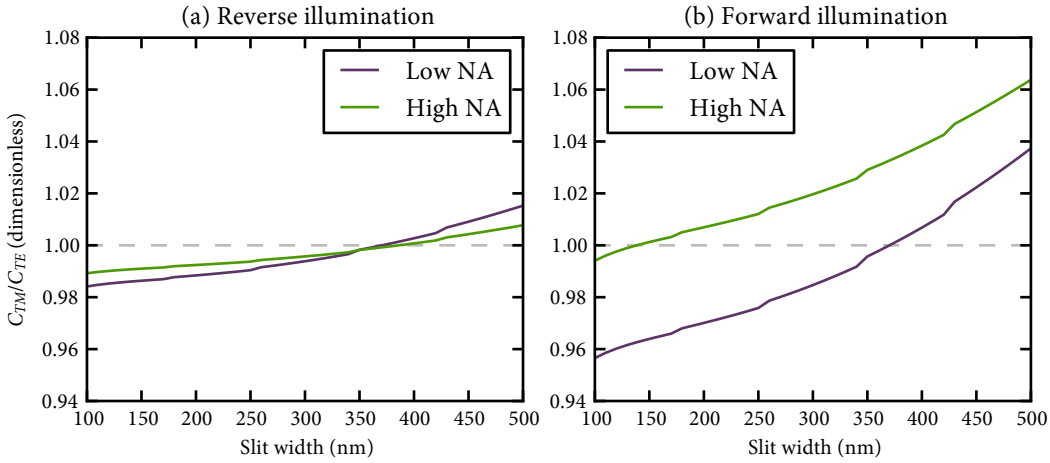


Figure 2.10: Ratio of the two NA correction factors $C_{\mathrm{TM}}/C_{\mathrm{TE}}$: less than 1 means that the detector picks up more TE light, more than 1 means the detector picks up more TM light. It is plotted for two different NA values: low ($NA = 0.2$, purple), and high ($NA = 0.8$, green). (a) Reverse illumination ($n_1 = 1.5$, $n_2 = n_3 = n_4 = 1.0$); (b) forward illumination ($n_1 = n_2 = n_4 = 1.0$, $n_3 = 1.5$). As usual, we take $\lambda = 800$ nm.

3
Spin-to-orbital angular momentum conversion in a subwavelength slit

We demonstrate partial conversion of circularly polarized light into orbital angular momentum-carrying vortex light with opposite-handed circular polarization. This conversion is accomplished in a novel manner using the birefringent properties of a circular subwavelength slit in a thin metal film. Our technique can be applied over a very wide range of frequencies and even allows the creation of anisotropic vortices when using a slit without circular symmetry.

3.1 Introduction

This chapter was previously published as: Chimento, Alkemade, 't Hooft, and Eliel (2012).

The curious phenomenon of optical vortices arising from axial symmetry in birefringent materials has been studied in uniaxial crystals of variable length,[1] birefringent plates with a spatially varying optical axis and half-wave retardation ("q-plates"),[2] and in annular concentric apertures which resonantly excite surface plasmons.[3] This interaction between spin and orbital angular momentum of light by way of a Berry-Pancharatnam phase has also been studied in space-variant gratings,[4] plasmonic nanostructures in the context of selection rules,[5] and also completely outside the domain of optics, in electron beams.[6]

We present here a novel method of accomplishing this conversion using a subwavelength slit in a metal film acting as a quarter-wave plate, described in chapter 2. We show how this method relaxes the axial symmetry requirement, allowing greater versatility in the form of the vortex created.

In optics, a spin angular momentum of $\pm\hbar$ is associated with a circularly polarized photon. Orbital angular momentum is often associated with an optical vortex beam, where the phase increases azimuthally around the optical axis. These beams have a topological charge Q, equal

[1] Ciattoni et al., 2003; Brasselet, Izdebskaya, et al., 2009.

[2] Marrucci et al., 2006; Karimi, Piccirillo, Marrucci, and Santamato, 2009; Brasselet and Loussert, 2011.

[3] Lombard et al., 2010.

[4] Bomzon et al., 2001.

[5] Gorodetski, Shitrit, Bretner, Kleiner, and Hasman, 2009.

[6] Karimi, Marrucci, et al., 2012.

to the number of full cycles the phase makes in one trip around the optical axis. The expectation value of the orbital angular momentum per photon is $Q\hbar$.[7] The difference between the two forms of angular momentum is beautifully apparent in the interaction of a beam with small particles: interaction with the spin angular momentum in the absence of absorption requires particles that are birefringent; they will start to rotate about their own axis, whereas interaction with a beam carrying orbital angular momentum causes particles, whether birefringent or not, to rotate about the beam's optical axis.[8]

[7] Allen, Beijersbergen, et al., 1992.

[8] O'Neil, MacVicar, Allen, and Padgett, 2002.

Chapter 2 describes how a subwavelength slit in a metal film can act as an optical retarder. A slit which is subwavelength in one direction, and extended in the other, has two eigenpolarizations: parallel and perpendicular to the slit. By careful design of the slit's width and depth, it is possible to construct a slit that behaves like a quarter-wave retarder for incident light of a certain wavelength, with its fast axis (i.e. axis with the lowest index of refraction) parallel to the orientation of the slit. One can achieve similar results using subwavelength structures with different resonances for orthogonal polarization components.[9] Illuminating the straight slit with circularly polarized light results in linearly polarized light emerging from the other side. The associated change in angular momentum means that a torque is exerted on the sample.[10]

[9] Roberts and Lin, 2012; Genevet et al., 2012.

[10] Beth, 1936.

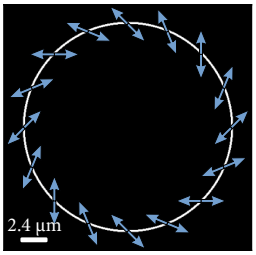


Figure 3.1: Diagram showing the expected local polarization state of light transmitted through the ring slit. The transmitted intensity is constant everywhere on the slit.

When the slit is circular, the fast and slow axes' orientations vary along the slit so that it acts as a space-variant quarter-wave plate. In this circularly symmetric configuration, photonic spin angular momentum cannot transfer to the sample, and must be converted to photonic orbital angular momentum in order to conserve the total angular momentum. This intuitive picture is confirmed by taking the expectation value of the spin and orbital angular momenta per photon,[11] respectively denoted S and L, averaged over the whole beam in the input and output states. Whereas the input state has $S = \hbar, L = 0$, the output state (shown in Fig. 3.1) has $S = 0, L = \hbar$. The total angular momentum per photon, $J = S + L$, is indeed conserved.

[11] Berry, Jeffrey, and Mansuripur, 2005.

3.2 *Near-field experiment*

To confirm this by experiment, we took a glass substrate of 0.5 mm thickness. On it we deposited a titanium adhesion layer of 10 nm thickness, and on that a gold film of 200 nm thickness. We milled a circular slit, 20 µm in diameter and (180 ± 10) nm wide, through the gold film

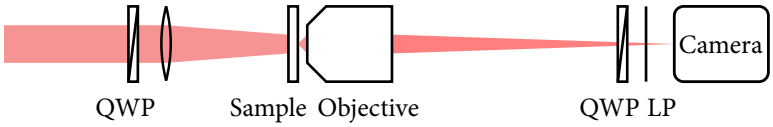


Figure 3.3: Sketch of the experimental setup used to image the ring slit. QWP: quarter wave plate; LP: linear polarizer. The quarter-wave plate and linear polarizer on the right-hand side of the figure measure the local polarization state of the light.

using a focused Ga^+ ion beam. Fig. 3.2 shows a sketch of the structure.

We conducted the experiment using a diode laser with a wavelength of 830 nm. We used a quarter-wave plate to give the beam from this laser a circular polarization state, $\hat{\boldsymbol{\sigma}}_+$. (We define the circular polarization basis unit vectors $\hat{\boldsymbol{\sigma}}_\pm = (\hat{\mathbf{x}} \pm i\hat{\mathbf{y}})/\sqrt{2}$.) We then focused the beam weakly onto the glass side of the sample. The beam diameter at the waist was 90 µm, much larger than the nanostructure diameter of 20 µm, so that, effectively, the structure was illuminated with a plane wave. We used a microscope objective (NA 0.4) to image the slit onto a CCD camera (Apogee Alta U1).

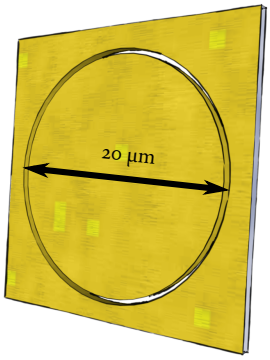


Figure 3.2: A sketch of the nanostructure milled into the sample.

NA: numerical aperture

CCD: charge-coupled device

We measured the polarization of the transmitted light as a function of the transverse position within the image. To determine this polarization, we used a fixed linear polarizer and a computer-controlled rotating quarter-wave plate, as shown in Fig. 3.3, from which we extracted the Stokes parameters according to the method described in Schaefer, Collett, Smyth, Barrett, and Fraher (2007) as a function of position. Fig. 3.4 shows the results of this experiment. We observe small variations in the transmitted intensity along the ring, which are probably caused by small variations in the slit width. The polarization state of the light emerging from the structure, however, shows excellent agreement with the result of our calculations, as shown in Fig. 3.5.

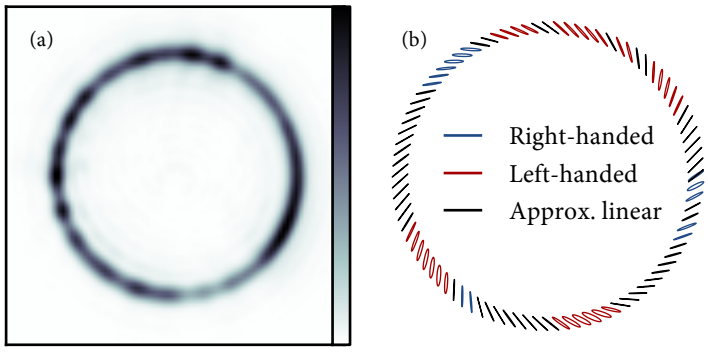


Figure 3.4: (a) Measured intensity transmitted through the ring slit. (b) Local polarization ellipses of the light transmitted through the ring slit. Blue ellipses indicate right-handed elliptical polarization, red ones indicate left-handed elliptical polarization, and black lines indicate polarization states with ellipticity less than 10%.

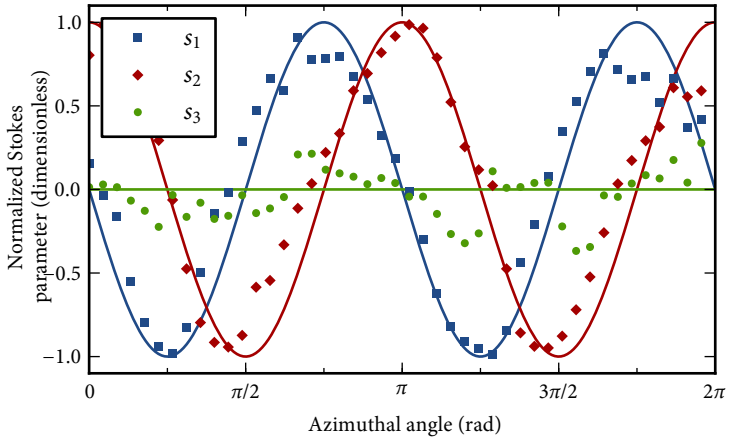


Figure 3.5: Measured normalized Stokes parameters $s_1 = S_1/S_0$, $s_2 = S_2/S_0$, $s_3 = S_3/S_0$ of the light transmitted through the ring slit as a function of azimuthal angle. This shows the same information as Fig. 3.4, but here it is easier to compare it to the expected results (solid lines), with which we observe quite good agreement. An angle of 0° corresponds to 3 o'clock in Fig. 3.4, and increases counterclockwise.

3.3 *Analytical model*

THE POLARIZATION measured in Fig. 3.4 suggests that the light emerging from the nanostructure is a superposition of radial and azimuthal polarization. Beams with such types of polarization, usually called vector beams, were first described as waveguide modes[12] with a dark spot in the center due to a polarization singularity. At first glance, one might expect our metallic nanostructure to produce a vector beam, and thus have a dark spot in the center of the far field. However, calculating the far field by numerical Fourier transform shows that there is no dark spot in the center; in fact, the local polarization state on the optical axis in the far field is purely $\hat{\boldsymbol{\sigma}}_+$, the same as the input polarization state.

[12] Marcatili and Schmeltzer, 1964.

In order to explore this further, we derived an analytical expression for the far field by Fourier-transforming the field shown in Fig. 3.1 and linearizing over the slit width ΔR,

$$\mathbf{E}_0^{\mathrm{FF}} \approx \frac{1+i}{\sqrt{2}} \pi R_0 \Delta R \left(J_0(R_0 k_\perp) \hat{\boldsymbol{\sigma}}_+ - i e^{2i\theta} J_2(R_0 k_\perp) \hat{\boldsymbol{\sigma}}_- \right), \tag{3.1}$$

where R_0 is the radius of the ring, $k_\perp$ the transverse component of the wave vector, and J_n denotes the Bessel function of the first kind of order n. This expression is valid for small ΔR in the paraxial approximation. These fields are visualized in Figs. 3.7(a) and (d). Note that the characteristic length scale in the far field is given by the radius R_0 of the circular structure — that is, the diffraction pattern does not arise from an aperture cutoff, but from the interference between opposite points on the circular slit.

This expression indicates that half of the transmitted beam energy has

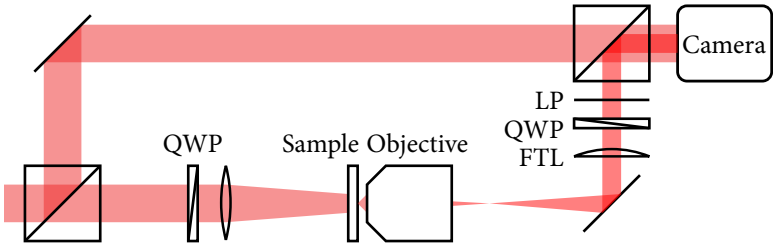


Figure 3.6: Sketch of the experimental setup used to measure the polarization and phase of the far field of the slit. QWP: quarter wave plate; FTL: Fourier-transforming ($2f$) lens; LP: linear polarizer. The objective's focus is now not on the camera but in the focus of the FTL. In this case, the quarter-wave plate and linear polarizer are simply used to view the $\hat{\boldsymbol{\sigma}}_+$ and $\hat{\boldsymbol{\sigma}}_-$ components separately. This configuration also includes a Mach-Zehnder interferometer which measures the phase of each polarization component. When not measuring the phase, we simply block the reference beam.

been converted from the $\hat{\boldsymbol{\sigma}}_+$ to the $\hat{\boldsymbol{\sigma}}_-$ state, while acquiring a topological charge of +2. (The integral of any $J_n(x)$ to infinite x is equal to 1 if $n \geq 0$.) This acquisition of topological charge by the opposite-handed component of the emerging beam can be seen as the result of spin-to-orbital angular momentum conversion, but it is equally instructive to consider it a Berry-Pancharatnam phase, the result of traveling from the north pole ($\hat{\boldsymbol{\sigma}}_+$) of the Poincaré sphere to the south pole ($\hat{\boldsymbol{\sigma}}_-$) through all possible points on the equator, twice.

We confirm this by calculating the expectation values of the spin and orbital angular momenta per photon for both polarization components separately. For the $\hat{\boldsymbol{\sigma}}_+$ component we have $S = \hbar, L = 0$, which is the same as the input state. For the $\hat{\boldsymbol{\sigma}}_-$ component, we have $S = -\hbar, L = 2\hbar$.

3.4 *Far-field experiment*

We performed further experiments to explore this, using a $2f$ system to examine the far field; see Fig. 3.6. We used a quarter-wave plate and a linear polarizer to measure the intensity distribution of the $\hat{\boldsymbol{\sigma}}_+$ and $\hat{\boldsymbol{\sigma}}_-$ components of the far field separately. We also used a misaligned Mach-Zehnder interferometer to visualize the phase of the light transmitted through the slit. The interference pattern consists of parallel interference fringes, which fork according to the topological charge carried by the beam.[13] Figure 3.7 shows the results of our measurements compared to the calculation of (3.1). The interferograms in Figs. 3.7(c) and (f)[14] show that the $\hat{\boldsymbol{\sigma}}_-$ component does indeed have a topological charge of +2, whereas the $\hat{\boldsymbol{\sigma}}_+$ component carries no topological charge.

[13] Basistiy, Soskin, and Vasnetsov, 1995.

[14] In (f), the interference fringe minima are marked in red, using the technique described in Cai, Liu, and Yang (2003).

3.5 *Discussion*

We also consider what happens when the amplitudes of the transmitted polarization components are unequal, or when the retardation is

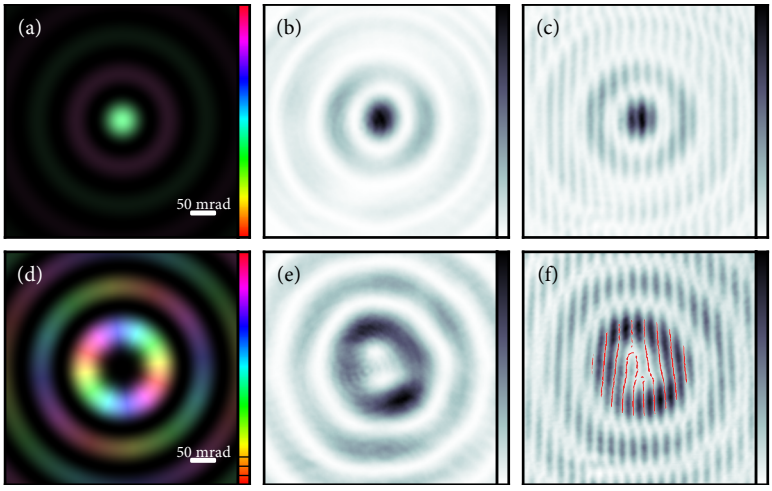


Figure 3.7: Calculated and measured far-field diffraction pattern of the circular slit, split into $\hat{\boldsymbol{\sigma}}_+$ (top row) and $\hat{\boldsymbol{\sigma}}_-$ (bottom row) components. (a, d) Calculated intensity and phase in the far field; luminance indicates intensity, and hue (cycling according to the color bars from 0 through 2π) indicates phase. The $\hat{\boldsymbol{\sigma}}_-$ component has $|Q| = 2$. (b, e): Measured intensity of both components, showing good agreement with the calculations. (c, f): Interferograms using reference beams with appropriate polarization, demonstrating the phase of both components. In (c), the fringes are parallel, indicating a flat wavefront with $Q = 0$. In (f), on the other hand, one fringe splits into three, indicating a helical wavefront with $|Q| = 2$, as in the calculations.

not exactly a quarter wave. We find that the polarization conversion efficiency η is *independent* of the slit's dichroism but depends on the relative phase retardation $\Delta\varphi$ between the polarization components as follows:

$$\eta = I_-/I_{\text{total}} = \sin^2(\Delta\varphi/2), \tag{3.2}$$

where I_- is the intensity of the $\hat{\boldsymbol{\sigma}}_-$ component. If the slit were to behave like a half-wave retarder, then η would become unity. However, designing a half-wave-like slit would once again require careful research to find a suitable width, depth, and material.

This last result suggests that optical spin-orbit conversion is a universal property of a circular nanoslit as long as the local polarization eigenmodes have different propagation constants and are not damped too differently. In order to obtain 100% conversion efficiency one obviously has to adjust the properties of the slit to the wavelength of the incident light in a way similar to the design of a liquid-crystal based q-plate[15] for a certain wavelength. An attractive benefit of this approach to optical spin-orbit conversion is that it is universal, i.e. it can be used at wavelengths from the deep UV to the far infrared.

[15] Marrucci et al., 2006.

One may wonder what happens when the metallic nanoslit is no longer cylindrically symmetric but encircles a singly connected domain. Since the circular symmetry is broken, transfer of angular momentum to the sample is no longer forbidden. For a quarter-wave-like slit with a circularly polarized Gaussian beam incident on it, half of the emerging light will have the opposite circular polarization and carry a charge 2 vortex

with a broad orbital angular momentum spectrum. Contrary to the case of a circular slit, this vortex will be anisotropic.

3.6 Summary

We have demonstrated spin-to-orbital angular momentum conversion of an electromagnetic field upon transmission through a circular metallic nanoslit. When illuminated with circularly polarized light, part of the field transmitted through the slit is converted to the opposite handedness and its topological charge is increased or decreased by 2, corresponding to a conversion of spin angular momentum to orbital angular momentum. The conversion efficiency is a function of the relative phase delay that the slit imposes on orthogonal polarization components. This means that full spin-orbit conversion could be achieved simply by passing the light through a slit in a thin metal film, if the slit were to behave like a half-wave retarder. Using a slit without circular symmetry, on the other hand, opens up a new world of possibilities for creating anisotropic optical vortices.

Appendix 3.A Plasmon-assisted transmission

This section is an appendix that did not appear in the published paper.

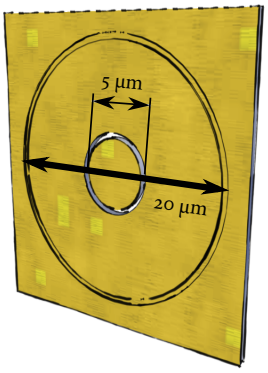


Figure 3.8: A sketch of the groove-slit nanostructure milled into the sample.

We also conducted experiments on a ring slit similar to the one in Fig. 3.2, with an added groove which serves as a surface plasmon outcoupler. On the same substrate, we milled a circular slit, 5 μm in diameter and 200 nm wide, and then a circular groove concentric with the slit, 20 μm in diameter and also 200 nm wide. The groove is essentially a slit which is not deep enough to reach all the way through the gold layer. Due to the focused-ion beam being depth-calibrated for silicon substrates and not gold, the exact depth of the groove is uncertain, but we estimate (100 ± 25) nm. Fig. 3.8 shows a sketch of this structure.

Since the slit's quarter-wave plate-like behavior relies strongly on the loss due to surface plasmon generation, we can expect surface plasmons to travel radially outwards from the slit. When these surface plasmons reach the groove, they are partly scattered into free space as propagating light. We expect this light to be radially polarized around the symmetry axis of the slit. To measure this scattering, we conducted the experiment in exactly the same way as described in section 3.2, except that we overexposed the CCD camera in order to detect the much weaker scattering from the groove.

Figures 3.9 and 3.10 show the results of this experiment. The transmitted intensity (Fig. 3.9a) is more complicated to interpret in this case. For one thing, it exhibits blooming[16], which blots out a small section of the groove.

[16] Blooming is the vertical streaking visible when overexposure causes too many electrons to accumulate in the potential well of a CCD pixel, making them overflow to neighboring pixels.

Also, the slit and groove are subwavelength, making it impossible to image them perfectly. In practice, this means that the crisp boundaries of the slit and groove are softened and widened, and unwanted garbage shows up on the camera outside of the slit and groove. The usual way

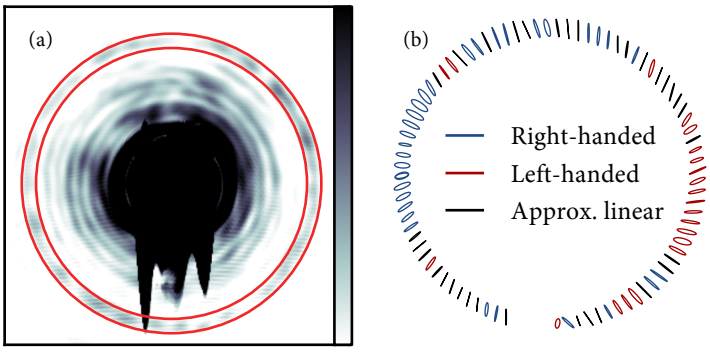


Figure 3.9: (a) Measured intensity emitted by the ring groove (delineated in red). The light transmitted through the smaller ring slit is obscured by blooming due to overexposure. (b) Local polarization ellipses of the light emitted by the ring groove. Blue ellipses indicate right-handed elliptical polarization, red ones indicate left-handed elliptical polarization, and black lines indicate polarization states with ellipticity less than 10%.

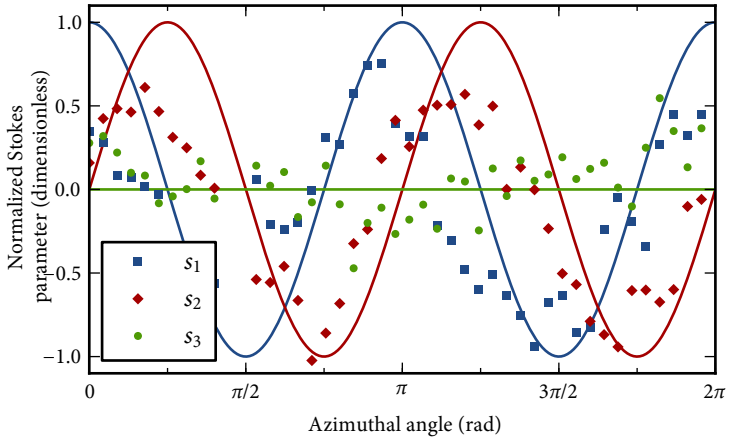


Figure 3.10: Measured normalized Stokes parameters $s_1 = S_1/S_0$, $s_2 = S_2/S_0$, $s_3 = S_3/S_0$ of the light emitted by the ring groove as a function of azimuthal angle. This shows the same information as Fig. 3.9, but here it is easier to compare it to the expected results (solid lines). Compare Fig. 3.5.

of explaining this phenomenon is to define a PSF for the imaging system. The PSF can also be viewed as the impulse response of the imaging system, the impulse being an infinitesimal point source. The source field can then be viewed as a superposition of infinitesimal point sources, and the field at the image plane of the imaging system is a superposition of point spread functions. In other words, the output field is the convolution of the input field with the point spread function. The intensity point spread function of an ideal imaging system is an Airy disc.[17] In Fig. 3.4a, the point spread function is barely visible, because the outer rings of the Airy function are very faint, but since the light emerging from the groove is much fainter than the slit, they are of comparable intensity. Therefore, the groove is marked in Fig. 3.9a in between two concentric circles.

PSF: point-spread function

Taking into account that the polarization measurements in Fig. 3.9 are less accurate than those in Fig. 3.4, we still note that the groove emits light that is more or less radially polarized. We compare the measurements to the expectation in Fig. 3.10.

Here, also, we calculate the expectation value of the spin and orbital angular momenta per photon averaged over the whole beam in the output state of the plasmon-assisted transmitted light, shown in Fig. 3.10. This output state has $S = 0, L = \hbar$, again showing that the total angular momentum per photon, $J = S + L$, is conserved.

[17] When dealing with the field, one should actually use a complex 3-vector-valued point spread function (Marian et al., 2007), but here we will assume that there is no coupling between TE and TM components due to the imaging system.

Appendix 3.B Plasmonic cross-talk between points on the ring

This section is an appendix that did not appear in the published paper.

CONSIDERING THE PLASMONIC CONTRIBUTION in the slit-only system may also help to explain why the polarization in Fig. 3.4b is not purely

linear. The light incident on the slit is converted linearly to a surface plasmon, barring an unknown attenuation and retardation factor which we will ignore for now. These surface plasmons travel from one side of the circle to the other.

TM: transverse magnetic

Only the TM component excites a plasmon, with $\hat{\mathbf{z}}$-polarization. The plasmon propagates across the gold surface, undergoing diffraction, and hits the slit again, scattering once again into TM, or $\hat{\mathbf{r}}$-polarized, light. Since the plasmons only couple to radial polarization, the plasmonic contribution to the transmission has a different polarization than the direct contribution. The smaller plasmonic contribution should therefore be visible as a deviation in the polarization of the light emerging from the slit.

To calculate the diffraction the surface plasmons undergo during the transit from one side of the ring to the other, we look at the Fresnel-Kirchhoff diffraction integral: [18]

[18] As in Griffiths (1999), $\boldsymbol{\imath}$ denotes the separation vector between a source point $\mathbf{r}'$ and a field point $\mathbf{r}$: $\boldsymbol{\imath} \equiv \mathbf{r} - \mathbf{r}' = (x-x')\hat{\mathbf{x}} + (z-z')\hat{\mathbf{z}}$.

$$E(x,z) = \frac{1}{i\sqrt{\lambda_{SP}}} \int E(x',0) \frac{e^{ik\boldsymbol{\imath}}}{\sqrt{\boldsymbol{\imath}}} \cos\eta \, dx'$$

Since the diffraction takes place in two dimensions, the Huygens waves scattered by each point on the wavefront are not spherical (e^{ikr}/r) but instead damped cylindrical waves[19] ($e^{ikr}/\sqrt{r}$). Here, z is the diffraction distance along the propagation axis. The separation vector $\boldsymbol{\imath}$ is the distance between a source point x' in the $z = 0$ plane and the point x that we are interested in in the image plane. The angle $\eta = \arccos z/\boldsymbol{\imath}$ is the angle between the propagation vector and the separation vector, so $\cos\eta$ can also be written as $\hat{\mathbf{k}} \cdot \hat{\boldsymbol{\imath}}$.

[19] Teperik, Archambault, Marquier, and Greffet, 2009.

Based on this, we can construct the following diffraction integral in polar coordinates for our ring-slit geometry, shown in Fig. 3.11:

$$E_{SP}(R_0,\theta) = \frac{1}{i\sqrt{\lambda_{SP}}} \int_{-\pi/2}^{\pi/2} E_{0,SP}(R_0, \theta+\zeta) \frac{e^{ik_{SP}\boldsymbol{\imath}}}{\sqrt{\boldsymbol{\imath}}} \hat{\mathbf{k}} \cdot \hat{\boldsymbol{\imath}} R_0 \, d\zeta \tag{3.3}$$

The physical meaning of this integral is that for each point on the ring, the surface plasmon-assisted field is a sum of the contributions from points elsewhere on the ring. The point (R_0, θ) that we are interested in only receives contributions from the facing inner side of the ring: angles $\theta+\pi/2$ to $\theta + 3\pi/2$.

After some calculation, we can write:

$$E_{SP}(R_0,\theta) = iE_{0,SP}(R_0,0)\tilde{f}(k_{SP}R_0) \tag{3.4}$$

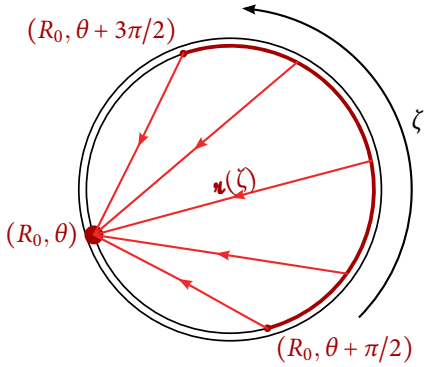


Figure 3.11: Diffraction geometry of surface plasmons traveling inside the circular slit from one side to the other. The plasmon-assisted transmission at each point on the ring (R_0, θ) is the sum of contributions from the plasmonic field launched on the semicircle opposite on the ring. When converted back to light, this plasmonic contribution should be purely radially polarized, which should be visible as an alteration of the polarization direction of the directly transmitted contribution.

where

$$\tilde{f}(q) = \frac{1}{4}\sqrt{\frac{q}{\pi}} \int_{-\pi/2}^{\pi/2} e^{i(\zeta + 2q\cos(\zeta/2))} \frac{1+\cos\zeta}{(\cos(\zeta/2))^{3/2}}\, d\zeta$$

Adding the direct and plasmonic contributions, we see that the light emerging from the slit is not necessarily linearly polarized anymore:

$$\tilde{\mathbf{E}}_{0,\text{out}}(R_0, \theta) = \frac{i}{\sqrt{2}} e^{i\theta} \left(\left(1 + \tilde{A}\tilde{f}(k_{\text{SP}} R_0)\right) \hat{\mathbf{r}} + \hat{\boldsymbol{\theta}} \right) \tag{3.5}$$

where $\tilde{A}$ is the unknown attenuation and retardation due to conversion between light and surface plasmons and vice versa. The plasmonic contribution adds a small degree of ellipticity to the polarization everywhere, depending on the phase of $\tilde{A}$.

4
Plasmonic optical vortex tomography

We present a novel method for analyzing the wavefront of optical vortices which does not involve interferometry, but rather uses surface plasmons. We employ a subwavelength slit in a gold film to cut slices from an optical vortex beam, and measure the diffraction of the generated surface plasmons by scattering them off a second slit. By moving the slits across the vortex beam, we create a tomogram, from which we can determine the vortex charge of the incident beam at a glance. We present results for vortex beams of integer and half-integer vortex charge.

4.1 Introduction

Portions of this chapter were previously published as: Chimento, 't Hooft, and Eliel (2010a), Chimento, 't Hooft, and Eliel (2010b).

Vortices occur in branches of physics as varied as superconductors, superfluids, Bose condensates, fluid flow, and optics. A property that all vortices share is that, when traversing a closed path around a vortex, an order parameter of the system changes by $2\pi Q$, with Q the "charge" of the vortex, the sign of which is associated with a direction of circulation. In this chapter, we will concern ourselves with phase vortices in the transverse field distribution of an optical beam, a subject that has attracted considerable attention in recent years.[1] Vortex beams have found applications in optics at both microscopic[2] and astronomical[3] scales. They also occur naturally in speckle fields scattered from inhomogeneous or rough surfaces.[4]

An optical vortex in its simplest form, namely the transverse cross section of a vortex beam, manifests itself as a doughnut-shaped intensity distribution; the phase increases azimuthally around the doughnut and the intensity vanishes at the center because the phase is undefined there. The number of cycles with which the phase increases on a closed loop around the doughnut equals the vortex charge Q. The photons in a vortex beam of charge Q carry $Q\hbar$ orbital angular momentum.[5]

There are several ways to measure the charge of an optical vortex beam.

[1] Soskin and Vasnetsov, 2001; Allen, Barnett, and Padgett, 2003.

[2] Foo, Palacios, and Swartzlander, 2005.

[3] Jesacher et al., 2007.

[4] Baranova et al., 1981.

[5] Allen, Beijersbergen, et al., 1992.

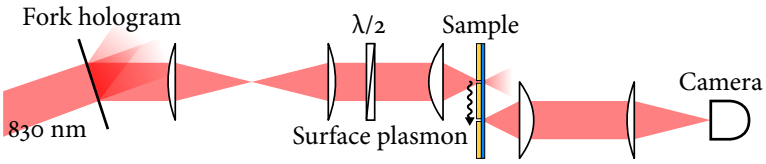


Figure 4.1: Schematic of the experimental setup. The appropriate diffraction order of the fork hologram is selected by means of an aperture (not shown); the others are blocked. A typical fork hologram is shown in Fig. 4.2, while a typical nanostructure on the sample is shown in Fig. 4.3.

Since it is a property of the light beam's phase, some sort of interferometry must be used. One way is to interfere the light beam with itself [6] or with a plane wave[7] and examine the fringe pattern, which contains a dislocation at the position of the vortex. Other ways are to build a mode sorter,[8] or use a multipoint interferometer and calculate the vortex charge from the resulting interference pattern.[9]

[6] Harris, Hill, Tapster, and Vaughan, 1994.

[7] Padgett, Arlt, Simpson, and Allen, 1996.

[8] Mair, Vaziri, Weihs, and Zeilinger, 2001; Leach, Padgett, Barnett, Franke-Arnold, and Courtial, 2002.

[9] Berkhout and Beijersbergen, 2009.

HERE WE PRESENT a simple and elegant method of determining the vortex charge of an optical beam. It is based on the use of surface plasmons. These surface plasmons are generated by scattering the incoming vortex beam off a narrow emitter[10] slit milled in a surface plasmon-supporting gold film. A second receiver slit, which is some distance from the emitter, picks up the diffracted surface plasmon wave, converting it back to a free-space optical beam. By translating the gold film across the vortex beam, we construct a tomographic pattern of the plasmonic diffraction that allows direct visualization of the vortex charge if it is an integer. If the vortex charge is not an integer, it is still possible to estimate it.

[10] The slits are physically identical, but we call them "emitter" and "receiver" to distinguish their role in the experiment.

Surface plasmons are a convenient tool for this tomography, for three reasons. For one, tomography, at its most fundamental, entails slicing three-dimensional data into two-dimensional sections without loss of information. Surface plasmons propagate in two dimensions, providing a means for tomography; surface plasmon diffraction provides a means of analysis of the sliced field. Second, we can achieve subwavelength resolution in our tomograms by translating the subwavelength slits in subwavelength steps. Finally, the coherent conversion of light to surface plasmons and vice versa allows transportation from the emitter to the receiver without loss of information, except for some power loss.

4.2 *Integer vortex experiment*

FIGURE 4.1 SHOWS our experimental setup. We create a linearly polarized beam of integer vortex charge by diffracting a Gaussian beam (λ = 830 nm) off of a computer-generated fork hologram,[11] shown in Fig. 4.2. The beams diffracted from this grating carry a vortex charge dependent on

[11] Bazhenov, Vasnetsov, and Soskin, 1990.

the diffraction order and the vortex inscribed in the hologram; we select a diffraction order that carries the desired vortex charge. Once the beam propagates to the far field of the hologram, it has the doughnut-shaped intensity and azimuthal phase as described above.

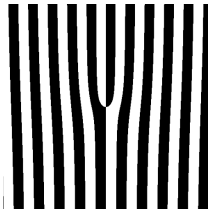


Figure 4.2: Typical design of a fork hologram with vortex charge 2.

A 4*f* lens system transports the beam to the back focal plane of a microscope objective, which focuses the far field of the beam onto the sample, down to a size of several microns. On its way, the beam passes through a half-wave plate which allows the experiment to be conducted with any desired linear polarization.

The sample (Fig. 4.3) consists of a gold film, 200 nm thick, attached to a glass substrate by a 10 nm titanium adhesion layer. The strongly dissipative titanium layer ensures that surface plasmons can only propagate on the gold-air interface.[12] The sample contains pairs of double slits, which are ion-beam milled through the gold. The slits used in this experiment are all 50 µm long and 100 nm wide, and the pairs are separated by 25, 50, and 75 µm. For comparison, the damping length of surface plasmons on gold at this wavelength is around 50 µm. We illuminate the emitter slit with our vortex beam and image the light emerging from the receiver slit onto a CCD camera. The incident beam is polarized so as to couple optimally with surface plasmons.

[12] Schouten, Kuzmin, et al., 2005.

CCD: charge-coupled device

We illuminate the emitter slit with the beam, causing it to launch a surface plasmon wave along the gold-air interface towards the receiver slit.[13] Its amplitude at the emitter is given by the local field amplitude incident on the slit. In between the slits, the wave diffracts freely. The receiver slit scatters the diffracted plasmonic field into free space. The diffraction pattern contains information on both the phase and the amplitude of the light incident on the emitter slit.

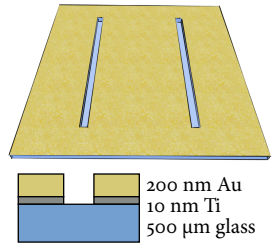


Figure 4.3: Sketch of a typical nanostructure on the sample. The slits are 100 nm wide and 50 µm long, and separated by 25, 50, or 75 µm.

[13] The emitter also launches a wave in the other direction, but we do not detect this wave.

The sample is mounted so as to allow translation transverse to the optical axis. At the start of a measurement, the beam is incident to one side of the slit pair. We translate the sample along the positive z-axis in 100 nm increments so that the emitter slit travels through the beam. At each position of the sample, we record the intensity profile of the light scattered from the receiver slit. We then assemble these profiles together in a tomogram, so that each vertical slice of the tomogram corresponds to one slice of the incident vortex beam after propagation from emitter to receiver. Figure 4.4 shows a sketch of the tomography process.

We calculate the expected tomograms by modelling the emitter slit as a plasmonic line source with its field amplitude given by the incident vortex beam's free-space field amplitude at that point on the sample. We then calculate the evolution of this field under propagation from emitter to re-

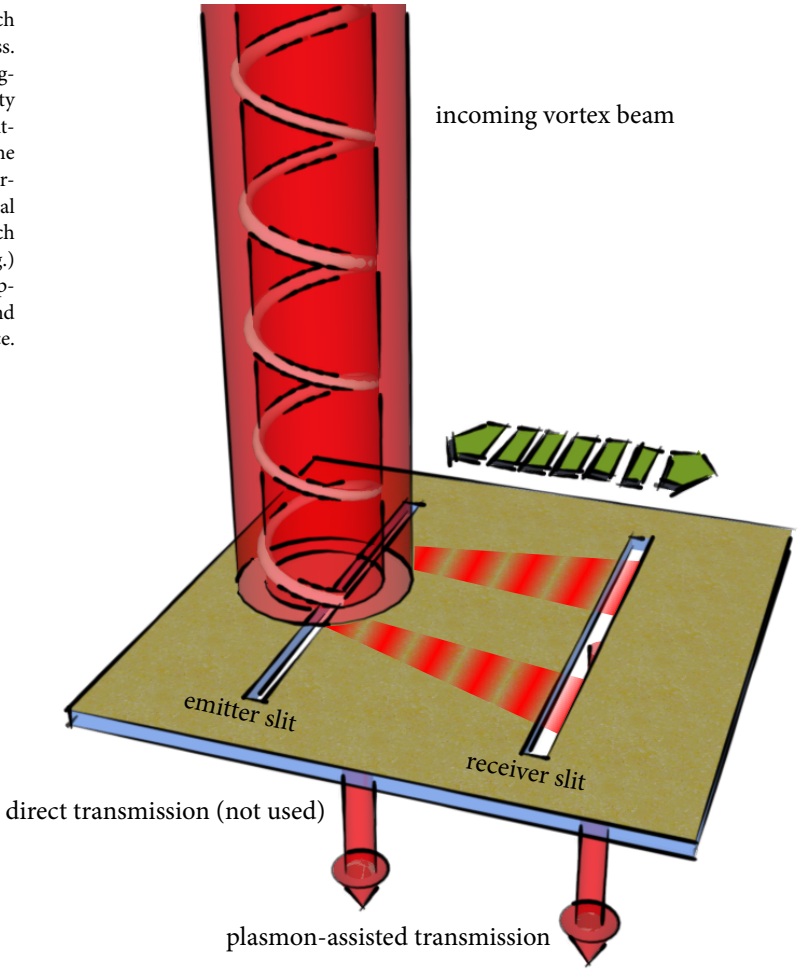


Figure 4.4: Close-up sketch of the tomography process. A vortex beam, with a ring-shaped transverse intensity profile, is incident on the emitter slit. The slit scatters the incident light, launching surface plasmons, proportional to the field amplitude (which is very low inside the ring.) The surface plasmons propagate to the receiver slit and are scattered into free space.

ceiver, using the Fresnel-Kirchhoff diffraction integral, modified for surface plasmons.[14] We model the receiver slit as another line, which scatters light into free space proportional to the plasmonic amplitude it receives.

[14] Teperik et al., 2009.

4.3 *Tomograms*

FIGURE 4.5 SHOWS the calculated and measured tomograms for incident beams of vortex charge $Q = +1$, -1, and -3, respectively, using slits separated by 25 μm. The tomographic patterns are very different from the original ring-shaped vortex beams. First, the pattern is no longer rotationally symmetric, but has a two-fold symmetric, elongated shape. Second, the patterns for beams with vortex charge $Q = +1$ and $Q = -1$ are

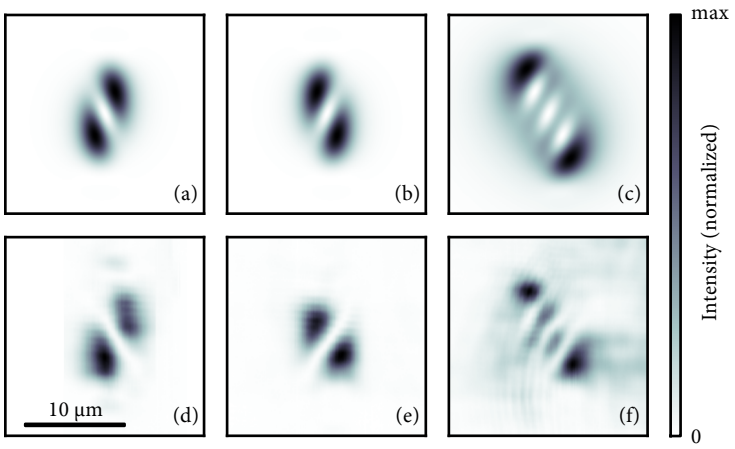


Figure 4.5: Calculated (a–c) and experimental tomograms (d–f) for beams on slits separated by 25 μm, with vortex charge $Q = +1$ (a and d), $Q = -1$ (b and e), and $Q = -3$ (c and f). Note that the intensity zeroes in (c) and (f) occur nicely along a straight line.

each other's mirror image. Third, the orientation of the long axis of the pattern carries the sign of the vortex charge. Finally, the magnitude of the vortex charge can be read off directly from the number of spatially separated intensity zeroes in the pattern. Our calculations, which are in excellent agreement with our experimental results, indicate that these intensity zeroes correspond to phase vortices in the tomogram.

Figure 4.6 shows calculated and measured tomograms for a $Q = -1$ incident beam for slit pairs with increasingly larger separations. As the distance between the slits increases, the tomographic pattern remains essentially the same but spreads out more, approaching Fraunhofer diffraction of the surface plasmons.

4.4 *Interpretation*

There is an intuitive way of explaining why the observed tomographic patterns look as they do. The surface plasmon field amplitude at the emitter slit diffracts as the plasmons travel from the emitter to the receiver. We can consider the field's amplitude at the emitter equivalent to an amplitude mask in a screen, with bright areas corresponding to slits. The screen is illuminated from behind by a plane wave at an angle determined by the field's phase at the emitter. If we place a second screen at some distance, corresponding to the receiver, then the positions of the bright and dark spots in the tomogram follow directly by considering where constructive and destructive interference occur on the second screen.

We discuss the $Q = -3$ case in some detail with the aid of Fig. 4.7. In

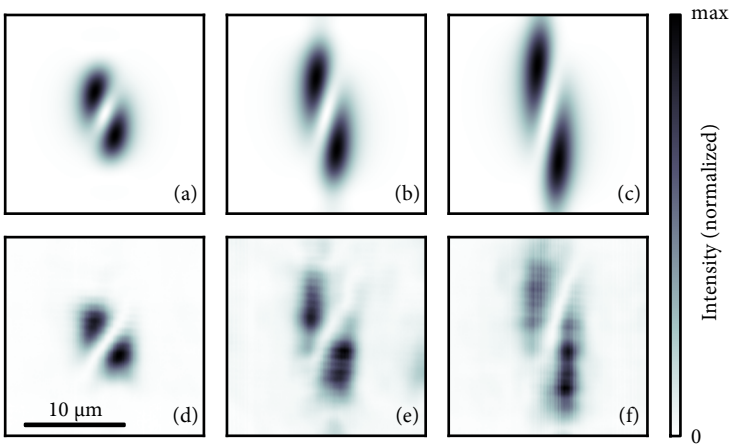


Figure 4.6: Calculated (a–c) and experimental tomograms (d–f) for a $Q = -1$ vortex beam on slits separated by distances of 25 µm (a and d), 50 µm (b and e), and 75 µm (c and f).

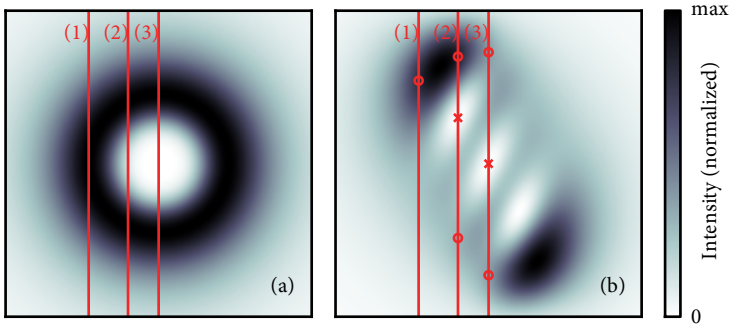


Figure 4.7: (a) Calculated intensity pattern of a $Q = -3$ vortex beam. Three important cross sections are indicated by vertical lines. (b) Calculated tomogram for a $Q = -3$ vortex beam on slits separated by 25 µm (cf. Fig. 4.5c). The indicated cross sections are identical to those of Fig. 4.7a. Local minima and maxima in the tomograms are marked by × and ○, respectively. These marks correspond to those in the diffraction patterns depicted schematically in Figs. 4.8, 4.9, and 4.10.

Fig. 4.7a we depict the intensity pattern of the incident beam at the surface of the gold film; it is intersected by three vertical lines labeled 1, 2 and 3, corresponding to three different positions of the emitter slit. In Fig. 4.7b we show the tomogram of the $Q = -3$ beam, with the equivalent three positions of the receiver slit. Line 1 is tangent to the ring of maximum intensity of the $Q = -3$ input beam. In the tomogram, the intense diffraction spot along line 1 corresponds to the case that the emitter slit picks up an essentially single-spot intensity distribution with a slanted phase front, arising from the azimuthal phase dependence of the incident beam. The corresponding diffraction pattern is that of a plane wave incident at an angle through a single slit: a single off-axis spot results (see Fig. 4.8).

When the emitter slit is positioned more towards the center of the incident beam, along line 2, the generated plasmonic field is bimodal with

Figure 4.8: Diffraction pattern of light at a single slit under oblique incidence. Corresponds to line 1 in Fig. 4.7.

a dip between the intensity maxima. In this region the phase of the plasmonic field varies steeply and linearly with position along the emitter slit. At this particular position of this slit, the relative phase of the two maxima of the bimodal intensity distribution equals 2π. Conceptually, the plasmonic diffraction pattern should therefore be similar to the pattern arising from a double slit illuminated at an angle in such a way that the two slits have equal phase, up to a factor 2π (see Fig. 4.9).

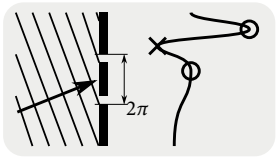


Figure 4.9: Diffraction pattern of light at a double slit under oblique incidence, such that the local field at the slits is in phase. Corresponds to line 2 in Fig. 4.7.

Finally, when the emitter slit is at line 3, the center of the incident beam, the plasmonic field will again be bimodal, now with a phase difference of 3π. The diffraction pattern at the receiver slit will be double-slit-like with a zero in the center as a result of destructive interference (see Fig. 4.10).

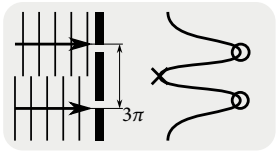


Figure 4.10: Diffraction pattern of light at a double slit in antiphase. Corresponds to line 3 in Fig. 4.7.

4.5 *Non-integer vortex experiment*

In order to consider the application of this method to more complex vortex-carrying fields, we also conducted experiments with a non-integer vortex beam, using a spiral phase plate[15] to generate the desired field with vortex charge $Q = 3\frac{1}{2}$. Figure 4.11 is a sketch of a spiral phase plate. The far-field diffraction pattern of such a beam is not rotationally symmetric, so we oriented the phase plate to produce an incident intensity pattern as shown in Fig. 4.12a, with the slits oriented vertically. The incident pattern shows three close-lying, separated $Q = +1$ vortices in the center, with an additional one intruding from the bottom. Figures 4.12b and 4.12c show two measured tomograms at different slit separations, while Figs. 4.12d and 4.12e show the corresponding calculations. These tomograms are devoid of any symmetry. Specifically, the three vortices are not arranged along a straight line, unlike the integer-vortex case. They also show a fourth vortex intruding from the side, although the visibility of the fourth vortex in the measurements is somewhat marginal. Currently we do not account for any inhomogeneity in the slit width; this problem might be solved by using a ptychographical algorithm, which iteratively reconstructs a field's complex amplitude, and the transfer function of the object used to probe it.[16]

For a better understanding of the relation between the fractional part of the vortex charge and the presence and position of the fourth vortex in the tomogram, we calculated the tomograms of beams with various vortex charges between +3 and +4, shown in Fig. 4.13. We see that as Q increases, the fourth vortex, as seen in Fig. 4.12, approaches the three existing vortices, and eventually joins them in a straight line at $Q = +4$. The vortices are arranged in a straight line only when Q is an integer. This suggests

[15] Oemrawsingh et al., 2004.

[16] Maiden and Rodenburg, 2009.

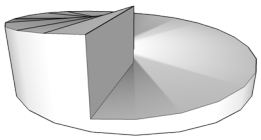


Figure 4.11: Sketch of a typical spiral phase plate.

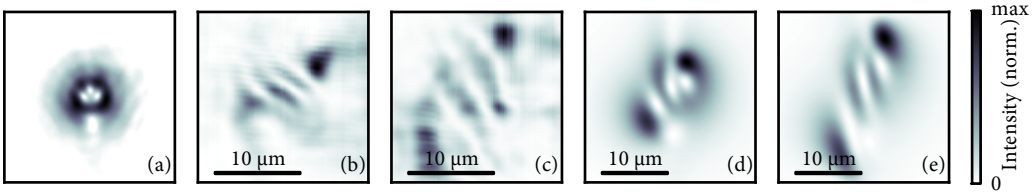


Figure 4.12: (a) Far-field diffraction pattern of a $Q = +3\frac{1}{2}$ vortex beam; (b) experimental tomogram of this beam on slits separated by 25 μm; (c) 75 μm; (d) calculated tomogram of an ideal $Q = +3\frac{1}{2}$ vortex beam on slits separated by 25 μm; (e) 75 μm. Three intensity nodes are visible, but they are not arranged along a straight line.

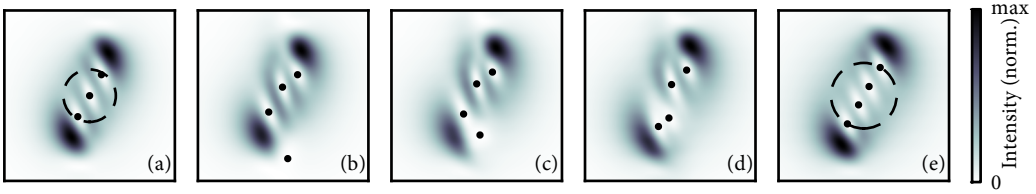
that a non-integer vortex charge can be determined from the deviation of the vortices' arrangement from a straight line, the angle of which is determined mainly by the ratio of the distance between the slits to the spot size of the beam. Our calculations indicate that any dependence on Q is less than 1° and may be a numerical artifact.

4.6 *Summary*

We have used the diffraction of surface plasmon polaritons to analyze a vortex-carrying light beam slice by slice, in order to recover information about the beam's phase: specifically, its vortex charge. Although the determination of non-integer vortex charges is not possible at a glance, we have shown through calculations that the magnitude of a non-integer vortex charge may be determined by measuring how much the arrangement of the vortices deviates from a straight line.

Phase retrieval normally requires some technique such as interferometry or a combination of near-field and far-field measurements. The current experiment's two slits can be considered to measure the surface plasmons' near field and far field. Therefore, the technique might be generalized to phase retrieval of arbitrary fields. Also, the sample's small size and the small distances between the optics involved suggest that the experiment can be easily miniaturized. Therefore, it has a potential application as a wavefront sensor.

Figure 4.13: Calculated tomograms of vortex beams with Q ranging from +3 to +4 on slits separated by 50 μm. (a) $Q = +3$; (b) $Q = +3\frac{1}{4}$; (c) $Q = +3\frac{1}{2}$; (d) $Q = +3\frac{3}{4}$; (e) $Q = +4$. The locations of the intensity zeroes are indicated with dots. In (a) and (e), the size of the original vortex ring is superimposed on the tomogram.



PART II

ANOMALOUS SURFACE PLASMON DISPERSION IN ALUMINUM

5
Surface plasmon coupling by attenuated total reflection for Drude-like metals

We discuss the influence of the optical properties of the metal used in a surface plasmon resonance experiment on the lineshape of the measured resonance curve. We also examine whether it is better to perform such experiments in the Kretschmann or Otto configuration, and find that the Otto configuration has some often-overlooked advantages. In addition, we present a phenomenological method for analyzing all possible lineshapes of surface plasmon resonance curves, that yields the complex surface plasmon mode index without a priori *knowledge of the composition of layers of metal and dielectric in the experiment.*

5.1 Introduction

This chapter has been submitted to *Optics Express* for publication.

A COMMON CONFIGURATION for studying surface plasmons at the interface between a metal and a dielectric is the attenuated total reflection setup. The most widely used variant is known as the Kretschmann configuration;[1] it is used in many applications, for instance in the bio-analytical sciences.[2] Various companies offer fully automated SPR analyzers for this purpose, starting with Biacore (now GE Healthcare) in the early 1990s;[3] Rich and Myszka[4] give an overview of recent devices.

An alternative to the Kretschmann configuration, known as the Otto configuration,[5] is much less frequently employed because it is generally presumed to be considerably more awkward experimentally. However, there are experimental systems where the Otto configuration outperforms the Kretschmann approach. One of the aims of the present chapter is to investigate when this applies and why that is so. We will also discuss the

[1] Kretschmann, 1971.

[2] Liedberg, Nylander, and Lunström, 1983.

SPR: surface plasmon resonance

[3] Rich and Myszka, 2008; Fivash, Towler, and Fisher, 1998.

[4] Rich and Myszka, 2010.

[5] Otto, 1968.

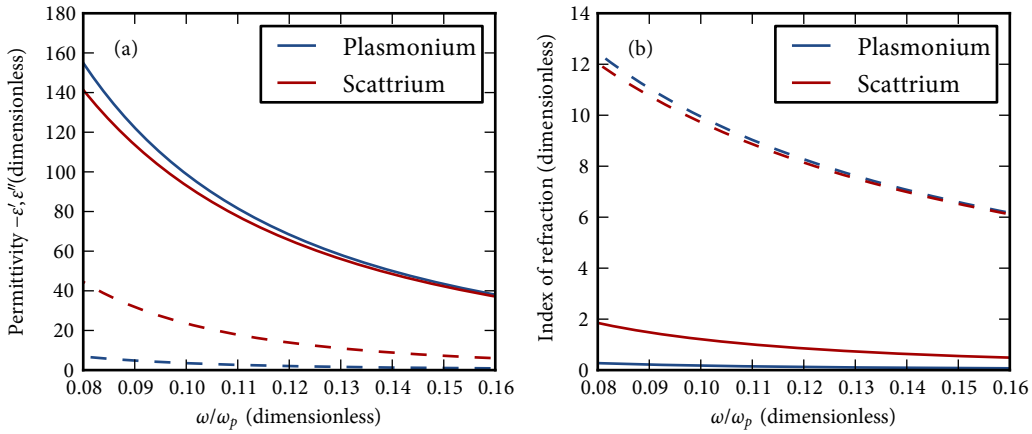


Figure 5.1: Dielectric function (a) and index of refraction (b) of the fictitious Drude metals plasmonium (blue) and scattrium (red). The solid lines, both left and right, indicate the real parts of the displayed quantity, and the dashed lines the imaginary parts. The real part of the dielectric function (solid line on left) is plotted with its sign flipped, i.e. as $-\varepsilon'$, so as to fit both quantities into a similar scale.

proper interpretation of SPR measurements when straying from the often-used metals of gold and silver.

We will base our discussion on the Drude model for the dielectric function of a metal:

$$\varepsilon(\omega) = 1 - \frac{\omega_p^2}{\omega(\omega + i\gamma)}, \tag{5.1}$$

where ω_p is the bulk plasma frequency of the metal and γ is the damping frequency related to the electron scattering time τ by $\gamma = 1/\tau$. The Drude model is a good approximation for many metals, in particular for the alkali metals such as lithium, sodium, and potassium. It applies also quite well to more mundane metals such as silver and aluminum for frequencies sufficiently far removed from an interband transition. We shall define dimensionless frequencies $\Omega = \omega/\omega_p$ and $\Gamma = \gamma/\omega_p$ so that the Drude model has only one material-dependent parameter:

$$\varepsilon(\Omega) = 1 - \frac{1}{\Omega(\Omega + i\Gamma)}.$$

In all interesting cases $\Gamma \ll 1$. The metallic regime is characterized by $\Omega < 1$.

FOR THE DISCUSSION AT HAND it is useful to introduce two fictitious Drude metals, which we will name plasmonium ($\Gamma = 0.0035$) and scattrium ($\Gamma = 0.025$). In this chapter, we will illustrate our findings with an octave of frequencies from $\Omega = 0.08$ to $\Omega = 0.16$, which is a relevant range for the analysis of our experiments on aluminum discussed

in the next two chapters. Plasmonium is similar to an idealized version of silver, while for short wavelengths, scattrium is an idealized version of aluminum. The dielectric functions $\varepsilon = \varepsilon' + i\varepsilon''$ of plasmonium and scattrium are shown in Fig. 5.1. Note that for this choice of parameters, we can approximate $\varepsilon' \approx -\Omega^{-2}$ and $\varepsilon'' \approx \Gamma\Omega^{-3}$: the real parts of the dielectric functions ε' of the two Drude metals are nearly equal, but the imaginary parts ε'' differ by the ratio of the electron scattering times.

5.2 *Surface plasmons on an interface between two semi-infinite materials*

THE QUINTESSENTIAL SETTING for studying surface plasmons is the interface between a half-space ($z < 0$) of metal (with relative permittivity ε_1) and a half-space ($z > 0$) of dielectric (with relative permittivity ε_2). Figure 5.2 is an illustration of this situation. With the interface at $z = 0$, and assuming that the surface plasmons travel in the x direction, the surface plasmon field is fully determined by the y component of the magnetic field $\mathbf{H}$. To determine this field we calculate the transfer matrix[6] for incoming and outgoing H_y amplitudes from both sides of the interface:

$$\begin{bmatrix} H_{y2}^{+} \\ H_{y2}^{-} \end{bmatrix} = \frac{1}{t_{21}} \begin{bmatrix} 1 & r_{21} \\ r_{21} & 1 \end{bmatrix} \begin{bmatrix} H_{y1}^{+} \\ H_{y1}^{-} \end{bmatrix}, \tag{5.2}$$

where $H_{yn}^{\pm}$ indicates the wave traveling in the $\pm z$ direction, and 1 and 2 represent the two half spaces. The coefficients r_{21} and t_{21} represent the interface reflection and transmission amplitudes, respectively.[7] These complex amplitudes are given by the well-known Fresnel relations (which imply that the appropriate fields are continuous across the interface):

$$r_{pq} = \frac{k_{zp}/\varepsilon_p - k_{zq}/\varepsilon_q}{k_{zp}/\varepsilon_p + k_{zq}/\varepsilon_q}, \qquad t_{pq} = \frac{2k_{zp}/\varepsilon_p}{k_{zp}/\varepsilon_p + k_{zq}/\varepsilon_q}. \tag{5.3}$$

Here $k_{zp}^2 = \varepsilon_p k_0^2 - k_x^2$, where $k_0 = \omega/c$ is the wave vector in vacuum. To find the plasmon mode we choose the sign of k_z in each half space such that the field H_y decays away from the interface. The allowed modes traveling along the interface in the x direction, i.e. the surface plasmons, follow from the requirement that they exist even if all incident fields (H_{y1}^{+}, H_{y2}^{-}) vanish. This requirement yields two surface plasmon modes traveling in the $\pm x$ directions, respectively:

$$\frac{1}{t_{21}} = 0 \implies \frac{k_{z1}}{\varepsilon_1} + \frac{k_{z2}}{\varepsilon_2} = 0 \implies k_x^{\infty} = \pm k_0 \sqrt{\frac{\varepsilon_1 \varepsilon_2}{\varepsilon_1 + \varepsilon_2}}. \tag{5.4}$$

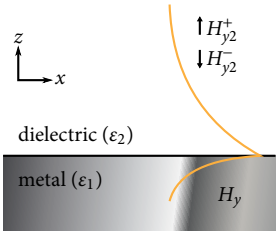


Figure 5.2: Sketch of an interface between two half-spaces of dielectric and metal. A typical H_y amplitude for the surface plasmon mode is sketched in orange. H_y must be continuous across the interface.

[6] Davis, 2009.

[7] We note that $t_{12}t_{21} - r_{12}r_{21} = 1$ and $r_{12} = -r_{21}$.

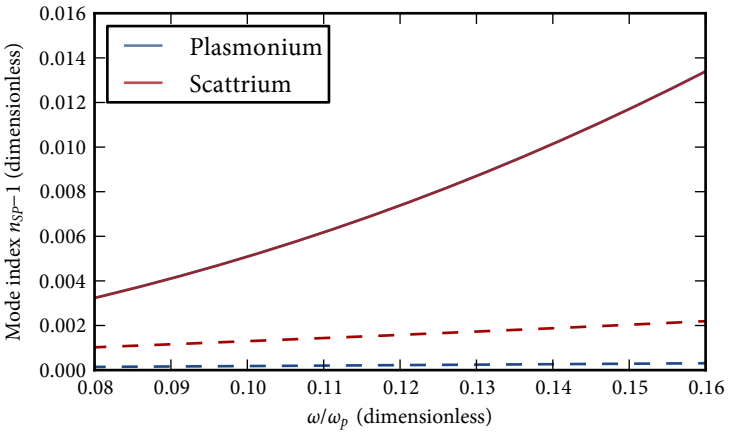


Figure 5.3: Effective mode index for a surface plasmon on an interface between vacuum and one of the fictitious metals plasmonium (blue) and scattrium (red). The solid lines (which coincide almost exactly) indicate the real part, and dashed lines the imaginary part. The real part of the index minus 1 is displayed so as to fit both quantities in a similar scale.

The result of (5.4) is the well-known surface plasmon dispersion relation on a flat interface between half-spaces; we use the notation k_{SP}^{∞} to emphasize that the materials are semi-infinite. Since ε_1 is complex-valued, the value for k_{SP}^{∞} that follows from (5.4) is also complex and can be written $k_{\mathrm{SP}}^{\infty} = k_{\mathrm{SP}}^{\infty\prime} + ik_{\mathrm{SP}}^{\infty\prime\prime}$: the surface plasmon propagates as a damped harmonic wave, with wavelength $2\pi/k_{\mathrm{SP}}^{\infty\prime}$ and $1/e$ amplitude damping length $1/k_{\mathrm{SP}}^{\infty\prime\prime}$.[8] It is convenient to introduce the complex surface plasmon mode index $n_{\mathrm{SP}}^{\infty} = k_{\mathrm{SP}}^{\infty}/k_0$. Figure 5.3 shows the dependence of the real and imaginary parts of this index for surface plasmons travelling along a metal–vacuum interface for plasmonium and scattrium, as a function of the frequency ratio Ω.

[8] This should not be confused with the intensity damping length $1/2k_{\mathrm{SP}}^{\infty\prime\prime}$, which some authors prefer.

When $|\varepsilon_1'|^2 > \varepsilon_1''^2$ and $\varepsilon_2 = 1$, we can approximate the mode index, by expanding the square root of a complex number, as

$$n_{\mathrm{SP}}^{\infty} \approx \sqrt{\frac{\varepsilon_1'}{\varepsilon_1'+1}}\left(1+\frac{i\varepsilon_1''}{2\varepsilon_1'(\varepsilon_1'+1)}\right),$$

which shows that the real part of the mode index only depends on the real part of ε_1.[9] This explains why the real parts of the mode indices are almost exactly the same for plasmonium and scattrium.

[9] Raether, 1988, p. 5.

5.3 *Surface plasmons on a thin metal layer in the Kretschmann configuration*

Investigating surface plasmons on a flat interface with semi-infinitely extending materials is confined to the realm of theory. In reality,

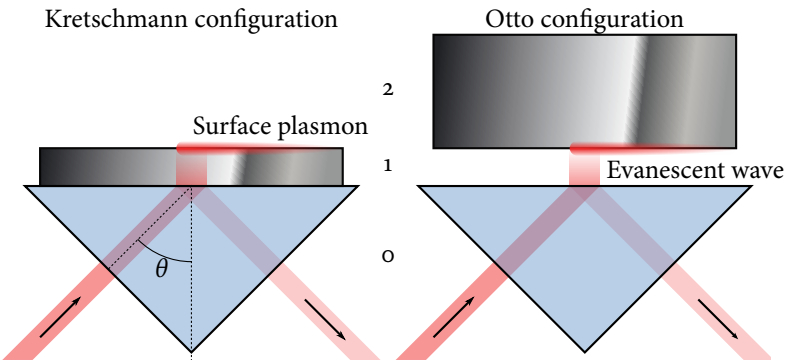


Figure 5.4: The Kretschmann and Otto variants of the attenuated total reflection method for exciting surface plasmons. In both cases, the evanescent wave from total internal reflection in the high-index dielectric (blue) phase-matches (in the direction parallel to the interface) to the surface plasmon mode on the interface between the metal (gray) and low-index dielectric (white).

one needs a way of coupling from freely propagating light to the confined surface plasmon mode and vice versa. Since the surface plasmon mode's wave vector (5.4) is larger than the free-space wave vector for a light wave of the same frequency, the difference in wave vector needs to be made up somehow. Popular methods of coupling to surface plasmons[10] include scattering from a corrugation on the metal surface,[11] increasing the wave vector by using one of the nonzero diffraction orders of a grating on the metal surface,[12] or having the light enter from a dielectric with an index of refraction n_0 that is higher than that of the dielectric that the surface plasmon travels on, so that the wave vector is increased by a factor of n_0.[13]

[10] Sambles, Bradbery, and Yang, 1991.
[11] Jasperson and Schnatterly, 1969.
[12] Ritchie, Arakawa, Cowan, and Hamm, 1968.
[13] Kretschmann, 1971; Otto, 1968.

The latter method, which uses frustrated total internal reflection, has two variations, known as the Kretschmann and Otto configurations, illustrated in Fig. 5.4. Both involve a high-index dielectric substrate, medium 0, and a metal-dielectric interface 1–2. The metal is the thin layer 1 in the Kretschmann configuration, while the Otto configuration has a thin dielectric layer 1 and a bulk metal on top as medium 2. If the light is incident in medium 0 at an angle θ larger than the critical angle θ_{cr} for total internal reflection at the interface 0–1 (Otto) or 0–2 (Kretschmann), then the field[14] at the interface 0–1 can phase-match with the surface plasmon mode at the interface 1–2. When this happens, the reflection from the interface 0–1 takes a sharp dive, since the energy is instead transferred to the surface plasmon mode. This yields SPR curves such as that in Fig. 5.5. This is the principle behind SPR sensing. The angle at which the reflection is most attenuated is known as the resonance angle θ_{SPR}.

[14] In the Otto configuration, this field is evanescent.

The depth of the reflectance dip is a measure for the coupling efficiency, and is a function of the metal layer's thickness. For each wavelength of the incident light, there is an optimum for the metal thickness at which the coupling is critical. At critical coupling, the internal damping is equal to the reradiation losses.[15]

[15] Raether, 1988, p. 12.

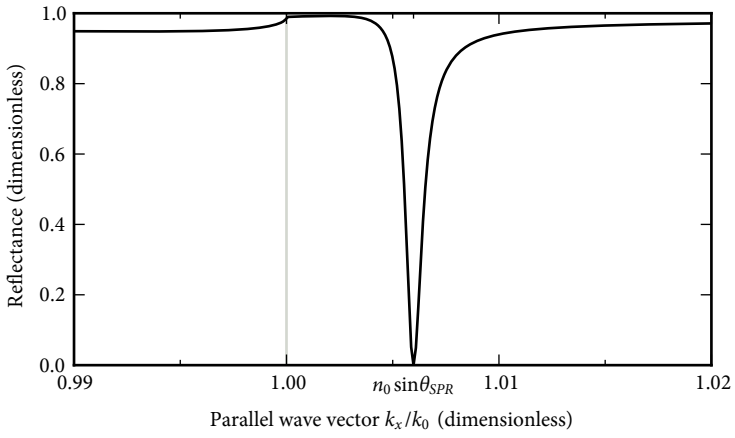


Figure 5.5: Typical SPR curve, here calculated for a plasmonium layer of critical coupling thickness on a glass-like substrate with $n = 1.5$, with vacuum on the outside. The critical angle for total internal reflection from substrate to vacuum is indicated by the gray line at $k_x = k_0$. The resonance angle θ_{SPR}, corresponding to a wave vector parallel to the interface $k_x = k_0 n_0 \sin\theta_{\text{SPR}}$, is the angle at which the largest fraction of the incident light is absorbed into the surface plasmon mode.

In order to come to a set of equations for the surface plasmon in these multilayer stacks, we generalize (5.2) to N layers.[16] This is a powerful set of equations that contains everything we need to know about the system:

[16] Davis, 2009.

$$\begin{bmatrix} H^+_{yN} \\ H^-_{yN} \end{bmatrix} = \begin{bmatrix} M_{00} & M_{01} \\ M_{10} & M_{11} \end{bmatrix} \begin{bmatrix} H^+_{y0} \\ H^-_{y0} \end{bmatrix}, \tag{5.5}$$

where

$$\begin{bmatrix} M_{00} & M_{01} \\ M_{10} & M_{11} \end{bmatrix} = \left(\prod_{n=1}^{N-1} \frac{1}{t_{(n+1)n}} \begin{bmatrix} 1 & r_{(n+1)n} \\ r_{(n+1)n} & 1 \end{bmatrix} \begin{bmatrix} e^{ik_{zn}d_n} & 0 \\ 0 & e^{-ik_{zn}d_n} \end{bmatrix} \right) \times \frac{1}{t_{10}} \begin{bmatrix} 1 & r_{10} \\ r_{10} & 1 \end{bmatrix}. \tag{5.6}$$

As in the two-layer case, we can use this set of equations to calculate various properties of the system. The requirement of having a solution in the absence of incident fields yields

$$M_{11} = 0, \tag{5.7}$$

and solving this for complex k_x gives us the wave vector of the surface plasmon mode. The reflectance, on the other hand, is obtained by calculating the ratio of outgoing to incident power on the side of layer 0, with the condition of no incident field on the side of layer N,

$$R = \left| \frac{-M_{10}(k_x(\theta))}{M_{11}(k_x(\theta))} \right|^2, \tag{5.8}$$

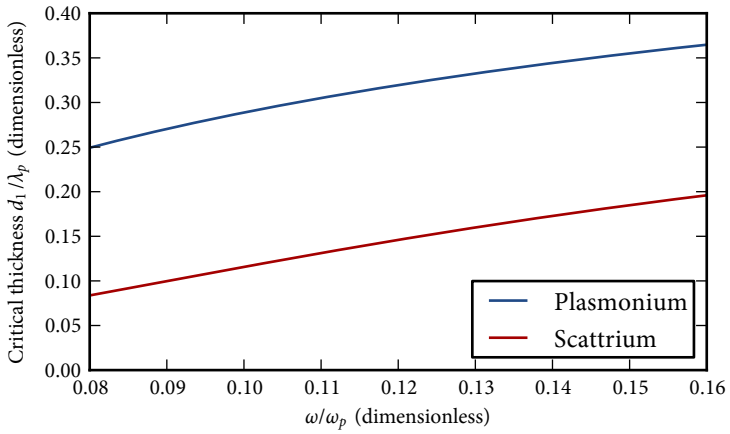


Figure 5.6: Free-space wavelength dependence of the metal thickness for critical coupling in a three-layer Kretschmann configuration of plasmonium or scattrium on a glass substrate with $n = 1.5$ and vacuum on the other side. Note that the critical thickness is given in units of the bulk plasmon wavelength $\lambda_p = \omega_p/2\pi c$.

with $k_x(\theta)$ the real-valued wave vector of the light incident from layer 0. On a related note, the thickness d_1 of layer 1 for which the reflectance vanishes is obtained by solving $M_{10} = 0$ for d_1 with the constraint that k_x is real.

The reflectance of a three-layer Kretschmann system can be written as:

$$R = \left| \frac{r_{01} + r_{12}\delta}{1 + r_{01}r_{12}\delta} \right|^2 , \tag{5.9}$$

with $\delta = e^{2ik_{z1}d_1}$. The condition for surface plasmons (5.7) works out to $r_{01}r_{12}\delta = -1$.

Conversely, the condition for zero reflection and thus critical coupling, $M_{10} = 0$, is equivalent to setting the numerator to zero. It is instructive to write it thus:

$$-r_{01} = r_{12}\delta . \tag{5.10}$$

For unit field amplitude incident on the multilayer stack, r_{01} on the left-hand side of this equation gives the complex amplitude of the field as reflected from the glass-metal interface. The right-hand side represents the complex field amplitude at the glass-metal interface that has passed up and down through the metal film and has been reflected off the metal-air interface. Equation (5.10) means these two reflected waves with equal amplitudes interfere destructively in the direction of the reflected beam, yielding zero reflectance. All the power of the incident beam is coupled into the surface plasmon, which dissipates it away.[17] The critical coupling thickness for the two fictitious metals are shown in Fig. 5.6.

Figure 5.6 shows that in the Kretschmann configuration, critical cou-

[17] Note that the critical coupling condition requires equality of two complex quantities.

pling is not easily lost when changing the wavelength of the incident light for a constant metal layer thickness. The difference in critical thickness between plasmonium and scattrium in Fig. 5.6 reflects the accelerated decay of the field in the lossier metal, and so a thinner layer is required to balance the two reflected fields.[18]

[18] As a reality check, we show that Fig. 5.6 does represent realistic numbers: for silver, a plasmonium-like metal with $\lambda_p = 138$ nm, (5.10) predicts a critical thickness of $0.34 \times 138 = 47$ nm for $\lambda = 1000$ nm ($\Omega = 0.138$). Likewise, for aluminum, a scattrium-like metal with $\lambda_p = 79$ nm, it predicts a critical thickness of $0.2 \times 79 = 14$ nm for $\lambda = 500$ nm ($\Omega = 0.16$).

[19] Raether, 1988, p. 12.

5.4 *Effect of electron scattering rate on the Kretschmann lineshape*

In traditional Kretschmann experiments where the dip in reflectance is very narrow because $|\varepsilon_1'| \gg 1$ and $\varepsilon_1'' \ll |\varepsilon_1'|$,[19] the dip is often approximated by a Lorentzian resonance subtracted from a constant background of unit magnitude. In this limit, the resonance angle gives the real part of the surface plasmon wave vector, and the half-width of the reflectance dip reveals the imaginary part. Repeating the measurement over a range of wavelengths yields the surface plasmon dispersion relation.

However, if $\varepsilon_1'' \ll |\varepsilon_1'|$ does not apply, then the reflectance yields much less information about the surface plasmon wave vector. Figure 5.7 illustrates this point by showing calculated Kretschmann reflectance curves at two different frequencies for a layer of plasmonium, which fulfills the conditions above, and a layer of scattrium, which does not fulfill $\varepsilon_1'' \ll |\varepsilon_1'|$. In the case of scattrium, even though the resonance angle and resonance width vary with the wavelength, it is difficult to say exactly how the resonance width should be defined, since the resonance is highly asymmetric. For example, the linewidths of the two curves in Fig. 5.7b are obviously different, but there is no apparent way to quantify them, since the lineshapes are asymmetric.

In fact, the rule of thumb that holds for plasmonium — that the resonance width yields information about the imaginary part of the surface plasmon wave vector — fails even on a basic intuitive level for scattrium: in Fig. 5.7, the purple curve's linewidth is, if anything, wider than that of the green curve, whereas one would expect it to be narrower because the resonance is more heavily damped at the higher frequency of the green curve, as we see from Fig. 5.3. The discrepancy is caused by the phase difference between the resonance and the background.

As we will show in the next section, the parallel wave vector at the resonance angle, $k_0\sqrt{\varepsilon_0}\sin\theta_{\mathrm{SPR}}$, does not necessarily correspond to the actual surface plasmon wave vector, contrary to what is usually assumed in Kretschmann experiments. In the case of a metal with low ε'' such as plasmonium, the difference is slight; but in scattrium, the actual surface

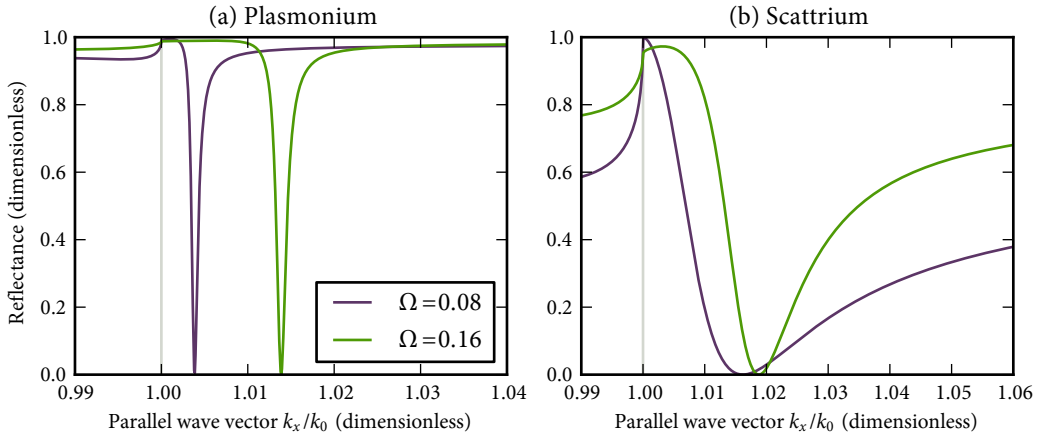


Figure 5.7: Calculated SPR curves for the Kretschmann configuration at two far-apart frequencies. The layers are plasmonium (a) and scattrium (b) on a $n = 1.5$ glass-like substrate. At each frequency, the layer is taken to have the proper thickness for critical coupling. The critical angle for total internal reflection in the substrate, at $k_x = k_0$, is indicated by a gray line. The curves in (a) are typical in SPR experiments. The position of the reflectance minimum corresponds to the real part of the wave vector of the surface plasmon, and its linewidth corresponds to the imaginary part. The scattrium-type curves in (b) are asymmetric without a well-defined linewidth. Their minimum does not correspond to the real part of the surface plasmon wave vector, and their width does not correspond to the imaginary part.

plasmon wave vector in scattrium is quite far removed from the parallel wave vector at the resonance angle. This finding is similar to a damped driven harmonic oscillator, where the damping parameter is related to ε''. It is well-known that a sufficiently damped, driven oscillator has its maximum response at a different frequency from the undamped resonance frequency. In fact, as we will see in the next section, a damped driven oscillator on a coherent background is precisely what describes the surface plasmon resonance.

5.5 *Analyzing Kretschmann lineshapes*

In his original paper, Kretschmann suggested considering the surface plasmon resonance a lightly damped driven oscillator, elegantly described by a Lorentzian lineshape.[20] The reflectance, in the neighborhood of the resonance angle, is then the resonance subtracted from a constant background of unit magnitude:[21]

$$R = 1 - \frac{4k_{\mathrm{SP}}^{\infty\prime\prime}\Delta k_{\mathrm{SP}}''}{(k_x - (k_{\mathrm{SP}}^{\infty\prime} + \Delta k_{\mathrm{SP}}'))^2 + (k_{\mathrm{SP}}^{\infty\prime\prime} + \Delta k_{\mathrm{SP}}'')^2}, \tag{5.11}$$

where k_{SP}^{∞} is the surface plasmon wave vector on the semi-infinite interface, as given by (5.4), and Δk_{SP} is a displacement that the resonance undergoes due to the presence of the coupling prism, approximated by:[22]

$$\Delta k_{\mathrm{SP}} = \frac{2|k_{\mathrm{SP}}^{\infty}|^3}{k_0^2(|\varepsilon_1'| + \varepsilon_2)} e^{2ik_{z1}d_1} r_{01}(k_{\mathrm{SP}}^{\infty}). \tag{5.12}$$

[20] Kretschmann, 1971.
[21] Raether, 1988, p. 12.
[22] Raether, 1988, p. 12.

However, as Fig. 5.7 clearly shows, this approximation does not fit very well to metals that behave like scattrium. In addition, it assumes that the resonance angle is equal to the angle corresponding to the surface plasmon wave vector. Various improvements to this fitting function exist, including ones that drop the latter assumption,[23] but there is little motivation to expand the analysis beyond plasmonium-type metals, since gold is most often used in commercial SPR systems anyway.

[23] Kurihara, Nakamura, and Suzuki, 2002.

HERE WE PRESENT A DESCRIPTION that is valid over a larger range of angles, not just in the neighborhood of the resonance, and can be used to fit metals with larger Drude scattering parameters. We start from the expression in (5.9) and write it as the coherent addition of a resonance to a slowly varying background. In addition, we note that r_{12}^{-1} goes to zero when $k_x = k_{\mathrm{SP}}^{\infty}$ (5.4) (the denominators of r_{12} and t_{12} are the same), so we write the expression as a function of r_{12}^{-1}:

$$r_{012} = \frac{r_{01} + r_{12}\delta}{1 + r_{01}r_{12}\delta} = r_{01} + \frac{(1 - r_{01})^2 r_{12}\delta}{1 + r_{01}r_{12}\delta} = r_{01} + \frac{(1 - r_{01})^2\delta}{r_{12}^{-1} + r_{01}\delta} \tag{5.13}$$

Then we take a linear approximation of r_{12}^{-1} around the zero at k_{SP}^{∞}:

$$r_{12}^{-1} \approx \alpha(k_x - k_{\mathrm{SP}}^{\infty}), \qquad \alpha = \frac{\partial}{\partial k_x} r_{12}^{-1}\bigg|_{k_x = k_{\mathrm{SP}}^{\infty}} \tag{5.14}$$

with α a complex-valued constant. So far, this is the same approach by which Kretschmann derived the Lorentzian resonance. However, instead of taking unit background and resonance amplitudes, we make no more approximations, instead writing the expression as follows:

$$R(k_x) = \left| B + \frac{A e^{i\phi} k''_{\mathrm{SP}}}{k'_{\mathrm{SP}} + i k''_{\mathrm{SP}} - k_x} \right|^2, k_x > k_{cr}. \tag{5.15}$$

We neglect the part of the reflectance curve under the critical angle, since the linear approximation breaks down at that point.

We can use this expression for extracting the surface plasmon wave vector from SPR curves. There are five fit parameters in the expression: B, the background amplitude; A, the resonance amplitude; ϕ, the phase difference between the background and resonance; and k'_{SP} and k''_{SP}, the complex surface plasmon wave vector. The advantage of this expression is that it yields a surface plasmon wave vector without requiring any advance knowledge of the composition or thicknesses of the layer system: it is completely phenomenological.

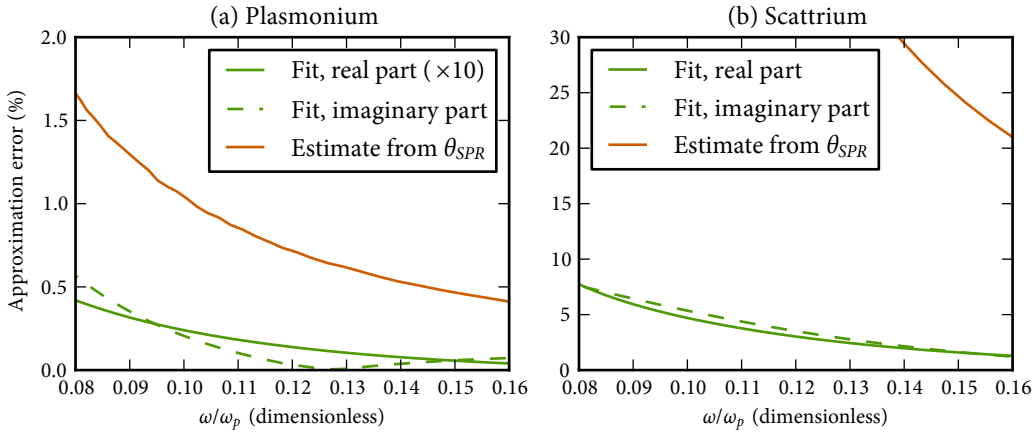


Figure 5.8: Plot of the relative error in the value of n_{SP} (green curves) made by using the value obtained from fitting the numerically calculated reflectance curves with (5.15). The orange curves are the result of simply estimating n'_{SP} from the resonance angle, which is a fairly good approximation for plasmonium (a), but not at all for scattrium (b). The error in the real part in (a) has been multiplied by 10 to improve visibility.

Figure 5.8 shows how effective the phenomenological fit is, compared to estimating the surface plasmon mode index from the resonance angle. On the vertical axis, we plot the approximation error, i.e. the deviation between the calculated and estimated mode index. For low-loss metals like plasmonium, the phenomenological fit proves excellent (error < 0.1%), but it is acceptable to use the resonance angle (error < 2%). For scattrium-like metals, the resonance angle is quite far off, whereas the phenomenological fit performs reasonably well.

As a comprehensive acid test for the phenomenological fitting procedure, we calculate SPR curves for a Kretschmann configuration experiment for a wavelength range from 500 to 800 nm. As a substrate we take BK7 glass with Sellmeier dispersion;[24] as metal we take a 40 nm layer of gold, the optical properties of which we approximate with a Drude model with added Lorentzian oscillators, fit to published values.[25] Gold is a plasmonium-like metal; however, around 500 nm, it has an interband absorption which increases the loss so that it enters a more scattrium-like regime. Therefore, this wavelength range nicely tests both symmetric and asymmetric SPR curves. We add an extra capping layer of 5 nm aluminum oxide with Sellmeier dispersion[26] to the calculations, in order to illustrate how the fitting procedure performs with more than three layers. Finally, we add Gaussian noise with a standard deviation of 1% to the signal. We then treat these data as measured results and fit them with (5.15). We show the results in Fig. 5.9.

In Fig. 5.9b, we see that the phenomenological fitting expression per-

[24] Schott AG, 2012.

[25] Rakić, Djurišić, Elazar, and Majewski, 1998.

[26] Babeva, Kitova, Mednikarov, and Konstantinov, 2002.

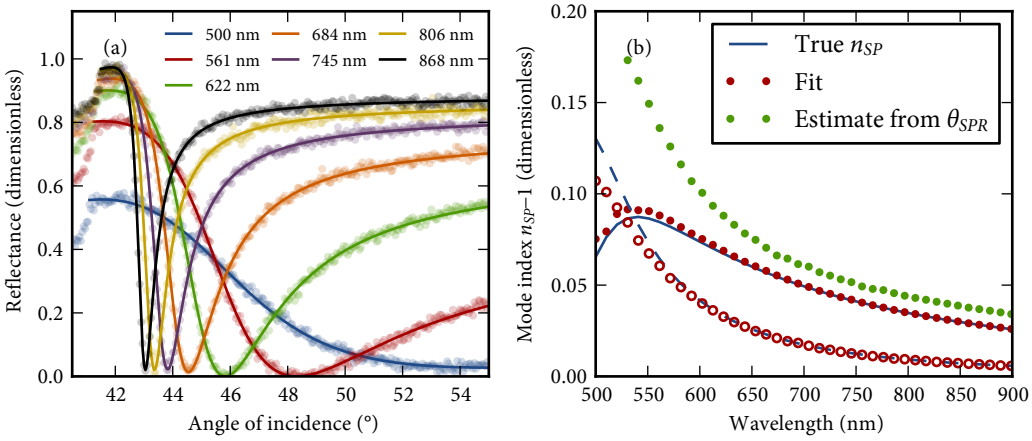


Figure 5.9: Performance of the phenomenological fitting expression of (5.15) on realistic, yet fictitious, SPR "measurements," calculated for a gold layer capped with Al_2O_3 on a BK7 substrate. (a) Generated fictitious data points for selected wavelengths, along with the corresponding fit as a solid line. Note the transition from a scattrium-like regime to a plasmonium-like regime as the wavelength increases. (b) Comparison of the calculated surface plasmon mode index n_{SP} (blue lines; real part solid, imaginary part dashed) to that obtained from the fitting procedure (red dots; real part closed, imaginary part open). For comparison, the green dots are the real part of the mode index estimated from the resonance angle.

forms admirably, much better in any case than estimating the real part of the mode index from the resonance angle. As the metal gets lossier and more scattrium-like, the fit gets slightly worse. It has a tendency to underestimate the imaginary part of the mode index, but that is not entirely surprising since the asymmetric dip is much wider than the "measured" angle range at 500 nm, as we see from the blue curve in Fig. 5.9a.

5.6 *Otto configuration*

THE OTTO CONFIGURATION (Fig. 5.4b) can be described by the same mathematics as the Kretschmann configuration. The only difference is that the thin layer ε_1 is a low-index dielectric and the metal ε_2 is on the outside. Because of this similarity, the Otto configuration is often considered equivalent to the Kretschmann configuration; but a common misconception is that there is practical difficulty in realizing it experimentally and it is therefore unattractive.

It is true that the original experimental realization of the Otto configuration, with air as the low-index dielectric, involves bringing the metal within a few microns of the prism and maintaining a constant air gap width over the entire surface, which was, and is even now, notoriously difficult to accomplish. For example, contamination by one or more dust particles of 75 μm[27] would make a one-micron air gap completely impossible. However, there is no reason why the low-index dielectric has to be air. For example, in chapters 6 and 7, we describe Otto experiments using

[27] ISO 4225:1994.

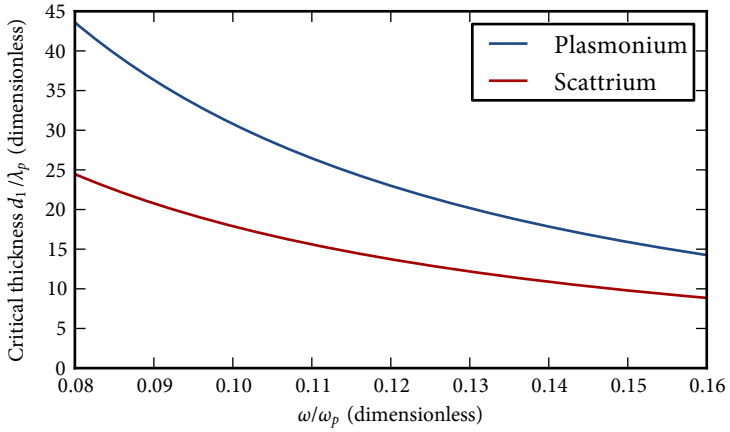


Figure 5.10: Free-space wavelength dependence of the gap thickness necessary for critical coupling in a three-layer Otto configuration of a glass substrate with $n = 1.5$ separated from bulk plasmonium or scattrium by a vacuum gap. Compare Fig. 5.6; the gap thickness is several orders of magnitude larger in the Otto configuration than the metal layer thickness in the Kretschmann configuration.

a high-index flint glass prism and magnesium fluoride as the low-index dielectric, where no gaps or moving parts are involved.

In this section, we will explore in which circumstances the Otto configuration is more appropriate for SPR measurements than the Kretschmann configuration. We will take the high-index dielectric to be a glass-like substance ($n = 1.5$) as before, the low-index dielectric to be vacuum, and the metal to be a bulk layer of either plasmonium or scattrium.

FIRST OF ALL, we use (5.10) to calculate the critical coupling thickness for the vacuum gap between the glass and the metal, shown in Fig. 5.10. In the Otto configuration, the middle layer must be several orders of magnitude thicker than the middle layer in the Kretschmann configuration in order to achieve critical coupling. This is because the surface plasmon's radiative losses must be equal to its damping losses at critical coupling, as we previously explained. A vacuum gap is lossless compared to a metal layer, and so the evanescent wave in the middle layer decays over a much larger distance in the vacuum gap than it does in the metal layer. If it has not decayed enough before bridging the gap, then the system is overcoupled. In addition, there is true total internal reflection at the glass-vacuum interface in the Otto configuration, whereas the wave in the metal layer in the Kretschmann configuration is not purely evanescent.

This is why the required layer thickness is many times that shown in Fig. 5.6; in order to balance the damping and reradiation losses, the field must cross a much larger distance compared to the Kretschmann case. Also, unlike the Kretschmann case, the critical coupling thickness now

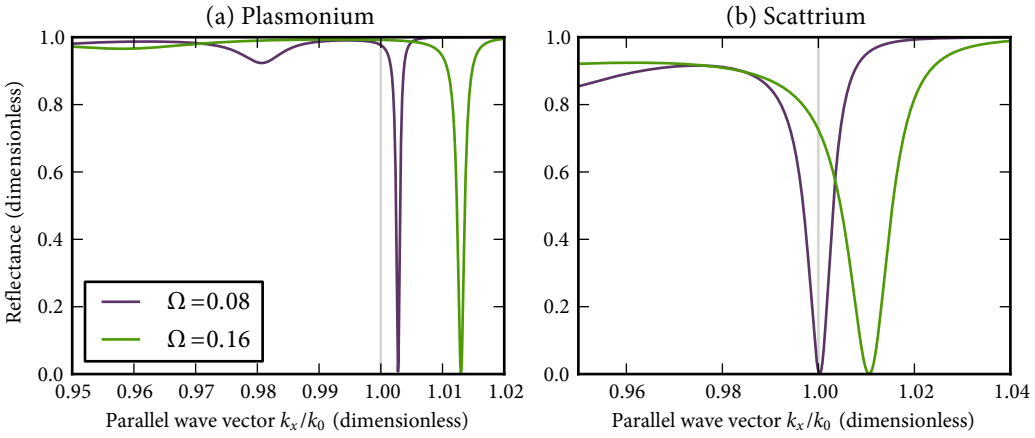


Figure 5.11: Calculated SPR curves for the Otto configuration at two far-apart frequencies. The outer layers are plasmonium (a) and scattrium (b), with a vacuum gap of critical coupling thickness (see Fig. 5.10 to read off the thickness) and an $n = 1.5$ glass-like substrate. The critical angle for total internal reflection in the substrate is indicated by a gray line. Compare Fig. 5.7.

exhibits a strong frequency dependence. The vacuum layer thickness d_1 scales approximately with i/k_{z1} (5.10); k_{z1}/k_0 is a small imaginary number, so $d_1 \propto 1/\omega$. This means that when designing an Otto experiment for a broad range of wavelengths, a middle layer of constant thickness will cause the resonance to become undercoupled or overcoupled much closer to the design wavelength than in the Kretschmann configuration.

WE NOW EXAMINE the wavelength-dependent SPR curves for the Otto system. We show a number of examples in Fig. 5.11. There are some notable differences from the equivalent SPR curves for the Kretschmann system (Compare Fig. 5.7.)

For both plasmonium and scattrium, nothing special happens at the critical angle, unlike the Kretschmann case. This is because the Otto configuration deals with true frustrated total internal reflection, where the incident wave, which is evanescent in the gap, can still excite a propagating wave in the metal for some angles. The Kretschmann configuration, on the other hand, has the vacuum on the outside, so whether the light couples into the surface plasmon mode or not, it cannot travel into the vacuum in any case; the total internal reflection is not frustrated, only perturbed.

In addition, secondary resonances are visible at lower angles ($k_x < k_0$) than the main plasmonic resonances. Calculating the mode profile[28] shows that these are waveguide modes in the vacuum gap, as suggested by Tillin and Sambles.[29] The Kretschmann configuration's metal layer is

[28] Davis, 2009.

[29] Tillin and Sambles, 1988.

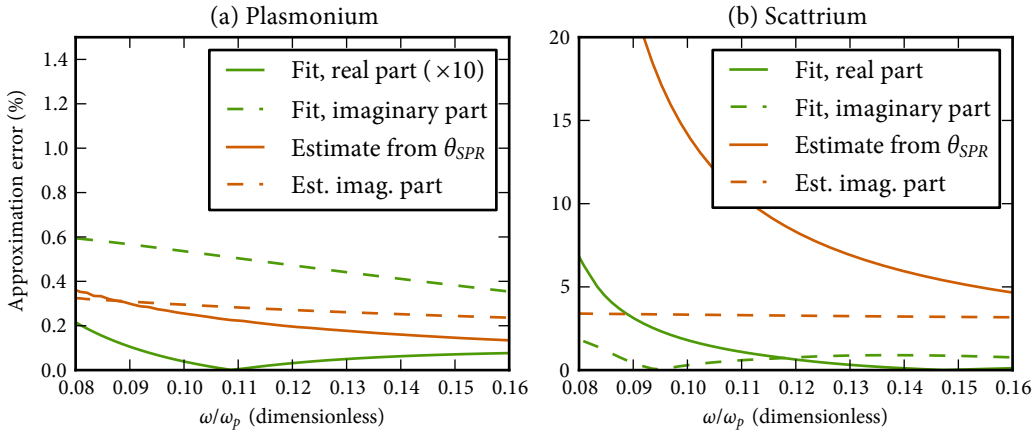


Figure 5.12: Plot of the relative error in the value of n_{SP} (green curves) made by using the value obtained from fitting the numerically calculated reflectance curves with (5.15). The orange curves are the result of simply estimating n'_{SP} from the resonance angle and n''_{SP} from the half-width, which is a good approximation for plasmonium (a), but not for scattrium (b). The error in the real part in (a) has been multiplied by 10 to improve visibility.

not thick enough to support such modes.

In both cases, the resonance lineshape is approximately Lorentzian and easy to interpret. Using the phenomenological fitting expression of (5.15), shown in Fig. 5.12, on the fictitious measurements of Fig. 5.11 shows that the analysis works well for both metals, performing comparably to the Kretschmann configuration. The resonance angle (making sure to take the resonance corresponding to the plasmon mode and not a waveguide mode) is a good indicator of the real part of the surface plasmon mode index for plasmonium, but not at all for scattrium. Again, this is because the approximation of (5.14) breaks down close to the critical angle.

Since the curves in Fig. 5.11 all have a reasonably well-defined linewidth, we can also estimate the imaginary part from the resonance's half-width at half-maximum. For scattrium, this yields reasonable results, but for plasmonium, this estimate is even slightly better than using (5.15). This indicates that, at least for the imaginary part of the mode index, the Otto configuration produces much more easily interpretable experimental results than the Kretschmann configuration when studying surface plasmons on a metal with large ε''.

5.7 *Conclusion*

We have discussed the influence of the optical properties of metals on the resonance lineshape in SPR measurements and examined the advantages and disadvantages of the Kretschmann and Otto configurations for SPR experiments.

We have demonstrated that there are advantages to the Otto configuration as a method of studying surface plasmons, contrary to what is often thought. In the case of a low-loss metal, such as our fictitious "plasmonium," it performs comparably to the Kretschmann configuration, although the Kretschmann configuration may be preferable if working with a large range of wavelengths. When working with high-loss metals such as our fictitious "scattrium," the resonance angle yields no information about the real part of the surface plasmon mode index in either configuration. However, the linewidth of an Otto curve is always a good indicator of the imaginary part of the mode index. Kretschmann curves, on the other hand, can be asymmetric for high-loss metals, in which case they do not have a well-defined linewidth.

In addition, the Otto configuration allows the use of arbitrarily thick layers of metal. This is important because a scattrium-type metal film must be very thin if used in the Kretschmann configuration: so thin, in fact, that the thickness is of the same order as the electron scattering length, possibly affecting the optical properties of the film.

We have also demonstrated a method for analyzing SPR curves that allows extraction of the complex surface plasmon mode index without any knowledge of the composition, thicknesses, or optical properties of the various layers of metal and dielectric involved in the attenuated total reflection coupling system. This phenomenological method of analysis yields values for the imaginary part of the mode index even when confronted with an asymmetric Kretschmann lineshape. It also yields more accurate values for the real part than can be obtained from the resonance angle, and works well for metals with large or small ε''.

Appendix 5.A Surface plasmon resonance curves beyond critical coupling

In this chapter, we have examined the critically coupled case of SPR, by adjusting the thickness of the middle layer (metal in the Kretschmann configuration, air in the Otto configuration) depending on the frequency of the incident light. This is not feasible in a real experiment, unless one is willing to deal with a layer of adjustable thickness, such as an oil layer whose thickness is changed by adjusting the mechanical pressure on the two surrounding solids;[30] even then, it is impossible to measure a broad range of wavelengths all at once.

[30] Quail, Rako, and Simon, 1983.

Therefore, in this appendix we evaluate a more realistic experimental situation for our fictitious metals plasmonium and scattrium. We take the plasma frequency of both metals to be $\hbar\omega_p = 15$ eV, allowing us to put in actual wavelengths; and we take a layer thickness appropriate to a wavelength of 800 nm ($\Omega = 0.10$). That is, in the Kretschmann configuration, 24.3 nm for plasmonium and 10.0 nm for scattrium; and in the Otto configuration, 2.4 μm for plasmonium and 1.4 μm for scattrium. We "measure" at six wavelengths, from 500 to 1000 nm.

Figure 5.13 shows the reflectance curves for the Kretschmann configuration. They are not so different from the curves in Fig. 5.7, bearing out our assertion that the coupling efficiency does not have a very strong dependency on the frequency in the Kretschmann configuration. It is interesting to note that the intuitive rule of thumb, that the resonance linewidth is related to the imaginary part of the surface plasmon wave vector, now seems to apply to scattrium; the blue ($\lambda = 500$ nm) resonance is broader in both plasmonium and scattrium, and this corresponds to more absorption in the metals at higher frequencies.

Figure 5.14, on the other hand, shows the reflectance curves for the Otto configuration. The secondary waveguide modes are once again visible at angles less than the critical angle, but this time they are closer to the critical angle. It is also immediately apparent that the coupling is much worse when the layer thickness is not optimized for critical coupling. The coupling to scattrium's surface plasmon mode at $k_x = 1.01k_0$ for $\lambda = 500$ nm is much weaker than the coupling to the waveguide mode at $k_x = 0.97k_0$. In plasmonium, the surface plasmon mode has all but disappeared at 500 nm, visible as a tiny blip at $k_x = 1.015k_0$.

Note that the main resonance in scattrium is not necessarily plasmonic at all wavelengths; it appears to the left of the critical angle, i.e. $k_{\text{mode}} < k_0$, for $\lambda = 900$ and 1000 nm. It is a hybrid between a surface plasmon and

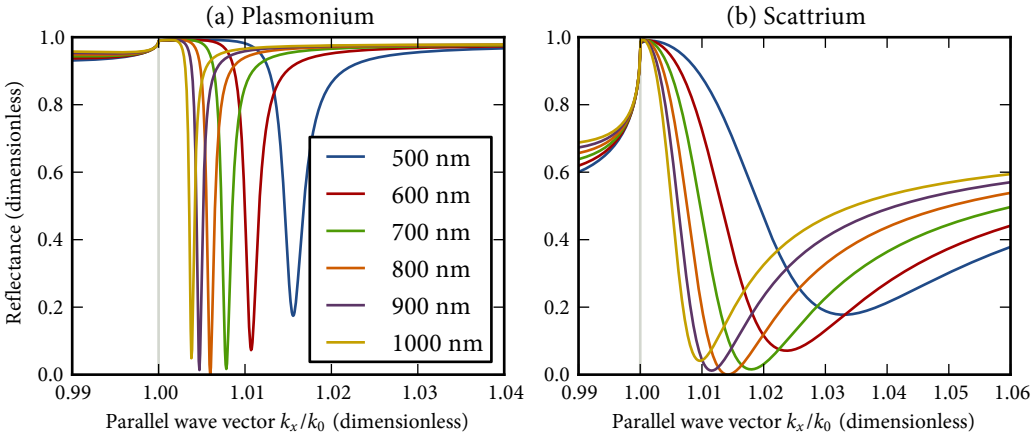


Figure 5.13: SPR curves for a number of representative wavelengths, for a Kretschmann experiment with (a) a 24.3 nm layer of plasmonium on a glass-like substrate with $n = 1.5$ surrounded by vacuum; (b) a 10.0 nm layer of scattrium on the same substrate, also in vacuum. The metal layers' thicknesses are designed for critical coupling at $\lambda = 800$ nm, but not for other wavelengths. Compare to Fig. 5.7.

waveguide mode, being strongly damped in the intermediate layer, but not entirely evanescent away from the interface. This emphasizes the modal character of the Otto plasmon; when the layer width is different from the critical layer width for surface plasmon mode, the effective index shifts.

In all cases, even for resonances away from critical coupling, the phenomenological expression of (5.15) still yields accurate results. The simpler method of extracting the real part of the surface plasmon wave vector from the resonance angle is a reasonable approximation in the Otto configuration, whereas it does not work in Kretschmann configuration if the lineshapes are asymmetric, as is the case with scattrium.

The conclusion stands, that the Otto configuration is a better experimental technique for probing surface plasmons on lossy scattrium-like metals. However, we note that if a broad enough wavelength range is required, then a constant gap width degrades the coupling to the surface plasmon mode, and in that case it is better to deal with the Kretschmann configuration's asymmetric lineshapes.

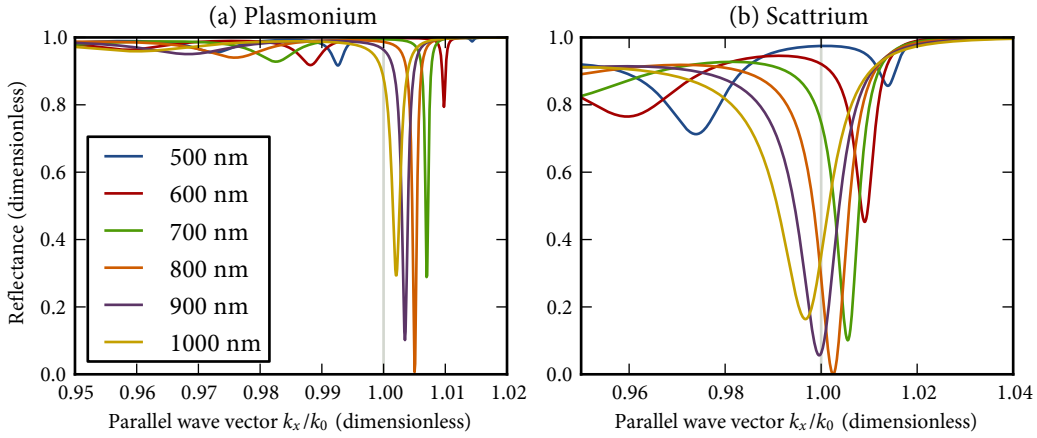


Figure 5.14: SPR curves for a number of representative wavelengths, for an Otto experiment with (a) plasmonium separated from a $n = 1.5$ glass-like substrate by a vacuum gap of 2.4 μm; (b) scattrium separated from the same substrate, by a vacuum gap of 1.4 μm. The gaps' thicknesses are designed for critical coupling at $\lambda = 800$ nm, but not for other wavelengths. Compare to Fig. 5.11.

6
Anomalous dispersion of surface plasmons

We demonstrate, using surface plasmon resonance experiments in the Kretschmann and Otto configurations, a region of anomalous dispersion in the effective mode index of surface plasmons on aluminum in the near-infrared. This phenomenon is a consequence of aluminum's parallel-band transition at 1.5 eV. Our results show that the transition is only weakly present in aluminum layers of the order of 10 nm.

6.1 Introduction

Radiative and dissipative losses play an important role in the field of plasmonics. While radiative loss is commonly considered useful, providing the coupling mechanism to the outside world, dissipative loss does not have this positive side. Apart from the drive towards lossless plasmonics,[1] the usual response to this challenge involves the use of materials, in this case metals, that have minimal dissipation.[2] Often, the metals that have low loss at optical frequencies have dielectric properties that exhibit Drude-like dispersion, that is, the dielectric properties are dominated by intraband electronic transitions. The Drude-like behavior implies that the dispersion of the effective surface plasmon mode index n_{SP} is always normal, i.e.:

[1] Berini and De Leon, 2012.

[2] West et al., 2010.

$$\frac{dn_{\mathrm{SP}}}{d\lambda} < 0 \text{ and } \frac{d^2 n_{\mathrm{SP}}}{d\lambda^2} > 0. \tag{6.1}$$

Anomalous dispersion in bulk materials such as atomic vapors is usually associated with a resonance in the absorption; there the vapor is essentially opaque so that the dispersion is usually difficult to measure directly.[3] A discussion of the dispersive properties of a surface plasmon is

[3] King, 1917.

slightly more involved because a surface plasmon is a mode and therefore it is the *modal dispersion* that we probe. Moreover, in a simple geometry, it is a mode that propagates along the interface between two materials and therefore it probes the dispersion of *both* materials. In quite a few experiments the surface plasmon field extends into a number of layers with different dielectric properties, and the anomalous dispersion can arise as a consequence of the dispersive properties of each material separately, or combined. For example, the dispersion relation of the surface plasmon can be influenced by interaction with a dye monolayer.[4] Here we address the simple case that the anomalous dispersion of a surface plasmon is caused by the dispersive properties of the metal, and not of the other materials into which it extends. Even in this simple case, the answer will turn out to depend on the experimental approach chosen.

[4] Wähling, 1981.

Following the reasoning of the previous paragraph, we are searching for a plasmonic material — i.e. a metal where surface plasmons are not damped too strongly — which exhibits a relatively narrow peak in the imaginary part of the dielectric constant, and thus in the bulk absorption. Actually, there are quite a few metals that exhibit a resonance of this kind, such as aluminum,[5] magnesium[6] and calcium,[7] to name just a few. For these metals, the spectral band associated with the extra absorption is known as parallel-band absorption:[8] the absorption arises as a consequence of an interband transition between parallel electronic bands near the Fermi energy. From an experimental point of view, aluminum is attractive since it is a metal that has wide applications in both optics and plasmonics. Moreover, compared to the other materials mentioned above, it is quite easily handled.

[5] Ehrenreich, Philipp, and Segall, 1963.
[6] Mathewson and Myers, 1973.
[7] Hunderi, 1976.
[8] Harrison, 1966.

The parallel-band absorption in aluminum is quite well known among optics experts as it is the effect underlying the small dip at $\lambda \approx 800$ nm ($\hbar\omega \approx 1.5$ eV) in the reflectance of aluminum, and thus of all aluminum mirrors.[9] Ehrenreich et al.[10] were first in explaining this dip as the result of a parallel-band transition. Their analysis of the optical properties of aluminum was extended and refined by Harrison[11] and Brust.[12] Almost a decade later, their model of the optical properties of aluminum, as a weakly perturbed free-electron metal,[13] was further refined by Ashcroft and Sturm[14] resulting in explicit expressions for the optical conductivity, which were shown to agree well with experimental data.

[9] Strong, 1936; Bennett et al., 1963.
[10] Ehrenreich et al., 1963.
[11] Harrison, 1966.
[12] Brust, 1970.
[13] Harrison, 1966.
[14] Ashcroft and Sturm, 1971.

In this chapter, we investigate the optical dispersion of surface plasmons traveling on a metallic aluminum interface in the spectral region of parallel-band absorption, using attenuated total reflection in the Kretschmann and Otto configurations.

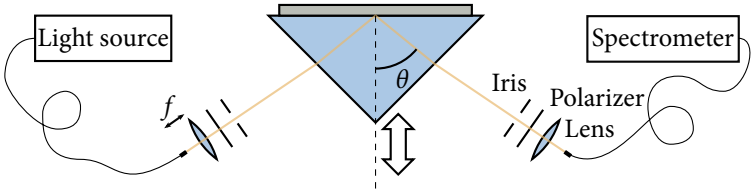


Figure 6.1: Sketch of the experimental setup. The setup is placed on a pair of rotation stages so that for a certain desired internal angle of incidence θ, the light source and detector are rotated to the correct angles. The prism is also placed on a translation stage aligned with the prism's axis of symmetry, so that the setup remains symmetric and the beam always probes the same spot on the coated face.

6.2 *Experiment*

THE LOCATION OF THE PARALLEL-BAND FEATURE in the spectrum suggests that in our search for plasmonic anomalous dispersion we need to study the spectral region between roughly 600 and 1000 nm. For the experiment this implies the use of a broadband light source such as a lamp, which, typically, has low spectral brightness. For this reason we choose to couple with the surface plasmon by traditional attenuated total reflection methods, as originally proposed by Kretschmann and Otto. The method put forward by Kretschmann[15] is almost universally considered preferable to that of Otto;[16] but in the present case the Otto configuration has some very strong advantages, as discussed in chapter 5. It was successfully used by Tillin and Sambles[17] to measure the dielectric function of aluminum. Here we will present experimental results for both configurations and we will see that the results are quite different, reflecting the fact that the two distinct experiments probe an entirely different surface plasmon mode.

[15] Kretschmann, 1971.

[16] Otto, 1968.

[17] Tillin and Sambles, 1988.

Figure 6.1 shows the generic experimental setup. Highly collimated *p*-polarized white light from a fiber-coupled tungsten halogen lamp (Ocean Optics HL-2000-HP-FHSA) is incident on one of the faces of a prism; the collimated beam that is internally reflected off the metal-covered face leaves the prism through its third face and is collected on the input facet of a fiber-coupled mini-spectrometer (Ocean Optics USB2000). During the experiment we rotate the prism in steps that increase the internal angle of incidence by 0.05°; the mini-spectrometer with the associated optics is mounted on an arm that co-rotates with the prism so that it catches the reflected beam. In order to ensure that we always probe the same spot on the aluminum layer, we adjust the prism's position after each rotation step. For each orientation of the prism we measure the reflected spectrum.

FOR THE KRETSCHMANN EXPERIMENT, we used a right-angled BK7 glass prism, with broadband near-infrared antireflection coatings (650–950 nm) on the two faces adjacent to the right angle. On the hypotenuse we deposited a thin layer of aluminum, using a Leybold Heraeus Z400

sputtering system. To prevent oxidation of the metal film, we used the same sputtering system to immediately cap it with a Si_3N_4 layer of a few nanometers thickness. The critical angle for total internal reflection on the BK7–air interface is around 41.5°, slightly less than 45°, meaning that the beam is almost normally incident on the antireflection-coated entrance and exit faces.

In the Kretschmann configuration, the surface plasmon that can be excited by the attenuated total reflection technique resides on the Si_3N_4-capped outer surface of the aluminum film. Because of the high optical loss in bulk aluminum, the calculated layer thickness that yields critical coupling is approximately 6 to 10 nm in the spectral region we are studying.

However, thin aluminum films have optical properties quite different from those of bulk aluminum.[18] In the measurements of Novotny et al.,[19] the parallel band resonance is not even in evidence for layers thinner than 9 nm, supposedly because the electron scattering length is then of the same order of magnitude as the layer thickness. In order to find the best layer thickness for critical coupling, we performed experiments on aluminum layers of several different thicknesses. The results here are shown for a prism coated with a 10.8 ± 0.6 nm thick aluminum layer with 7.6 ± 0.3 nm Si_3N_4 on top. We illustrate the layer stack on the prism and the plasmonic mode we probe in Fig. 6.2. Figure 6.3 shows experimental results for this prism for a number of representative wavelengths.

The most striking feature of the experimental results is that the reflectance is relatively small for a wide range of angles of incidence and seems not to recover at large angles, quite in contrast to the well-known case of surface plasmons on gold or silver films. As discussed in the previous chapter, this is typical for a metal where the ratio of the imaginary and real parts of the dielectric constant $|\varepsilon''/\varepsilon'|$ is of order 1.

[18] Du et al., 2006; Novotny, Bulir, Lancok, Pokorny, and Bodnar, 2011.

[19] Novotny, Bulir, et al., 2011.

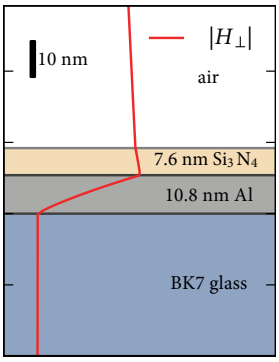


Figure 6.2: Schematic of the various layers deposited on the BK7 prism used in the Kretschmann setup. The calculated amplitude of the magnetic component of the surface plasmon field $H_\perp$ is shown in red. The left edge of the figure corresponds to zero amplitude. The field mode decays exponentially into the glass and air with decay lengths of 9.1 μm and 0.76 μm, respectively. Hence, the decay is not visible at this scale.

FOR THE OTTO EXPERIMENT, we used an equilateral prism made of F2 glass ($n \approx 1.608$ at 800 nm), again with coatings deposited on all three faces: broadband near-infrared antireflection coatings ($R < 2\%$ for 500–940 nm) on the two faces through which light enters and exits the prism; and a compound layer on the third side consisting of 570 nm MgF_2 (the low-index dielectric), 100 nm aluminum, and 110 nm SiO_2 as a capping layer to prevent oxidation. For this prism, the critical angle for total internal reflection at the F2–MgF_2 interface is around 59°, depending on wavelength. Again, the light beam is almost normally incident on this prism's antireflection-coated faces.

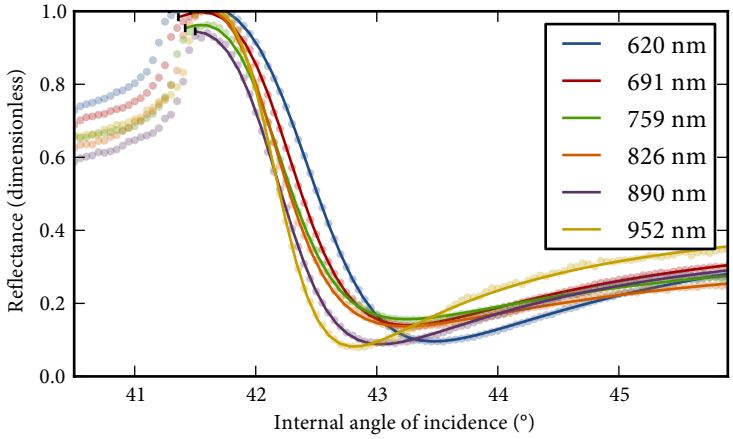


Figure 6.3: Experimental results for the angle-dependent reflectance for several representative wavelengths, measured using the Kretschmann configuration and the sample illustrated in Fig. 6.2. The dots indicate measured values, and the solid lines are fits to the data using (6.2), discussed later in this chapter. The fits cut off at the critical angle for total internal reflection from BK7 glass to air, as indicated by the small black markers.

For our experiments we chose a MgF_2 spacer layer of 570 nm thickness. Theoretically, we expect this layer to provide critical coupling in the 750–850 nm wavelength range. Note that in this setup we excite the surface plasmon at the interface between the aluminum film and the MgF_2 layer. In contrast to the case of the Kretschmann configuration, the metallic layer is sufficiently thick that it may be considered to be bulk aluminium.

Experimental results for the reflectance as a function of internal angle for the Otto setup are shown in Fig. 6.5, for a number of representative wavelengths across the wavelength range of interest. The reflectance curves are narrower than in the Kretschmann configuration, but for the longer wavelengths the coupling is far from critical.

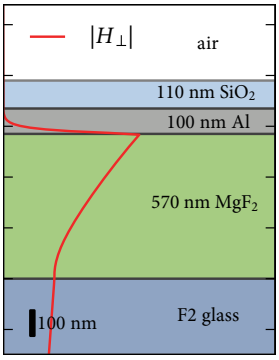


Figure 6.4: Illustration of the various layers deposited on the F2 prism used in the Otto setup. The calculated amplitude of the magnetic field $H_\perp$ associated with the surface plasmon mode is shown in red, in order to illustrate the plasmonic mode that we are probing in this experiment. The left edge of the figure corresponds to zero amplitude.

6.3 *Results and interpretation*

FROM THE EXPERIMENTAL RESULTS, as shown in Figs. 6.3 and 6.5, we extract the resonance angle, i.e. angular position of the minimum of the reflectance curves. We plot the resonance angle for both configurations as a function of wavelength in Fig. 6.6.

The first thing to notice is that the two curves are quite similar qualitatively. In both cases, the resonance angle decreases as a function of wavelength at the shorter and longer wavelengths measured, while at intermediate wavelengths the resonance angle goes through a local maximum. In the Kretschmann configuration, the maximum is less pronounced than in the Otto configuration. More careful inspection of the data brings

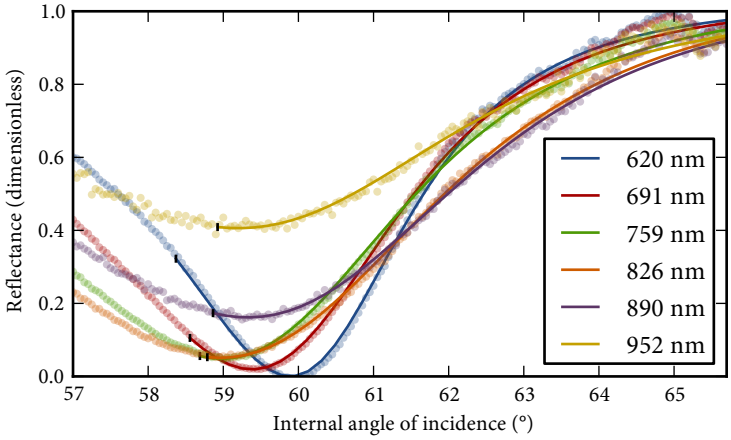


Figure 6.5: Example reflectance curves for several representative wavelengths, measured using the Otto configuration and the sample illustrated in Fig. 6.4. The dots indicate measured values, and the solid lines are fits to the data using (6.2), discussed later in this chapter. The fits cut off at the critical angle for total internal reflection from F2 glass to MgF_2, as indicated by the small black markers.

out the differences: in the Kretschmann setup the maximum lies somewhere between 800 and 850 nm, while the Otto data exhibit a maximum near 900 nm. Note that the Kretschmann data are noisier than the Otto data; this reflects the very broad minimum of the experimental reflectance curves for the Kretschmann configuration, the exact position of which is somewhat difficult to determine, in particular when the signal to noise ratio is not good, as is to be expected in a reflectance minimum. The Otto curves exhibit a much narrower dip, the center of which is therefore easier to determine. At longer wavelengths ($\lambda > 900$ nm) the angular reflectance spectra in the Otto configuration also exhibit a shallow dip, as our experimental system is far from critical coupling there. The difficulty in finding the minimum of these shallow curves is reflected in the noise in Fig. 6.5.

IT IS TEMPTING to draw the conclusion from the data of Fig. 6.6 that the surface plasmon modes in the Kretschmann and Otto configurations exhibit anomalous dispersion. However, the resonance angle is not a good measure for the surface plasmon wave vector, as discussed in chapter 5. In order to analyze the data properly, obtaining both the real and imaginary parts of the surface plasmon mode index, we fit the reflectance curve above the critical angle with the Fano-type lineshape (5.15, 6.2), for each wavelength studied. As shown in chapter 5, this lineshape is an excellent description of the surface plasmon resonance curves for angles of incidence above the critical angle.

$$R(k_x) = \left| B + \frac{A e^{i\phi} k''_{SP}}{k'_{SP} + i k''_{SP} - k_x} \right|^2 , k_x > k_{cr}. \tag{6.2}$$

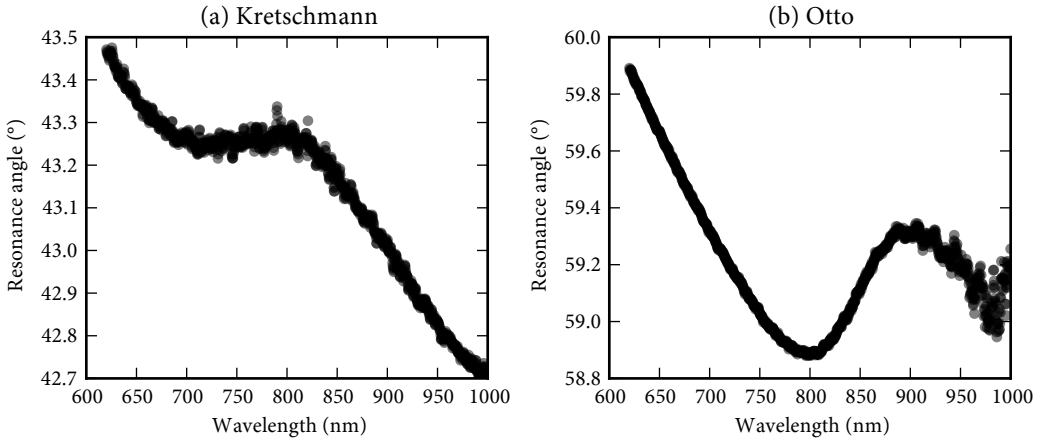


Figure 6.6: Measured resonance angle as a function of wavelength for the Kretschmann experiment (a) and the Otto experiment (b).

Here, R is the reflectance; A, B, ϕ, k'_{SP}, and k''_{SP} are fit parameters. The parameter B can be identified with the modulus of the reflection coefficient r_{01} at the first interface. Since we are only fitting the part of the curve above the critical angle, we can set $B = 1$ for the Otto configuration. From the complex-valued k_{SP} obtained in this way, we plot the real and imaginary parts of the surface plasmon mode index $n_{SP} = k_{SP}/k_0$ in Fig. 6.7.

For comparison, we show calculated curves for the respective layer systems using Sellmeier models for the dielectric functions of BK7 glass, F2 glass,[20] MgF_2,[21] SiO_2,[22] and Si_3N_4.[23] We used Ashcroft and Sturm's model for the dielectric function[24] of the thin layer of aluminum in the Kretschmann experiment, with modified values for the electron scattering times, as discussed in chapter 7. We measured the dielectric function of the thin aluminum layer by ellipsometry and fit Ashcroft and Sturm's expression to these data, with the electron scattering times as variable parameters, yielding 9.7 fs for free electron scattering and 1.9 fs for parallel-band scattering. For the thicker aluminum used in the Otto experiment, we used a Drude-Lorentz model for the dielectric function[25] based on tabulated values.[26] These estimates exhibit good qualitative agreement with the behavior shown in Fig. 6.7.

The Ashcroft-Sturm expression reflects the parallel-band resonance's much heavier damping in the thinner layer; the electron scattering time is greatly reduced because the scattering length is of the same order as the layer thickness. Based on a Fermi velocity of 2.03×10^6 m/s,[27] a parallel-band scattering time of 3.8 fs in bulk aluminum[28] yields a mean free path of 7.7 nm, which is close enough to the layer thickness of 10.8 nm that

[20] Schott AG, 2012.
[21] Dodge, 1984.
[22] Malitson, 1965.
[23] Bååk, 1982.
[24] Ashcroft and Sturm, 1971.
[25] Rakić et al., 1998.
[26] Smith, Shiles, and Inokuti, 1985.
[27] Ashcroft and Mermin, 1976, p. 38.
[28] Mathewson and Myers, 1972.

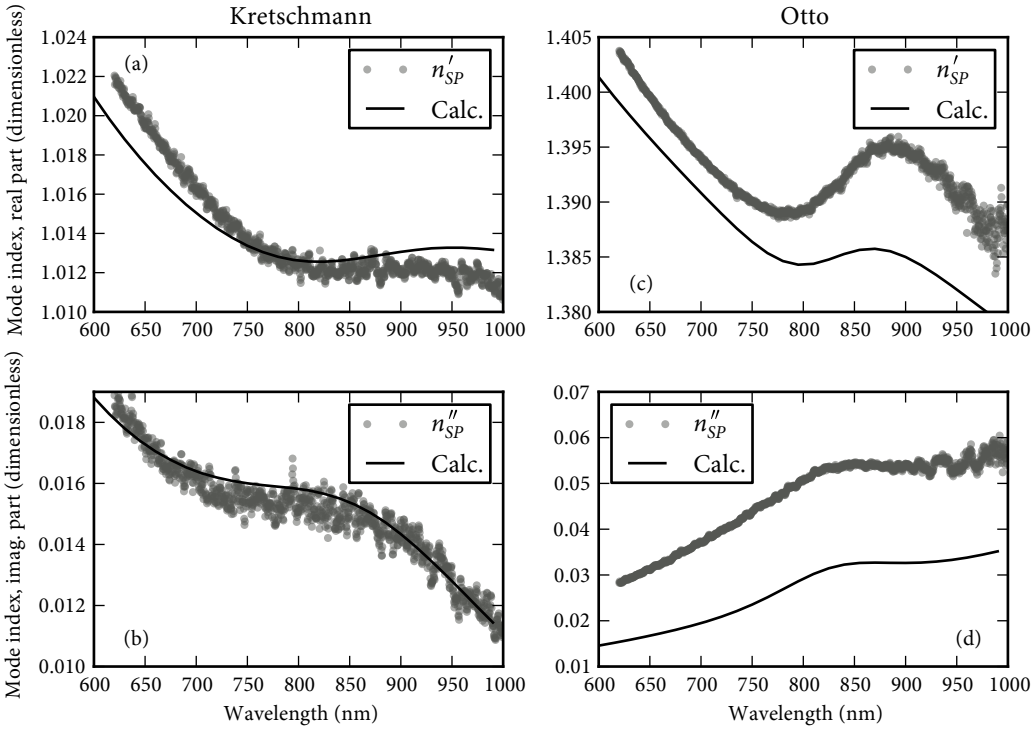


Figure 6.7: Real and imaginary parts of the surface plasmon mode index $n_{SP} = k_{SP}/k_0$ obtained from our measurements on aluminum films in the Kretschmann (a, b) and Otto (c, d) configurations, as a function of free-space wavelength. In (a) and (b), we also show calculated values for the mode index in the Kretschmann layer system, using expressions for the dielectric constant of aluminum as given by Ashcroft and Sturm (1971). In (c) and (d), we show the calculated mode index in a similar manner, but using published values for the dielectric function of bulk aluminum metal.

surface scattering effects play more of a part. Our fitted value of 1.9 fs for the parallel-band scattering time, on the other hand, corresponds to a mean free path of 3.8 nm. In Figs. 6.7c–d, we compare the measured results to published values for bulk aluminum; however, in chapter 7 we will also demonstrate an improved correspondence by adjusting the scattering times.

Figure 6.7 brings out a few aspects quite clearly: for both the Kretschmann and Otto configuration, there is a wavelength region where the dispersion of the modal index of the surface plasmon is identifiably anomalous. However, the dispersion for the two configurations is quite different, reflecting the large difference of the surface plasmon's modal profile in the two cases studied. One reason is that the 'Otto plasmon' resides largely on the interface between MgF_2 and aluminum, whereas the mode in the Kretschmann configuration essentially has the character of an aluminum-air surface plasmon. Our Otto plasmon is therefore much more 'metallic' than our Kretschmann plasmon; the former is a better vehicle for prob-

ing the metal's material dispersion. A second reason that may play an important role in explaining the difference between the dispersion of the two modes is the effect of metal film thickness. The Kretschmann plasmon resides on an aluminum layer so thin that its optical properties seem to be quite different from those of bulk metallic aluminum. In the Otto setup, the film is sufficiently thick that its behavior is closer to that of bulk metal.

Another difference apparent from the figure is that the modal loss is much smaller for the Kretschmann plasmon than for its Otto counterpart. This difference can also be attributed to the Otto plasmon being considerably more metallic.

6.4 Conclusion

We have demonstrated anomalous dispersion in the effective mode index of surface plasmons on metallic aluminum layers, paired with air in the Kretschmann configuration and MgF_2 in the Otto configuration. The anomalous dispersion is a direct consequence of aluminum's optical properties, and not those of the dielectric: specifically, the absorption due to aluminum's interband transition at 1.5 eV.

The surface plasmon modes we have measured are not, in either of the two configurations, the same as the simple surface plasmon mode on an infinite interface between two half-spaces. Each configuration has its own plasmonic mode with a slightly different mode index. We find that the properties of the 'Otto plasmon' have more in common with the metal than those of the 'Kretschmann plasmon,' which is situated more in the dielectric.

Our Kretschmann results suggest, by the weak anomalous dispersion, that aluminum's interband transition is only weakly present if the aluminum layer is thin enough. However, our Otto results show the stronger anomalous dispersion expected based on various models for the optical properties of aluminum.

Appendix 6.A Comparison of dielectric functions obtained by reflection and ellipsometry

In order to check the consistency of our results, we also analyzed the Kretschmann reflection curves using the entire Fabry-Perot expression for reflection, $R = |r_{0123}|^2$, with the real and imaginary parts of ε_1, the dielectric function of the aluminum layer, as fit parameters. We used the layer thicknesses obtained by ellipsometry, and Sellmeier models for the dielectric functions of BK7[29] and Si_3N_4.[30] Figure 6.8 shows the aluminum dielectric function we obtained in this way, compared to the aluminum dielectric function we obtained by ellipsometry.

[29] Schott AG, 2012.
[30] Bååk, 1982.

We conclude that the methods are in excellent agreement about the real part of the dielectric function, but that the Kretschmann measurement has a tendency to overestimate the imaginary part compared to the ellipsometric measurement.

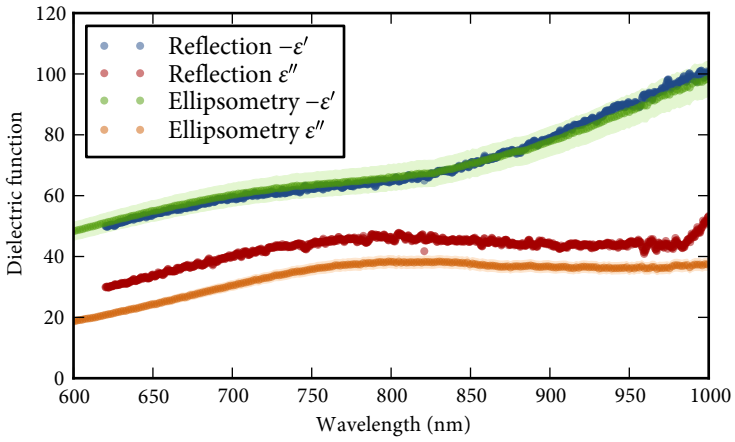


Figure 6.8: Dielectric function of our 10.8 nm layer of aluminum, measured by ellipsometry (green, negative real part, and orange, imaginary part) and Kretschmann reflectance (blue, negative real part, and red, imaginary part). The shaded areas indicate the confidence interval output by the ellipsometer's fitting routine.

7
Enhancing the anomalous surface plasmon dispersion in aluminum

The effective index of the surface plasmon mode an an aluminum surface has a region of anomalous dispersion in the near-infrared as a consequence of aluminum's parallel-band transition at 1.5 eV. By cooling aluminum to 86 K and performing surface plasmon resonance experiments in the Otto configuration, we demonstrate a sizeable enhancement of this anomalous dispersion. The second-order dispersion parameter derived from our measurements increases from its room temperature value of 9000 ps/nm·km to 25000 ps/nm·km.

7.1 Introduction

It is widely appreciated that aluminum is a good approximation to a free-electron metal. Aluminum is also eminently suitable for use in plasmonic applications at wavelengths shorter than 600 nm, where gold's absorption starts to be a hindrance.[1] However, aluminum has some surprising properties in the vicinity of its absorption peak around $\lambda \approx 800$ nm. In chapter 6, we reported anomalous dispersion of the surface plasmon mode on an aluminum interface, exploiting this absorption. This feature is unusual, since the dispersion of the surface plasmon mode is usually normal, a consequence of the Drude-like behavior of most plasmonic metals.

[1] West et al., 2010.

Aluminum's absorption peak manifests most visibly as a reflectance dip familiar to anyone who has used aluminum mirrors, and it is due to a parallel-band transition at 1.5 eV.[2] More accurate analysis turned up another parallel-band transition at 0.5 eV ($\lambda \approx 2500$ nm),[3] which had gone undiscovered until then because the Drude-like absorption dominates at that point. Shortly after the latter work, Ashcroft and Sturm developed a theoretical model for aluminum's parallel-band transitions, and pointed

[2] Ehrenreich et al., 1963.

[3] Bos and Lynch, 1970.

out that parallel bands in the vicinity of the {200} set of crystal planes cause the 1.5 eV absorption peak, and likewise, bands near the {111} planes are responsible for the 0.5 eV peak.[4]

[4] Ashcroft and Sturm, 1971.

Interestingly, the parallel-band absorption peaks in aluminum have a pronounced temperature dependence. When the temperature is reduced, they shift toward slightly higher energies, narrow, and become stronger.[5] A pressure dependence was also discovered: at 30 GPa, both absorption peaks move by almost a whole electron volt toward higher energies.[6] In addition, parallel-band transitions occur only in crystalline aluminum, not in liquid aluminum[7] or in other solid states.[8] After the 1980s, the subject of aluminum's parallel-band transitions seems to have been largely forgotten, but the data are far from buried. They are readily available in Palik's well-known *Handbook of Optical Constants of Solids*.[9]

[5] Liljenvall et al., 1971; Mathewson and Myers, 1972; Benbow and Lynch, 1975.

[6] Tups and Syassen, 1984.

[7] Miller, 1969.

[8] Bernland, Hunderi, and Myers, 1973.

[9] Smith et al., 1985.

In this chapter, we present experimental measurements of the temperature-dependent surface plasmon dispersion relation in aluminum in the neighborhood of the resonance at 1.5 eV, using the method of attenuated total reflection in the Otto configuration. We cool the aluminum to liquid-nitrogen temperatures, causing a giant increase in the anomalous dispersion that accompanies the resonance, compared to what we reported in chapter 6. Finally, we discuss the feasibility of surface plasmon solitons on a liquid nitrogen-cooled aluminum surface.

7.2 *Temperature dependence of the parallel-band absorption*

THE 1960S AND 1970S saw a considerable amount of labor expended on measuring and understanding the optical properties of metallic aluminum. Much of that work was focused on the energy range below 2.5 eV, where interband absorption plays an important role. Work by several authors[10] provided a reasonably accurate description of the interband features and their impact on the optical properties, in particular the interband feature at 1.5 eV. The picture that arose is that most of the interband absorption is due to aluminum's parallel bands near the Fermi energy.[11]

[10] Bennett et al., 1963; Ehrenreich et al., 1963; Dresselhaus, Dresselhaus, and Beaglehole, 1971.

[11] Harrison, 1966.

In 1971, building on this work, Ashcroft and Sturm proposed a more detailed description of the intraband and interband absorption[12] which was in good agreement with experimental data over a relatively large range of energies. Moreover, it allowed them to predict a temperature dependence of the interband absorption peak. Mathewson and Myers[13] employed this theoretical model to analyze the optical constants of metallic aluminum over a wide range of energies, at various temperatures, work which was extended by Benbow and Lynch.[14]

[12] Ashcroft and Sturm, 1971.

[13] Mathewson and Myers, 1972.

[14] Benbow and Lynch, 1975.

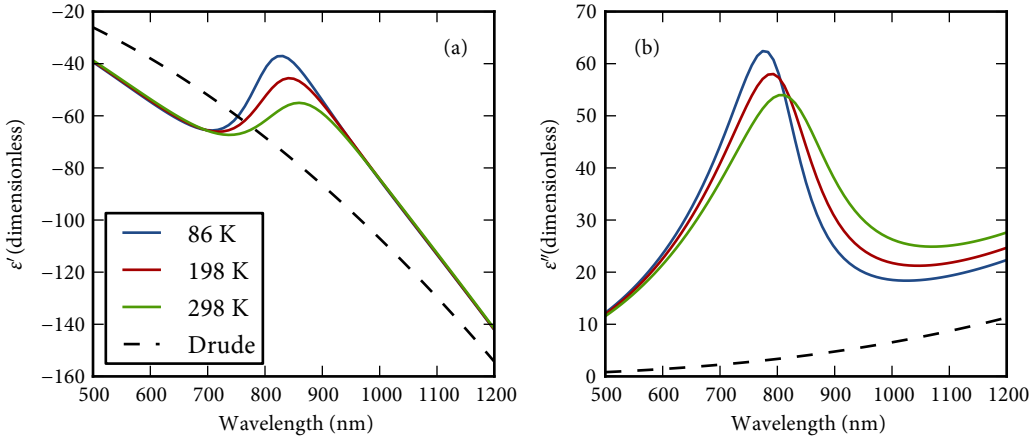


Figure 7.1: Real (a) and imaginary (b) parts of the dielectric function of aluminum, calculated using the model of Ashcroft and Sturm (1971) with temperature-dependent parameters extrapolated from the experimental results of Mathewson and Myers (1972) and Benbow and Lynch (1975). The dashed line is the Drude contribution, without the parallel-band resonance's influence.

Taking the Ashcroft and Sturm model as a starting point, we show the wavelength dependence of the real and imaginary parts of the dielectric function of aluminum in Fig. 7.1. We calculate the curves using parameters derived from Mathewson and Myers, supplemented by data for 4.2 K from Benbow and Lynch. The green (room temperature) and red (198 K) curves correspond to measurements reported in Mathewson and Myers. The blue curve is an estimate, according to the model, of the dielectric function of the aluminum in our liquid nitrogen-cooled experiments at 86 K. The dashed curve shows the Drude contribution to the real and imaginary parts of the dielectric constant. Note that these results suggest that the parallel-band transition's influence can be felt over the whole visible spectrum. While the real part ε' is dominated by the Drude response, the imaginary part ε'' is not.

The resonant feature at $\lambda \approx 850$ nm (1.5 eV) in ε' and ε'' becomes more pronounced as the temperature is reduced. Possibly more important, in particular from the point of view of anomalous dispersion, is that the feature sharpens in both the real and imaginary parts when the temperature is reduced.

7.3 *Experiment*

In order to measure the dispersion relation of the surface plasmon mode on an aluminum surface, we used the method of attenuated total reflection in the Otto configuration,[15] illustrated schematically in Fig. 7.2. The relative merits of the Kretschmann and Otto experimental

[15] Otto, 1968.

configurations have been extensively discussed in chapters 5 and 6; the most important reason for using the Otto configuration is the ability to employ a thick layer of aluminum, with supposedly known optical constants. The thin layers required for the Kretschmann configuration do not always exhibit the parallel-band resonance we are trying to probe,[16] or it is there to a lesser degree.[17]

[16] Novotny, Bulir, et al., 2011.

[17] We have performed preliminary experiments on the low-temperature response of aluminum in a Kretschmann configuration, which also suggest an increase of the anomalous dispersion of the surface plasmon mode.

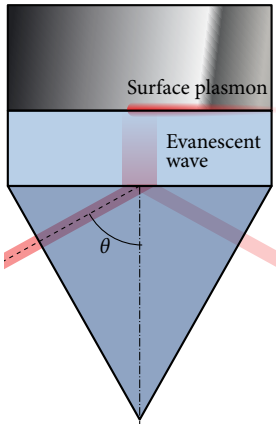


Figure 7.2: The evanescent tail of a light beam, incident in a high-index dielectric (dark blue) at an angle θ greater than the critical angle, excites a surface plasmon on the interface between a metal and a low-index dielectric (light blue).

The Otto configuration involves three materials, a high-index dielectric (dielectric function ε_0), a low-index dielectric (dielectric function ε_1), and a metal (complex dielectric function ε_2, with real part $\varepsilon_2' < 0$). A light beam is incident at an angle θ greater than the critical angle for total internal reflection at the 0–1 interface and undergoes total internal reflection, its evanescent tail extending into medium 1. When the parallel component of the wave vector of the incident beam in medium 0 phase-matches to the surface plasmon mode on the 1–2 interface, then the incident wave can couple to this surface plasmon mode, which is visible as an attenuation of the reflected wave. This evanescent wave coupling is similar to the phenomenon of frustrated total internal reflection.

We used an equilateral prism made of Schott F2 glass ($n \approx 1.608$ at 800 nm) with coatings deposited on all three sides: broadband near-infrared antireflection coatings ($R < 2\%$ for 500–940 nm) on two sides, and three layers on the third side: 570 nm MgF_2 (the low-index dielectric), 100 nm aluminum, and 110 nm SiO_2 as a capping layer to prevent the aluminum from oxidizing.

We placed the prism in a θ–2θ reflectometry setup, already described in chapter 6, and illustrated in Fig. 6.1. This time, the prism was enclosed within a home-built liquid nitrogen bath cryostat. The prism was clamped between two copper cold plates in thermal contact with the liquid nitrogen bath. Monitoring the temperature of the cold plates during the experiment proved them to be stable at 77 K; simulations with the heat flow module of COMSOL suggest that the aluminum layer's temperature reaches a steady state of 86 K within twenty minutes of the cold plates reaching 77 K. The prism assembly was shielded with fiberglass insulation on the sides where optical access was not required, and the entire cryostat was placed on top of the rotation stage of the θ–2θ reflectometer.

As one can see in Fig. 7.1, we need to probe a relatively broad spectral region between approximately 700 and 1000 nm. Therefore, as a light source we used a collimated beam from a fiber-coupled tungsten halogen fiber source (Ocean Optics HL-2000-FHSA-HP), and measured the reflectance spectrum from 620–1000 nm as a function of angle of incidence using a fiber-coupled mini-spectrometer (Ocean Optics USB2000). We

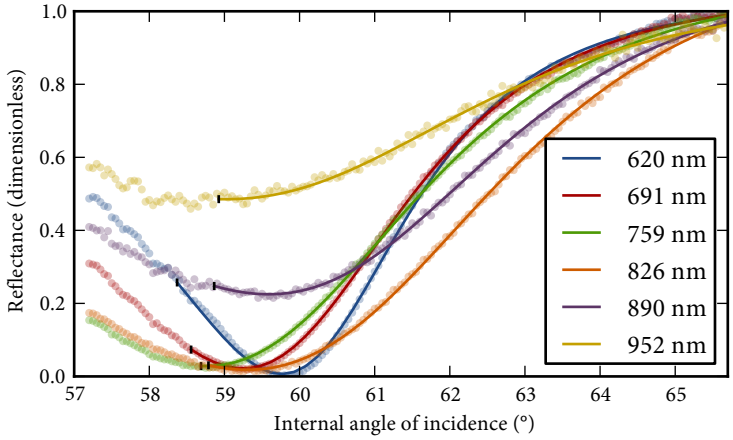


Figure 7.3: Measured reflectance curves for several representative wavelengths. The dots indicate measured values, and the solid lines are fits to the data using (7.1), discussed later in this chapter. The fits cut off at the critical angle for total internal reflection from F2 to MgF_2, where the fit function no longer applies. This is indicated by the small black markers.

show experimental results for selected wavelengths across the wavelength range of interest in Fig. 7.3.

From these measurements we extract the resonance angle by determining the angle at which the reflection reaches a minimum.[18] Figure 7.4 shows the resonance angle as a function of wavelength, for each spectrometer wavelength bin.

[18] For this, we use the minimum of the fitting function (7.1), discussed in the next section.

The resonance angle decreases as a function of wavelength at both long and short wavelengths, as if aluminum were a Drude-like metal in these wavelength ranges. However, the slope of the curve reverses sign in between, and, for the low-temperature curve, the resonance angle increases from about 770 to 870 nm. The data become more noisy at $\lambda > 950$ nm; this reflects the layer thickness getting farther and farther from the thickness necessary for critical coupling. This makes the surface plasmon resonance dip shallower, and its minimum therefore harder to pinpoint exactly.

7.4 *Analysis*

We cannot directly obtain the dispersion of the surface plasmon mode from the results shown in Fig. 7.4, as discussed in chapter 5, since the resonance angle is only an approximate indicator of the surface-plasmon wave vector at a particular wavelength. Therefore, we analyze our angle-dependent wavelength curves using the method outlined in chapter 5 in order to obtain both the real and imaginary parts of the wave vector. We

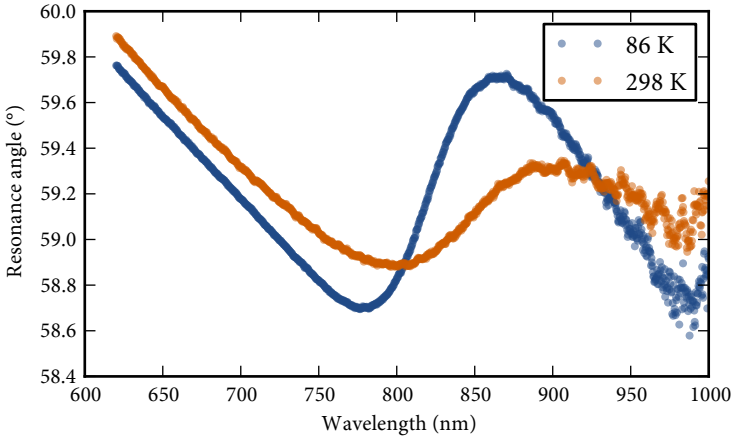


Figure 7.4: Angle of incidence at which the reflectance reaches a minimum, as a function of free-space wavelength. For comparison, we show the same quantity for room temperature (see Fig. 6.6b.)

fit the data with a Fano-type lineshape:

$$R(k_x) = \left| 1 + \frac{A e^{i\phi} k''_{\mathrm{SP}}}{k'_{\mathrm{SP}} + i k''_{\mathrm{SP}} - k_x} \right|^2 , k_x > k_{cr}. \tag{7.1}$$

In this expression, there are four fit parameters: the resonance amplitude A, the phase difference ϕ between the resonance and the background, and the real and imaginary parts of the mode index k'_{SP} and k''_{SP}. We discard the part of the curve measured at angles less than the critical angle, which allows us to set a unit background and fit with only four parameters instead of five as in (5.15). This procedure yields a complex-valued surface plasmon wave vector k_{SP}, and we define the effective surface plasmon mode index to be $n_{\mathrm{SP}} = k_{\mathrm{SP}}/k_0$. This effective mode index is plotted as a function of wavelength in Figure 7.5, where we see a region of anomalous dispersion from about 750 nm to 850 nm, accompanied by an absorption peak centered around 800 nm. The solid lines are the mode index calculated for this system, taking the Ashcroft-Sturm model for the dielectric function of aluminum, as shown in Fig. 7.1. We took as the pseudopotential $U_{200} = 0.795$ eV, as the Drude scattering time $\tau_D = 11.5$ fs, and as the parallel-band scattering time $\tau_I = 6.0$ fs. For the purpose of this calculation, we used a Sellmeier model for the dispersion of F2 glass,[19] MgF_2,[20] and SiO_2.[21] Both the real and imaginary parts of the experimentally determined mode index are slightly higher than the calculated prediction, but exhibit very similar overall behavior.

[19] Schott AG, 2012.
[20] Dodge, 1984.
[21] Malitson, 1965.

Fig. 7.5 demonstrates that the parallel-band absorption resonance in metallic aluminum and its associated anomalous dispersion do indeed influence the dispersion relation of surface plasmons traveling along an

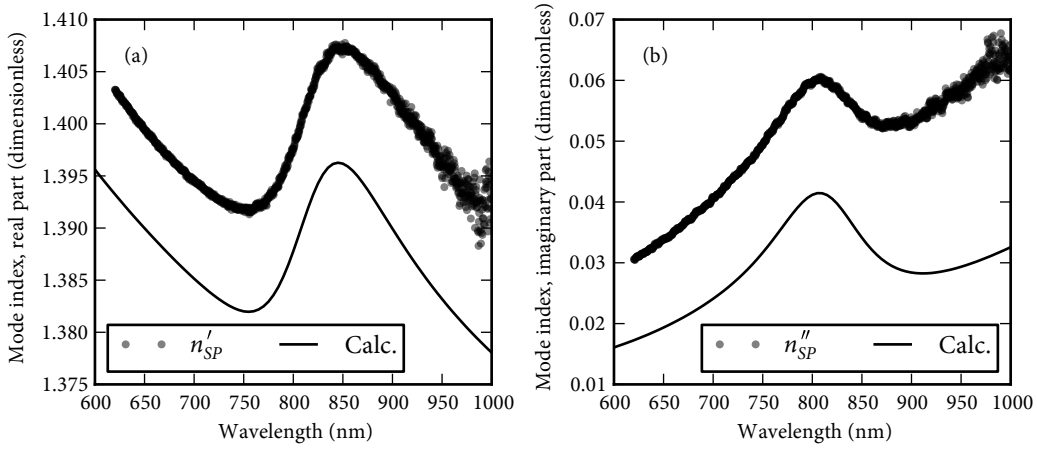


Figure 7.5: Measured real part (a) and imaginary part (b) of the surface plasmon mode index, determined by analysis of the angle-dependent reflectance curves, measured at $T = 86$ K. The solid lines are a calculation of the mode index to be expected for this particular Otto-configuration system, based on the Ashcroft-Sturm model for aluminum's dielectric function and published values for the other materials' optical properties. Compare to Fig. 6.7.

interface between aluminum and a dielectric.

The dispersive features shown in Fig. 7.5 suggest that a short (less than 100 fs) surface plasmon pulse with its center wavelength in the vicinity of the parallel-band absorption will experience substantial pulse reshaping. In order to obtain some insight into the pulse propagation we evaluate the modal group index,

$$n_g = \frac{d\omega}{dk} = n - \lambda \frac{dn}{d\lambda}, \tag{7.2}$$

and the second-order dispersion parameter, which is proportional to the group index's derivative,

$$D = -\frac{\lambda}{c} \frac{d^2 n}{d\lambda^2}. \tag{7.3}$$

In both of the above equations, n refers to the real part of the surface-plasmon mode index. The dispersion is called normal for $D < 0$, and anomalous when $D > 0$.

In order to find the group index and second-order dispersion parameter we smooth the data of Fig. 7.5 using a Savitzky-Golay filter,[22] allowing us to estimate the curves' derivatives. We plot the surface plasmon group index estimated from our data in Fig. 7.6, and for comparison, we also show the group index obtained in the same way from the room-temperature Otto data of chapter 6. We clearly see that a plasmonic pulse would experience strong dispersion in the parallel-band region, which is greatly enhanced by cooling down the aluminum.

[22] Savitzky and Golay, 1964.

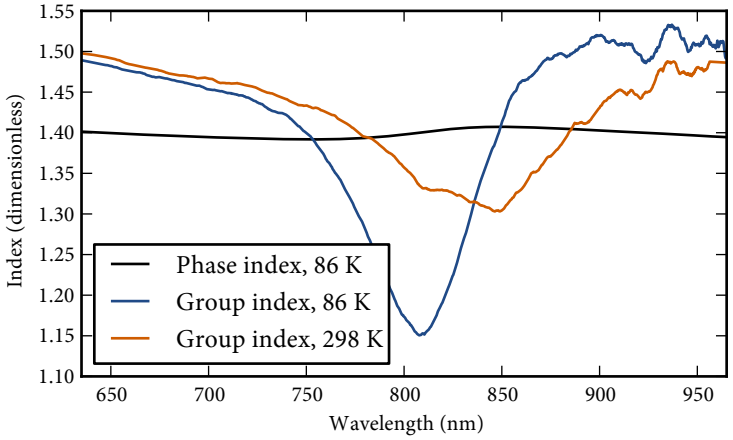


Figure 7.6: Real parts of the group and phase index (smoothed version of Fig. 7.5a) of the surface plasmon mode, as a function of free-space wavelength. For comparison, the same quantity at room temperature is shown (orange), as calculated from the data of chapter 6.

Our estimates for the second-order dispersion parameter D, expressed in ps/nm·km,[23] are shown in Fig. 7.7. At both temperatures, there is a region where D is positive, meaning that the group velocity dispersion is anomalous.

[23] A unit for D common in fiber optics, meaning picosecond delay per nanometer bandwidth and kilometer propagation length.

At the peak of the anomalous dispersion regime, the second-order dispersion reaches the large value of 25000 ps/nm · km. The peak value at room temperature is about 9000 ps/nm · km, meaning that cooling to liquid nitrogen temperatures causes almost a threefold increase of this peak value. The wavelength range in which anomalous dispersion occurs, however, is narrower (about 70 nm versus 100 nm at room temperature.) This is consistent with Mathewson and Myers'[24] earlier measurements demonstrating that the parallel-band resonance became stronger and narrower at lower temperatures.

[24] Mathewson and Myers, 1972.

We now consider using this anomalous dispersion to create a surface plasmon soliton. The anomalous dispersion region is about 70 nm wide, centered around 830 nm. This bandwidth corresponds approximately to a 14 fs Fourier-limited pulse width. Surface plasmons generally have a short decay length, typically less than 50 μm in the relevant wavelength range. The broadening of a 14 fs pulse associated with $D = 25000$ ps/nm · km over a length of 50 μm is 88 fs. However, the $1/e$ surface plasmon amplitude damping length at 830 nm in our measurements is a mere 2.2 μm, leading to a pulse broadening of 3.8 fs over one damping length.

However, this does not mean that a surface plasmon soliton is possible on an unmodified aluminum-vacuum interface. A soliton pulse in

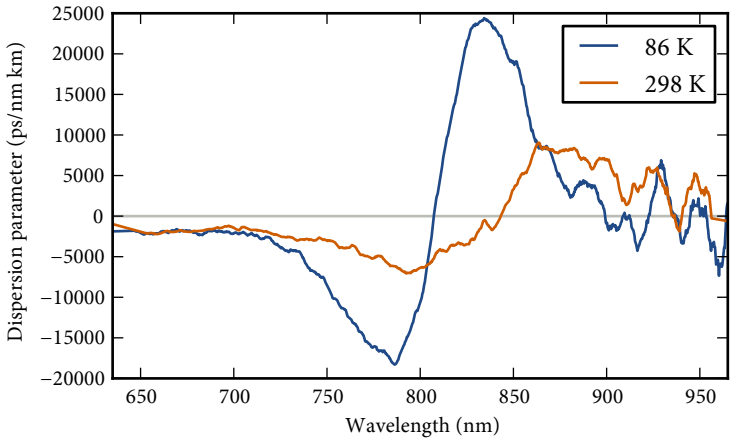


Figure 7.7: Second-order dispersion parameter D estimated from the measured surface plasmon dispersion relation, as a function of free-space wavelength. Note the region where $D > 0$, where the group dispersion is anomalous.

conventional optics requires both anomalous dispersion and Kerr nonlinearity in order to maintain its shape. For a surface plasmon soliton, the aluminum must be paired with a dielectric exhibiting the Kerr effect, changing the dispersion relation. Although Sámson et al.[25] suggest using a pulse with a peak power high enough that the metal's intrinsic Kerr nonlinearity suffices, we suspect that this would prove prohibitively damaging. Huang, Chang, Leung, and Tsai[26] suggest a possible expression for the dispersion relation of a surface plasmon on the interface between an ideal metal and a Kerr medium.

[25] Sámson et al., 2011.

[26] Huang et al., 2009.

The propagation losses are an additional consideration for a surface plasmon soliton. When the $1/e$ amplitude damping length of a surface plasmon is little more than 2 μm, then to be of any use there must be some form of loss compensation added to the surface plasmon mode. Pairing the metal with a pumped gain medium as the dielectric[27] seems a promising way to achieve this.

[27] Noginov et al., 2008.

7.5 *Conclusion*

By cooling aluminum to 86 K, we have demonstrated a sizeable enhancement of the anomalous dispersion in the effective surface plasmon mode index previously reported in chapter 6. The measured dispersion relation indicates that there is a wavelength region between approximately 810 and 880 nm where the second-order dispersion is anomalous, meaning this system is capable of supporting a plasmon soliton, provided that some form of nonlinear Kerr effect can be added and the propagation losses can be compensated.

Appendix 7.A Ashcroft-Sturm model for the temperature-dependent parallel-band conductivity

[28] Ashcroft and Sturm, 1971.

ASHCROFT AND STURM[28] were the first researchers to derive explicit expressions for the contribution to the optical conductivity, and therefore to the dielectric function, from parallel-band transitions. However, their paper unfortunately contains several misprints, making it nearly impossible to use their results simply by reading their paper. In this appendix we hope to correct this state of affairs by reprinting the expressions we obtained by going through their derivation carefully and simplifying the notation here and there.

Ashcroft and Sturm cast their model in the form of the optical conductivity, which describes the metal's optical properties equivalently to the dielectric function ε. They are related as follows:

$$\varepsilon(\omega) = 1 + i\sigma(\omega)/\omega\varepsilon_0. \tag{7.4}$$

The parallel-band model, $\sigma(\omega) = \sigma_D(\omega) + \sigma_I(\omega)$, consists of an interband contribution $\sigma_I(\omega)$ and a free-electron contribution $\sigma_D(\omega)$,

$$\sigma_D(\omega) = \frac{\sigma_{DC}}{1 - i\omega\tau_D}. \tag{7.5}$$

DC: direct current

Equation (7.5) is equivalent to the more familiar form of the Drude model, (5.1), with the Drude electron scattering time $\tau_D = 1/\gamma$ and the DC conductivity $\sigma_{DC} = \varepsilon_0\tau_D\omega_p^2$. The interband contribution takes the form of $\sigma_I(\omega) = \sum_K \sigma_{I,K}(\omega)$, where K is the reciprocal lattice vector corresponding to the set of planes in which the parallel bands occur, and the expression takes the sum over all such sets of planes. The contribution from one set of planes $\sigma_{I,K}(\omega)$ is a complicated expression, but it is governed by the interband electron scattering time τ_I, and the two energies between which parallel-band absorption occurs: the lower bound

$$\hbar\omega_- = 2U_K, \tag{7.6}$$

where U_K is the Fourier component of the pseudopotential for K; and the upper bound

$$\hbar\omega_+ = (\hbar^2K^2/2m_*)(2k_F/K - 1), \tag{7.7}$$

where m_* is the effective optical mass of a conduction electron and k_F is the Fermi wave vector. There is also a "normal" interband contribution to the absorption (i.e. not due to parallel bands) which dominates for $\omega > \omega_+$, but it is negligible within the frequency regime that we are examining.

In short, the model is an integral over a density of states in the portion of k-space in which the parallel-band absorption occurs. The two scattering times τ_D and τ_I are important, because they remove all sorts of influences from the model. For example, the influence of aluminum film thickness discussed in chapter 6 manifests itself as a broadening of the interband scattering, which is why the resonance is less in evidence. Experiments suggest that in rough terms, $\tau_D \approx 2\tau_I$ for any temperature,[29] which suggests that the electron-phonon scattering dominates τ_I.[30] However, in other circumstances, experimental observations suggest that the values of τ_I for resonances associated with different sets of planes diverge, and separate values τ_{200} and τ_{111} are necessary to explain the observed data.[31]

[29] Mathewson and Myers, 1972.

[30] Benbow and Lynch, 1975.

[31] Benbow and Lynch, 1975; Tups and Syassen, 1984.

THE FULL EXPRESSIONS for the real and imaginary parts of the parallel-band conductivity for one set of planes, reproduced from Ashcroft and Sturm in slightly simplified form, are:

$$\sigma(\omega)' = \frac{\sigma_0 z J(\omega)}{\rho(z^2 + b^2)}; \tag{7.8}$$

$$\begin{aligned}\sigma(\omega)'' = \frac{\sigma_0}{2\pi b\rho}\Bigg(&\tfrac{1}{2}\sin\varphi_+ \ln\frac{t_0^2 + 2t_0\rho\cos\varphi_+ + \rho^2}{t_0^2 - 2t_0\rho\cos\varphi_+ + \rho^2} \\ &+ \cos\varphi_+\left(\arctan\frac{t_0 + \rho\cos\varphi_+}{\rho\sin\varphi_+} + \arctan\frac{t_0 - \rho\cos\varphi_+}{\rho\sin\varphi_+}\right) \\ &+ \pi\frac{b^2 - z^2}{b^2 + z^2}J(\omega)\Bigg),\end{aligned} \tag{7.9}$$

where the quantities $z = \omega/\omega_-$, $b = 1/\omega_-\tau_I$, $\rho = \left((1 - b^2 + z^2)^2 + 4z^2b^2\right)^{1/4}$, and $t_0 = \sqrt{\omega_+/\omega_- - 1}$ are all dimensionless and represent frequency ratios, the angles

$$\varphi_\pm = \frac{\pi}{4} \pm \frac{1}{2}\arctan\frac{1 + b^2 - z^2}{2bz}, \tag{7.10}$$

σ_0 is a constant with units of conductivity defined as $e^2 M_K K/24\pi\hbar$ with e the elementary charge and M_K the multiplicity associated with the K set of planes, i.e. how many planes bound the first Brillouin zone: 48 for the {111} set of planes and 24 for the {200} set of planes. (The solid angle wedge defined by the *XULKW* points in the Brillouin zone is one forty-eighth of

the entire zone.) $J(\omega)$ is a complicated real-valued function defined as

$$\begin{aligned}\pi J(\omega) = {} & \frac{4zb\rho}{z^2+b^2}\arctan t_0 \\ & + \frac{1}{2}\left(\frac{z^2-b^2}{z^2+b^2}\cos\varphi_- + \frac{2zb}{z^2+b^2}\sin\varphi_-\right)\ln\frac{t_0^2+2t_0\rho\sin\varphi_-+\rho^2}{t_0^2-2t_0\rho\sin\varphi_-+\rho^2} \\ & + \left(\left(\frac{z^2-b^2}{z^2+b^2}\sin\varphi_- - \frac{2zb}{z^2+b^2}\cos\varphi_-\right)\right. \\ & \left.\times\left(\arctan\frac{t_0+\rho\sin\varphi_-}{\rho\cos\varphi_-} + \arctan\frac{t_0-\rho\sin\varphi_-}{\rho\cos\varphi_-}\right)\right). \end{aligned} \tag{7.11}$$

Contrary to what Ashcroft and Sturm assert on p. 1902, $J(\omega)$ does not tend to unity for $\omega \to \infty$.

In order to gain a little insight into these difficult expressions and to reproduce Ashcroft and Sturm's Figs. 4 and 8, we also consider the expressions for the limit in which there are no electron collisions ($\tau \to \infty$). In this case, $J(\omega)$ reduces to the much simpler

$$J(\omega) = H(\omega-\omega_-) - H(\omega-\omega_+), \tag{7.12}$$

where $H(\omega)$ is the Heaviside step function. It is now apparent that the physical meaning of $J(\omega)$ is simply the frequency range in which parallel-band absorption occurs (ω_- to ω_+): the parallel-band absorption is absent at lower energies, and at higher energies it is replaced by what Ashcroft and Sturm term the "normal" interband absorption, which we shall not cover in this appendix. The complicated expression in which electron collisions are included simply softens the sharp transition between absorption regimes. The real part for the collisionless case is then

$$\sigma(\omega)' = \begin{cases} \frac{\sigma_0}{z\sqrt{z^2-1}}, & \omega_- < \omega < \omega_+, \\ 0, & \text{otherwise}, \end{cases} \tag{7.13}$$

The imaginary part is given by two expressions, one for frequencies below ω_- and one for frequencies above:

$$\sigma(\omega)'' = \begin{cases} \frac{2\sigma_0}{\pi z}\left(\frac{1}{\sqrt{1-z^2}}\arctan\frac{\sqrt{z_0^2-1}}{\sqrt{1-z^2}} - \arctan\sqrt{z_0^2-1}\right), & \omega < \omega_-, \\ \frac{2\sigma_0}{\pi z}\left(\frac{1}{2\sqrt{z^2-1}}\ln\frac{\sqrt{z_0^2-1}-\sqrt{z^2-1}}{\sqrt{z_0^2-1}+\sqrt{z^2-1}} - \arctan\sqrt{z_0^2-1}\right), & \omega > \omega_- \end{cases}, \tag{7.14}$$

where $z_0 = \omega_+/\omega_-$.

Ashcroft and Sturm calculate their figures using values for the scattering times of $\tau_I = \tau_D = 0.6 \times 10^{-14}$ s,[32] pseudopotentials $U_{111} = 0.0179$ Ry (0.244 eV) and $U_{200} = 0.0562$ Ry (0.765 eV),[33] and $a_0 k_F = 0.9247$. We have made our computer code available with which we calculate the conductivity and dielectric function of aluminum.[34]

[32] Ashcroft and Sturm, 1971.

[33] Ashcroft, 1963.

[34] Chimento, 2013a.

Bibliography

Abbott, E. A. (1884). *Flatland: a romance of many dimensions*. London: Seeley & Co. http://www.gutenberg.org/ebooks/97

Allen, L., Barnett, S. M., & Padgett, M. J. (Eds.). (2003). *Optical angular momentum*. Bristol: Institute of Physics Publishing.

Allen, L., Beijersbergen, M. W., Spreeuw, R. J. C., & Woerdman, J. P. (1992, June). Orbital angular momentum of light and the transformation of Laguerre-Gaussian laser modes. *Physical Review A*, *45*(11), 8185–8189. doi:10.1103/PhysRevA.45.8185

Altewischer, E., van Exter, M. P., & Woerdman, J. P. (2002, July). Plasmon-assisted transmission of entangled photons. *Nature*, *418*(6895), 304–306. doi:10.1038/nature00869

Ashcroft, N. W. (1963, December). The Fermi surface of aluminium. *Philosophical Magazine*, *8*(96), 2055–2083. doi:10.1080/14786436308209098

Ashcroft, N. W. & Sturm, K. (1971, March). Interband absorption and the optical properties of polyvalent metals. *Physical Review B*, *3*(6), 1898. doi:10.1103/PhysRevB.3.1898

Ashcroft, N. W. & Mermin, N. D. (1976). *Solid state physics*. Philadelphia: Saunders College.

Astilean, S., Lalanne, P., & Palamaru, M. (2000, March). Light transmission through metallic channels much smaller than the wavelength. *Optics Communications*, *175*(4-6), 265–273. doi:16/S0030-4018(00)00462-4

Atwater, H. A. (2007, April). The promise of plasmonics. *Scientific American*, 56–63. http://www.scientificamerican.com/article.cfm?id=the-promise-of-plasmonics

Bååk, T. (1982, March). Silicon oxynitride; a material for GRIN optics. *Applied Optics*, *21*(6), 1069–1072. doi:10.1364/AO.21.001069

Babeva, T., Kitova, S., Mednikarov, B., & Konstantinov, I. (2002, July). Preparation and characterization of a reference aluminum mirror. *Applied Optics*, *41*(19), 3840–3846. doi:10.1364/AO.41.003840

Baranova, N., Zel'dovich, B. Ya., Mamaev, A. V., Pilipetskii, N. F., & Shkunov, V. V. (1981, February). Dislocations of the wavefront of a http://www.jetpletters.ac.ru/ps/1503/article_22988.shtml

speckle-inhomogeneous field (theory and experiment). *JETP Letters*, *33*(4), 195–199.

doi:10.1016/0030-4018(95)00267-C Basistiy, I. V., Soskin, M. S., & Vasnetsov, M. V. (1995, September). Optical wavefront dislocations and their properties. *Optics Communications*, *119*(5-6), 604–612.

doi:10.1364/OE.16.003420 Baudrion, A.-L., de León-Pérez, F., Mahboub, O., Hohenau, A., Ditlbacher, H., García-Vidal, F. J., Dintinger, J., Ebbesen, T. W., Martín-Moreno, L., & Krenn, J. R. (2008, March). Coupling efficiency of light to surface plasmon polariton for single subwavelength holes in a gold film. *Optics Express*, *16*(5), 3420–3429.

http://www.jetpletters.ac.ru/ps/1159/article_17529.shtml Bazhenov, V. Yu., Vasnetsov, M. V., & Soskin, M. S. (1990, October). Laser beams with screw dislocations in their wavefronts. *JETP Letters*, *52*(8), 429–31.

doi:10.1364/AO.14.001322 Bell, R. J., Goben, C. A., Davarpanah, M., Bhasin, K., Begley, D. L., & Bauer, A. C. (1975, June). Two-dimensional optics with surface electromagnetic waves. *Applied Optics*, *14*(6), 1322–1325.

doi:10.1103/PhysRevB.12.5615 Benbow, R. L. & Lynch, D. W. (1975, December). Optical absorption in Al and dilute alloys of Mg and Li in Al at 4.2 K. *Physical Review B*, *12*(12), 5615–5621.

doi:10.1364/JOSA.53.001089 Bennett, H. E., Silver, M., & Ashley, E. J. (1963, September). Infrared reflectance of aluminum evaporated in ultra-high vacuum. *Journal of the Optical Society of America*, *53*(9), 1089–1095.

doi:10.1038/nphoton.2011.285 Berini, P. & De Leon, I. (2012, January). Surface plasmon-polariton amplifiers and lasers. *Nature Photonics*, *6*(1), 16–24.

doi:10.1088/1464-4258/11/9/094021 Berkhout, G. C. G. & Beijersbergen, M. W. (2009, September). Using a multipoint interferometer to measure the orbital angular momentum of light in astrophysics. *Journal of Optics A: Pure and Applied Optics*, *11*(9), 094021.

doi:10.1103/PhysRevLett.31.363 Bernland, L. G., Hunderi, O., & Myers, H. P. (1973, August). Optical absorption in vapor-quenched aluminum. *Physical Review Letters*, *31*(6), 363–365.

doi:10.1088/1464-4258/7/11/011 Berry, M. V., Jeffrey, M. R., & Mansuripur, M. (2005, November). Orbital and spin angular momentum in conical diffraction. *Journal of Optics A: Pure and Applied Optics*, *7*(11), 685–690.

doi:10.1103/PhysRev.50.115 Beth, R. A. (1936, July). Mechanical detection and measurement of the angular momentum of light. *Physical Review*, *50*(2), 115–125.

doi:10.1103/PhysRev.66.163 Bethe, H. A. (1944, October). Theory of diffraction by small holes. *Physical Review*, *66*(7-8), 163–182.

doi:10.1103/PhysRevA.79.041803 Bliokh, K. Y., Bliokh, Y. P., & Ferrando, A. (2009, April). Resonant

plasmon-soliton interaction. *Physical Review A*, *79*(4), 041803.

Bogers, M. (2012, July). *Creating different polarisations using sub wavelength slits in a zigzag pattern* (Bachelor's thesis, Leiden). http://www.molphys.leidenuniv.nl/qo/thesis/Bachelor-Mark-Bogers.pdf

Bomzon, Z., Kleiner, V., & Hasman, E. (2001, September). Pancharatnam–Berry phase in space-variant polarization-state manipulations with subwavelength gratings. *Optics Letters*, *26*(18), 1424–1426. doi:10.1364/OL.26.001424

Bos, L. W. & Lynch, D. W. (1970, July). Low-energy optical absorption peak in aluminum and Al-Mg alloys. *Physical Review Letters*, *25*(3), 156. doi:10.1103/PhysRevLett.25.156

Bosman, J. (2011). *A subwavelength slit as a quarterwave retarder* (Bachelor's thesis, Leiden). http://www.molphys.leidenuniv.nl/qo/thesis/Bachelor-Johan-Bosman.pdf

Bouwhuis, G., Braat, J., Huijser, A., Pasman, J., Van Rosmalen, G., & Schouhamer Immink, K. (1985). *Principles of optical disk systems*. Bristol, England: Adam Hilger Ltd.

Bouwkamp, C. J. (1954, January). Diffraction theory. *Reports on Progress in Physics*, *17*(1), 35–100. doi:10.1088/0034-4885/17/1/302

Bozhevolnyi, S. I. & Pudonin, F. A. (1997, April). Two-dimensional micro-optics of surface plasmons. *Physical Review Letters*, *78*(14), 2823–2826. doi:10.1103/PhysRevLett.78.2823

Brasselet, E., Izdebskaya, Ya., Shvedov, V., Desyatnikov, A. S., Krolikowski, W., & Kivshar, Yu. S. (2009, April). Dynamics of optical spin-orbit coupling in uniaxial crystals. *Optics Letters*, *34*(7), 1021–1023. doi:10.1364/OL.34.001021

Brasselet, E. & Loussert, C. (2011, March). Electrically controlled topological defects in liquid crystals as tunable spin-orbit encoders for photons. *Optics Letters*, *36*(5), 719–721. doi:10.1364/OL.36.000719

Brust, D. (1970, August). Electronic structure effects in the Drude and interband absorption of aluminum. *Physical Review B*, *2*(4), 818–825. doi:10.1103/PhysRevB.2.818

Cai, L., Liu, Q., & Yang, X. (2003, June). A simple method of contrast enhancement and extremum extraction for interference fringes. *Optics & Laser Technology*, *35*(4), 295–302. doi:10.1016/S0030-3992(03)00022-7

Chimento, P. F., 't Hooft, G. W., & Eliel, E. R. (2010a, November). Plasmonic optical vortex analyzer. In J. Pozo, M. Mortensen, P. Urbach, X. Leijtens, & M. Yousefi (Eds.), *Proceedings of the 2010 annual symposium of the IEEE Photonics Benelux Chapter*, November 19, 2010 (pp. 17–20). 2010 Annual Symposium of the IEEE Photonics Benelux Chapter. Delft, Netherlands: Uitgeverij TNO. http://www.photonics-benelux.org/images/stories/media/proceedings/2010/s10p017.pdf

https://gist.github.com/ptomato/5262520 Chimento, P. F. (2013a, March). Ashcroft-Sturm model.

https://gist.github.com/ptomato/5262493 Chimento, P. F. (2013b, March). Waveguide model for a subwavelength slit as a quarter-wave retarder.

doi:10.1364/OL.37.004946 Chimento, P. F., Alkemade, P. F. A., 't Hooft, G. W., & Eliel, E. R. (2012, December). Optical angular momentum conversion in a nanoslit. *Optics Letters*, *37*(23), 4946–4948.

doi:10.1364/OE.19.024219 Chimento, P. F., Kuzmin, N. V., Bosman, J., Alkemade, P. F. A., 't Hooft, G. W., & Eliel, E. R. (2011, November). A subwavelength slit as a quarter-wave retarder. *Optics Express*, *19*(24), 24219–24227.

doi:10.1364/OL.35.003775 Chimento, P. F., 't Hooft, G. W., & Eliel, E. R. (2010b, November). Plasmonic tomography of optical vortices. *Optics Letters*, *35*(22), 3775–3777.

doi:10.1364/JOSAA.20.000163 Ciattoni, A., Cincotti, G., & Palma, C. (2003, January). Circularly polarized beams and vortex generation in uniaxial media. *Journal of the Optical Society of America A*, *20*(1), 163–171.

doi:10.1016/j.optcom.2008.09.043 Davis, T. (2009, January). Surface plasmon modes in multi-layer thin-films. *Optics Communications*, *282*(1), 135–140.

doi:10.1364/OE.17.021732 Davoyan, A. R., Shadrivov, I. V., & Kivshar, Y. S. (2009, November). Self-focusing and spatial plasmon-polariton solitons. *Optics Express*, *17*(24), 21732–21737.

doi:10.1364/OE.18.020610 Devaux, E., Laluet, J.-Y., Stein, B., Genet, C., Ebbesen, T., Weeber, J.-C., & Dereux, A. (2010, September). Refractive micro-optical elements for surface plasmons: from classical to gradient index optics. *Optics Express*, *18*(20), 20610.

doi:doi:10.1063/1.1506018 Ditlbacher, H., Krenn, J. R., Schider, G., Leitner, A., & Aussenegg, F. R. (2002, September). Two-dimensional optics with surface plasmon polaritons. *Applied Physics Letters*, *81*(10), 1762–1764.

doi:10.1364/AO.23.001980 Dodge, M. J. (1984, June). Refractive properties of magnesium fluoride. *Applied Optics*, *23*(12), 1980–1985.

http : / / books . google . com / books ? id = jX6LI73mUaUC & pg = PA33 Dresselhaus, G., Dresselhaus, M. S., & Beaglehole, D. (1971, December). Electronic density of states. In L. H. Bennett (Ed.), November 3–6, 1969 (pp. 33–36). 3rd Materials Research Symposium. Gaithersburg, Maryland: US National Bureau of Standards.

doi:10.1016/j.optmat.2005.04.011 Du, H., Xiao, J. Q., Zou, Y. S., Wang, T. G., Gong, J., Sun, C., & Wen, L. S. (2006, June). Optical properties of ultrathin aluminum films deposited by magnetron sputtering in visible band. *Optical Materials*, *28*(8–9), 944–949.

doi:10.1038/35570 Ebbesen, T. W., Lezec, H. J., Ghaemi, H. F., Thio, T., & Wolff, P. A. (1998, February). Extraordinary optical transmission through sub-wavelength hole arrays. *Nature*, *391*(6668), 667–669.

Ehrenreich, H., Philipp, H. R., & Segall, B. (1963, December). Optical properties of aluminum. *Physical Review*, *132*(5), 1918. doi:10.1103/PhysRev.132.1918

Feigenbaum, E. & Orenstein, M. (2007, March). Plasmon-soliton. *Optics Letters*, *32*(6), 674–676. doi:10.1364/OL.32.000674

Fienup, J. R. (1982, August). Phase retrieval algorithms: a comparison. *Applied Optics*, *21*(15), 2758–2769. doi:10.1364/AO.21.002758

Fivash, M., Towler, E. M., & Fisher, R. J. (1998, February). Biacore for macromolecular interaction. *Current Opinion in Biotechnology*, *9*(1), 97–101. doi:10.1016/S0958-1669(98)80091-8

Fizeau, M. H. (1861, December). Recherches sur plusieurs phénomènes relatifs à la polarisation de la lumière. *Annales de chimie et de physique*, *63*, 385. http://gallica.bnf.fr/ark:/12148/bpt6k348044/f382.image.langEN

Foo, G., Palacios, D. M., & Swartzlander, G. A., Jr. (2005, December). Optical vortex coronagraph. *Optics Letters*, *30*(24), 3308–3310. doi:10.1364/OL.30.003308

García-Vidal, F. J. (2006, March). Solid-state physics: light at the end of the channel. *Nature*, *440*(7083), 431–433. doi:10.1038/440431a

Genevet, P., Yu, N., Aieta, F., Lin, J., Kats, M. A., Blanchard, R., Scully, M. O., Gaburro, Z., & Capasso, F. (2012, January). Ultra-thin plasmonic optical vortex plate based on phase discontinuities. *Applied Physics Letters*, *100*(1), 013101. doi:doi:10.1063/1.3673334

Gómez Rivas, J., Kuttge, M., Kurz, H., Haring Bolivar, P., & Sánchez-Gil, J. A. (2006, February). Low-frequency active surface plasmon optics on semiconductors. *Applied Physics Letters*, *88*(8), PAGES. doi:doi:10.1063/1.2177348

González, M. U., Weeber, J.-C., Baudrion, A.-L., Dereux, A., Stepanov, A. L., Krenn, J. R., Devaux, E., & Ebbesen, T. W. (2006, April). Design, near-field characterization, and modeling of 45°surface-plasmon Bragg mirrors. *Physical Review B*, *73*(15), 155416. doi:10.1103/PhysRevB.73.155416

Gorbach, A. V. (2013, January). Nonlinear graphene plasmonics: Amplitude equation for surface plasmons. *Physical Review A*, *87*(1), 013830. doi:10.1103/PhysRevA.87.013830

Gorodetski, Y., Shitrit, N., Bretner, I., Kleiner, V., & Hasman, E. (2009, August). Observation of optical spin symmetry breaking in nanoapertures. *Nano Letters*, *9*(8), 3016–3019. doi:10.1021/nl901437d

Griffiths, D. J. (1999). *Introduction to electrodynamics*. Prentice Hall.

Harris, M., Hill, C. A., Tapster, P. R., & Vaughan, J. M. (1994, April). Laser modes with helical wave fronts. *Physical Review A*, *49*(4), 3119. doi:10.1103/PhysRevA.49.3119

Harrison, W. A. (1966, July). Parallel-band effects in interband optical absorption. *Physical Review*, *147*(2), 467–469. doi:10.1103/PhysRev.147.467

doi:10.1088/0022-3727/45/43/433001 Hayashi, S. & Okamoto, T. (2012, October). Plasmonics: visit the past to know the future. *Journal of Physics D: Applied Physics*, *45*(43), 433001.

doi:10.1016/j.optcom.2008.12.025 Huang, J.-H., Chang, R., Leung, P.-T., & Tsai, D. P. (2009, April). Nonlinear dispersion relation for surface plasmon at a metal-Kerr medium interface. *Optics Communications*, *282*(7), 1412–1415.

doi:10.1088/0305-4608/6/6/033 Hunderi, O. (1976, June). Optical properties of metallic calcium. *Journal of Physics F: Metal Physics*, *6*(6), 1223.

http://cise.aip.org/getabs/servlet/GetabsServlet?prog=normal&id=CSENFA000009000003000090000001 Hunter, J. D. (2007, May). Matplotlib: a 2D graphics environment. *Computing in Science & Engineering*, *9*(3), 90–95.

http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=10025 ISO 4225:1994: *Air quality — General aspects — Vocabulary*. Geneva, Switzerland: International Organization for Standardization.

doi:10.1103/PhysRev.188.759 Jasperson, S. N. & Schnatterly, S. E. (1969, December). Photon-surface-plasmon coupling in thick Ag foils. *Physical Review*, *188*(2), 759–770.

doi:10.1364/OE.15.005801 Jesacher, A., Schwaighofer, A., Fürhapter, S., Maurer, C., Bernet, S., & Ritsch-Marte, M. (2007, April). Wavefront correction of spatial light modulators using an optical vortex image. *Optics Express*, *15*(9), 5801–5808.

http://www.scipy.org/ Jones, E., Oliphant, T., & Peterson, P. (2001). SciPy: open source scientific tools for Python.

doi:10.1098/rspa.1954.0191 Jones, R. V. & Richards, J. C. S. (1954, August). The polarization of light by narrow slits. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *225*(1160), 122–135.

doi:10.1103/PhysRevLett.108.044801 Karimi, E., Marrucci, L., Grillo, V., & Santamato, E. (2012, January). Spin-to-orbital angular momentum conversion and spin-polarization filtering in electron beams. *Physical Review Letters*, *108*(4), 044801.

doi:10.1364/OL.34.001225 Karimi, E., Piccirillo, B., Marrucci, L., & Santamato, E. (2009, April). Light propagation in a birefringent plate with topological charge. *Optics Letters*, *34*(8), 1225–1227.

doi:10.1364/OL.36.002498 Khoo, E. H., Li, E. P., & Crozier, K. B. (2011, July). Plasmonic wave plate based on subwavelength nanoslits. *Optics Letters*, *36*(13), 2498–2500.

doi:10.1063/1.2840675 Kihm, H. W., Lee, K. G., Kim, D. S., Kang, J. H., & Park, Q.-H. (2008, February). Control of surface plasmon generation efficiency by slit-width tuning. *Applied Physics Letters*, *92*(5), 051115.

doi:10.1086/142326 King, A. S. (1917, May). A study with the electric furnace of the anomalous dispersion of metallic vapors. *The Astrophysical Journal*, *45*, 254.

Kretschmann, E. (1971, August). Die Bestimmung optischer Konstanten von Metallen durch Anregung von Oberflächenplasmaschwingungen. *Zeitschrift für Physik*, *241*(4), 313–324. doi:10.1007/BF01395428

Kurihara, K., Nakamura, K., & Suzuki, K. (2002, August). Asymmetric SPR sensor response curve-fitting equation for the accurate determination of SPR resonance angle. *Sensors and Actuators B: Chemical*, *86*(1), 49–57. doi:10.1016/S0925-4005(02)00146-6

Kuzmin, N. V. (2008, January). *Interference effects with surface plasmons* (Doctoral dissertation, Leiden). http://hdl.handle.net/1887/12551

Lalanne, P., Hugonin, J. P., & Rodier, J. C. (2005, December). Theory of surface plasmon generation at nanoslit apertures. *Physical Review Letters*, *95*(26), 263902. doi:10.1103/PhysRevLett.95.263902

Lalanne, P., Hugonin, J. P., & Rodier, J. C. (2006, July). Approximate model for surface-plasmon generation at slit apertures. *Journal of the Optical Society of America A*, *23*(7), 1608–1615. doi:10.1364/JOSAA.23.001608

Leach, J., Padgett, M. J., Barnett, S. M., Franke-Arnold, S., & Courtial, J. (2002, June). Measuring the orbital angular momentum of a single photon. *Physical Review Letters*, *88*(25), 257901. doi:10.1103/PhysRevLett.88.257901

Liedberg, B., Nylander, C., & Lunström, I. (1983, July). Surface plasmon resonance for gas detection and biosensing. *Sensors and Actuators*, *4*, 299–304. doi:10.1016/0250-6874(83)85036-7

Liljenvall, H. G., Mathewson, A. G., & Myers, H. P. (1971, January). The temperature dependence of the optical constants of aluminium. *Solid State Communications*, *9*(2), 169–170. doi:10.1016/0038-1098(71)90281-X

Lombard, E., Drezet, A., Genet, C., & Ebbesen, T. W. (2010, February). Polarization control of non-diffractive helical optical beams through subwavelength metallic apertures. *New Journal of Physics*, *12*(2), 023027. doi:10.1088/1367-2630/12/2/023027

Loo, C., Lowery, A., Halas, N., West, J., & Drezek, R. (2005, April). Immunotargeted nanoshells for integrated cancer imaging and therapy. *Nano Letters*, *5*(4), 709–711. doi:10.1021/nl050127s

Lord Rayleigh. (1897). On the passage of waves through apertures in plane screens, and allied problems. *Philosophical Magazine*, *43*(263), 259–272. doi:10.1080/14786449708620990

Maiden, A. M. & Rodenburg, J. M. (2009, September). An improved ptychographical phase retrieval algorithm for diffractive imaging. *Ultramicroscopy*, *109*(10), 1256–1262. doi:10.1016/j.ultramic.2009.05.012

Maier, S. A. & Atwater, H. A. (2005, July). Plasmonics: localization and doi:10.1063/1.1951057

guiding of electromagnetic energy in metal/dielectric structures. *Journal of Applied Physics*, *98*(1), 011101–10.

doi:10.1038/35085529 Mair, A., Vaziri, A., Weihs, G., & Zeilinger, A. (2001, July). Entanglement of the orbital angular momentum states of photons. *Nature*, *412*, 313–316.

doi:10.1364/JOSA.55.001205 Malitson, I. H. (1965, October). Interspecimen comparison of the refractive index of fused silica. *Journal of the Optical Society of America*, *55*(10), 1205–1208.

http://www.alcatel-lucent.com/bstj/vol43-1964/articles/bstj43-4-1783.pdf Marcatili, E. A. J. & Schmeltzer, R. A. (1964, July). Hollow metallic and dielectric waveguides for long distance optical transmission and lasers. *Bell Systems Technical Journal*, *43*(4), 1783–1809.

doi:10.1111/j.1365-2818.2007.01727.x Marian, A., Charrière, F., Colomb, T., Montfort, F., Kühn, J., Marquet, P., & Depeursinge, C. (2007, February). On the complex three-dimensional amplitude point spread function of lenses and microscope objectives: theoretical aspects, simulations and measurements by digital holography. *Journal of Microscopy*, *225*(2), 156–169.

doi:10.1364/OE.19.006616 Marini, A., Skryabin, D. V., & Malomed, B. (2011, March). Stable spatial plasmon solitons in a dielectric-metal-dielectric geometry with gain and loss. *Optics Express*, *19*(7), 6616–6622.

doi:10.1103/PhysRevLett.96.163905 Marrucci, L., Manzo, C., & Paparo, D. (2006, April). Optical spin-to-orbital angular momentum conversion in inhomogeneous anisotropic media. *Physical Review Letters*, *96*(16), 163905.

doi:10.1088/0305-4608/2/2/030 Mathewson, A. G. & Myers, H. P. (1972, March). Optical absorption in aluminium and the effect of temperature. *Journal of Physics F: Metal Physics*, *2*(2), 403–415.

doi:10.1088/0305-4608/3/3/021 Mathewson, A. G. & Myers, H. P. (1973, March). The optical properties of lithium-magnesium alloys. *Journal of Physics F: Metal Physics*, *3*(3), 623–639.

doi:10.1080/14786436908228198 Miller, J. C. (1969). Optical properties of liquid metals at high temperatures. *Philosophical Magazine*, *20*(168), 1115–1132.

doi:10.1038/nphys607 Molina-Terriza, G., Torres, J. P., & Torner, L. (2007, May). Twisted photons. *Nature Physics*, *3*(5), 305–310.

doi:10.1103/RevModPhys.57.783 Moskovits, M. (1985, July). Surface-enhanced spectroscopy. *Reviews of Modern Physics*, *57*(3), 783–826.

doi:10.1364/OE.16.001385 Noginov, M. A., Podolskiy, V. A., Zhu, G., Mayy, M., Bahoura, M., Adegoke, J. A., Ritzo, B. A., & Reynolds, K. (2008, January). Compensation of loss in propagating surface plasmon polariton by gain in adjacent dielectric medium. *Optics Express*, *16*(2), 1385–1392.

Novotny, L. & van Hulst, N. (2011, February). Antennas for light. *Nature Photonics*, *5*(2), 83–90. doi:10.1038/nphoton.2010.237

Novotny, M., Bulir, J., Lancok, J., Pokorny, P., & Bodnar, M. (2011, February). In-situ monitoring of the growth of nanostructured aluminum thin film. *Journal of Nanophotonics*, *5*(1), PAGES. doi:doi:10.1117/1.3543816

Nugrowati, A. M., Pereira, S. F., & van de Nes, A. S. (2008, May). Near and intermediate fields of an ultrashort pulse transmitted through Young's double-slit experiment. *Physical Review A*, *77*(5), 053810. doi:10.1103/PhysRevA.77.053810

Oemrawsingh, S. S. R., Eliel, E. R., Woerdman, J. P., Verstegen, E. J. K., Kloosterboer, J. G., & 't Hooft, G. W. (2004, May). Half-integral spiral phase plates for optical wavelengths. *Journal of Optics A: Pure and Applied Optics*, *6*(5), s288–s290. doi:10.1088/1464-4258/6/5/029

O'Neil, A. T., MacVicar, I., Allen, L., & Padgett, M. J. (2002, January). Intrinsic and extrinsic nature of the orbital angular momentum of a light beam. *Physical Review Letters*, *88*(5), 053601. doi:10.1103/PhysRevLett.88.053601

Otto, A. (1968, August). Excitation of nonradiative surface plasma waves in silver by the method of frustrated total reflection. *Zeitschrift für Physik*, *216*(4), 398–410. doi:10.1007/BF01391532

Padgett, M. J., Arlt, J., Simpson, N., & Allen, L. (1996, January). An experiment to observe the intensity and phase structure of Laguerre-Gaussian laser modes. *American Journal of Physics*, *64*(1), 77–82. doi:10.1119/1.18283

Padgett, M. J., Courtial, J., & Allen, L. (2004, May). Light's orbital angular momentum. *Physics Today*, *57*(5), 35. doi:10.1063/1.1768672

Powell, C. J. & Swan, J. B. (1959, August). Origin of the characteristic electron energy losses in aluminum. *Physical Review*, *115*(4), 869–875. doi:10.1103/PhysRev.115.869

Quail, J. C., Rako, J. G., & Simon, H. J. (1983, July). Long-range surface-plasmon modes in silver and aluminum films. *Optics Letters*, *8*(7), 377–379. doi:10.1364/OL.8.000377

Raether, H. (1988). *Surface plasmons on smooth and rough surfaces and on gratings*. Berlin: Springer-Verlag.

Rakić, A. D., Djurišić, A. B., Elazar, J. M., & Majewski, M. L. (1998, August). Optical properties of metallic films for vertical-cavity optoelectronic devices. *Applied Optics*, *37*(22), 5271–5283. doi:10.1364/AO.37.005271

Ray, R. (2006, July). TOPEX/Poseidon: revealing hidden tidal energy. Retrieved March 11, 2013. http://svs.gsfc.nasa.gov/stories/topex/tides.html

Rich, R. L. & Myszka, D. G. (2008, November). Survey of the year 2007 commercial optical biosensor literature. *Journal of Molecular Recognition*, *21*(6), 355–400. doi:10.1002/jmr.928

doi:10.1002/jmr.1004 Rich, R. L. & Myszka, D. G. (2010, January). Grading the commercial optical biosensor literature—Class of 2008: 'The Mighty Binders'. *Journal of Molecular Recognition*, *23*(1), 1–64.

doi:10.1103/PhysRev.106.874 Ritchie, R. H. (1957, June). Plasma losses by fast electrons in thin films. *Physical Review*, *106*(5), 874–881.

doi:10.1103/PhysRevLett.21.1530 Ritchie, R. H., Arakawa, E. T., Cowan, J. J., & Hamm, R. N. (1968, November). Surface-plasmon resonance effect in grating diffraction. *Physical Review Letters*, *21*(22), 1530–1533.

doi:10.1364/OL.37.001820 Roberts, A. & Lin, L. (2012, June). Plasmonic quarter-wave plate. *Optics Letters*, *37*(11), 1820–1822.

doi:10.1080/00107519108211048 Sambles, J. R., Bradbery, G. W., & Yang, F. (1991). Optical excitation of surface plasmons: an introduction. *Contemporary Physics*, *32*(3), 173–183.

doi:10.1364/OL.36.000250 Sámson, Z. L., Horak, P., MacDonald, K. F., & Zheludev, N. I. (2011, January). Femtosecond surface plasmon pulse propagation. *Optics Letters*, *36*(2), 250–252.

doi:10.1103/PhysRevB.60.8359 Sánchez-Gil, J. A. & Maradudin, A. A. (1999). Near-field and far-field scattering of surface plasmon polaritons by one-dimensional surface defects. *Physical Review B*, *60*(11), 8359.

doi:10.1021/ac60214a047 Savitzky, A. & Golay, M. J. E. (1964, July). Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, *36*(8), 1627–1639.

doi:10.1119/1.2386162 Schaefer, B., Collett, E., Smyth, R., Barrett, D., & Fraher, B. (2007, February). Measuring the Stokes polarization parameters. *American Journal of Physics*, *75*(2), 163.

http://www.schott.com/advanced_optics/english/download/schott_optical_glass_collection_datasheets_dec_2012_en.pdf Schott AG. (2012). Optical glass data sheets. Retrieved February 4, 2013.

doi:10.1103/PhysRevLett.94.053901 Schouten, H. F., Kuzmin, N., Dubois, G., Visser, T. D., Gbur, G., Alkemade, P. F. A., Blok, H., 't Hooft, G. W., Lenstra, D., & Eliel, E. R. (2005, February). Plasmon-assisted two-slit transmission: Young's experiment revisited. *Physical Review Letters*, *94*(5), 053901.

doi:10.1088/1464-4258/6/5/027 Schouten, H. F., Visser, T. D., Gbur, G., Lenstra, D., & Blok, H. (2004, May). The diffraction of light by narrow slits in plates of different materials. *Journal of Optics A: Pure and Applied Optics*, *6*(5), S277–S280.

doi:10.1103/PhysRevE.67.036608 Schouten, H. F., Visser, T. D., Lenstra, D., & Blok, H. (2003, March). Light transmission through a subwavelength slit: waveguiding and optical vortices. *Physical Review E*, *67*(3), 036608.

Smith, D. Y., Shiles, E., & Inokuti, M. (1985). The optical properties of metallic aluminum. In E. D. Palik (Ed.), *Handbook of optical constants of solids* (pp. 369–406). San Diego: Academic Press.

Snyder, A. W. & Love, J. D. (1983). *Optical waveguide theory*. London: Chapman & Hall.

Soliton wave receives crowd of admirers. (1995, August). *Nature*, *376*(6539), 373. doi:10.1038/376373b0

Soskin, M. S. & Vasnetsov, M. V. (2001). Singular optics. In E. Wolf (Ed.), *Progress in optics* (Vol. 42, pp. 219–276). Progress in optics. Amsterdam: Elsevier.

Strong, J. (1936, June). The evaporation process and its application to the aluminizing of large telescope mirrors. *The Astrophysical Journal*, *83*, 401. doi:10.1086/143734

Suckling, J. R., Hibbins, A. P., Lockyear, M. J., Preist, T. W., Sambles, J. R., & Lawrence, C. R. (2004, April). Finite conductance governs the resonance transmission of thin metal slits at microwave frequencies. *Physical Review Letters*, *92*(14), 147401. doi:10.1103/PhysRevLett.92.147401

Takakura, Y. (2001, June). Optical resonance in a narrow slit in a thick metallic screen. *Physical Review Letters*, *86*(24), 5601. doi:10.1103/PhysRevLett.86.5601

Teperik, T. V., Archambault, A., Marquier, F., & Greffet, J. J. (2009, September). Huygens-Fresnel principle for surface plasmons. *Optics Express*, *17*(20), 17483–17490. doi:10.1364/OE.17.017483

Tillin, M. D. & Sambles, J. R. (1988, December). A surface plasmon-polariton study of the dielectric constants of reactive metals: aluminium. *Thin Solid Films*, *167*(1–2), 73–84. doi:10.1016/0040-6090(88)90483-X

Tummers, B. (2006). DataThief III. http://datathief.org

Tups, H. & Syassen, K. (1984, November). Effect of pressure on the optical absorption in aluminium. *Journal of Physics F: Metal Physics*, *14*(11), 2753–2767. doi:10.1088/0305-4608/14/11/028

Wähling, G. (1981, April). Resonant interaction of surface plasmons with a dye monolayer. *Zeitschrift für Naturforschung*, *36a*, 588–594.

Walasik, W., Nazabal, V., Chauvet, M., Kartashov, Y., & Renversez, G. (2012, November). Low-power plasmon–soliton in realistic nonlinear planar structures. *Optics Letters*, *37*(22), 4579–4581. http://ol.osa.org/abstract.cfm?URI=ol-37-22-4579

West, P. R., Ishii, S., Naik, G. V., Emani, N. K., Shalaev, V. M., & Boltasseva, A. (2010, November). Searching for better plasmonic materials. *Laser & Photonics Reviews*, *4*(6), 795–808. doi:10.1002/lpor.200900055

Yang, F. & Sambles, J. R. (2002, July). Resonant transmission of microwaves through a narrow metallic slit. *Physical Review Letters*, *89*(6), 063901. doi:10.1103/PhysRevLett.89.063901

Samenvatting

Een oppervlakteplasmon is een lichtgolf die gebonden is aan een metaaloppervlak. Er bestaan twee soorten van: de ene komt voor bij metalen deeltjes met sub-micron afmetingen, de andere bij metaaloppervlakken die tenminste op de schaal van de optische golflengte vlak zijn. In dit proefschrift, *"Tweedimensionale optica: diffractie en dispersie bij oppervlakteplasmonen,"* gaat het over het laatste type. Deze soort plant zich langs het twee-dimensionale metaaloppervlak voort, in tegenstelling tot 'gewoon' licht dat door de driedimensionale ruimte reist. De binding van oppervlakteplasmonen aan het metaaloppervlak maakt het mogelijk optische signalen te sturen door kanalen met extreem kleine afmetingen.

Deel 1: Verschijnselen bij dunne spleten in metaalfilms

Uit eerder werk, o.a. dat van mijn voorganger Nikolay Kuzmin, was bekend dat onder bepaalde omstandigheden een zeer nauwe spleet of kras in een heel dunne metaallaag invallend licht met de juiste polarisatie om kan zetten in oppervlakteplasmonen, en omgekeerd.

In hoofdstuk 2 bekijken we bij een metaalfilm gemaakt van goud met een dikte van 0.0002 mm, hoeveel licht door spleten van verschillende breedte, varierend van 0.00005 tot 0.0005 mm, doorgelaten wordt, en hoe dit doorgelaten licht gepolariseerd is. We gebruiken het aanslaan van oppervlakteplasmonen bij één polarisatie om de polarisatie van het doorgelaten licht te kunnen beheersen. Het blijkt dat bij een bepaalde spleetbreedte en filmdikte de spleet lineair gepolariseerd licht kan omzetten in circulair gepolariseerd, en omgekeerd. Wij hebben een simpel model ontwikkeld dat deze uitkomst ook op intuïtieve wijze verklaart. Dit resultaat is een handige manier om op kleine schaal de functionaliteit van een zogenaamde kwart-lambdaplaat te realiseren. In hoofdstuk 3 gebruiken we dit verschijnsel nog eens, maar met cirkelvormige spleten, om een optische draaikolk te laten ontstaan uit circulair gepolariseerd licht, en daarmee

optisch spinimpulsmoment om te zetten in optisch baanimpulsmoment.

Hoofdstuk 4 beschrijft een proef met twee zeer nauwe spleten die in een heel dunne goudfilm parallel aan elkaar zijn gekerfd. De ene spleet wordt beschenen met licht; dat wordt daar gedeeltelijk omgezet in oppervlakteplasmonen. De oppervlakteplasmonen reizen naar de andere spleet waar ze weer omgezet worden in licht. We meten de lichtverdeling, maar onderweg is deze door buiging van vorm veranderd ten opzichte van de verdeling bij het invallende licht. Deze vormverandering gebruiken we om informatie over de fase (het golffront) van het invallende licht te achterhalen; de fase kan niet direct worden gemeten en wordt meestal gemeten met behulp van interferentie met een tweede lichtbundel. Als toepassing van deze techniek meten we de fase van een bundel die een optische draaikolk bevat, maar uiteindelijk kan de techniek leiden tot een golffrontsensor met een veel hogere ruimtelijke resolutie dan de gangbare technieken, wat interessant zou kunnen zijn voor de astronomie en UV-lithografie.

Deel 2: Anomale dispersie van oppervlakteplasmonen

Dispersie is het verschijnsel dat de snelheid waarmee licht zich door een materiaal voortplant afhangt van de golflengte van het licht (de kleur). Bijvoorbeeld, een puls van rood licht en één van blauw die op hetzelfde moment een blok glas worden ingestuurd, komen aan de andere kant op verschillende momenten aan. Gewoonlijk komt het rode licht eerder aan dan het blauwe (deze situatie wordt 'normale dispersie' genoemd), maar soms is het andersom: 'anomale' dispersie. Anomale dispersie heb je nodig om solitonen te laten ontstaan, lichtpulsen die een lange afstand kunnen afleggen zonder te vervormen.

Anomale dispersie komt vaak voor in de buurt van golflengtes die het materiaal absorbeert. Het metaal aluminium heeft zo'n absorptie in het nabije infrarood. In het tweede deel van dit werk proberen wij de vraag te beantwoorden of deze absorptie ook anomale dispersie van oppervlakteplasmonen aan een aluminiumoppervlak met zich meebrengt. Dit onderzoeken wij met een methode, waarbij de oppervlakteplasmonen aangeslagen worden door licht aan te voeren vanuit een prisma. Deze techniek kent twee varianten, genoemd naar de Duitse onderzoekers Kretschmann en Otto. Van de Ottoconfiguratie wordt vaak gedacht dat deze alleen maar nadelen biedt vergeleken met de Kretschmannconfiguratie. In hoofdstuk 5 laten we zien dat dit een misverstand is. Daarnaast introduceren we een analysemethode waarmee wij de resultaten van experimenten

aan verliesgevende metalen met beide opstellingen zinvol kunnen interpreteren, wat niet mogelijk is met de gebruikelijke aanpak.

Hoofdstukken 6 en 7 beschrijven de meetresultaten aan oppervlakteplasmonen met anomale dispersie. In hoofdstuk 6 tonen we inderdaad anomale dispersie aan bij oppervlakteplasmonen op een aluminiumoppervlak. Vervolgens maken we de mate van anomale dispersie nog veel groter door het metaal tot vloeibare stikstoftemperatuur af te koelen, ongeveer −200 °C, beschreven in hoofdstuk 7. Dit is echter een afweging tussen meer anomale dispersie en meer verlies, omdat de oppervlakteplasmonen sneller uitdoven op het gekoelde metaal.

Curriculum Vitæ

Philip Francis Chimento III

1981	Born in Raleigh, North Carolina, United States
1993–1994	Secondary education, Sherwood Githens Middle School, Durham, North Carolina, United States
1994–1999	Secondary education, Het Stedelijk Lyceum, Enschede, Netherlands
1999–2008	Bachelor's and Master's degree in applied physics Twente University, Enschede, Netherlands *Freshman year completed cum laude*
2009–2013	PhD research Leiden Institute of Physics, Leiden University, Leiden, Netherlands
2013–	Software engineer, Endless Mobile

List of Publications

- Chimento, P. F., Jurna, M., Bouwmans, H. S. P., Garbacik, E. T., Hartsuiker, L., Otto, C., Herek, J. L., & Offerhaus, H. L. (2009). High-resolution narrowband CARS spectroscopy in the spectral fingerprint region. *Journal of Raman Spectroscopy, 40,* 1229–1233.
- Chimento, P. F., 't Hooft, G. W., & Eliel, E. R. (2010). Plasmonic optical vortex analyzer. In J. Pozo, M. Mortensen, P. Urbach, X. Leijtens, & M. Yousefi (Eds.), *Proceedings of the 2010 annual symposium of the IEEE Photonics Benelux Chapter,* November 19, 2010 (pp. 17–20). 2010 Annual Symposium of the IEEE Photonics Benelux Chapter. Delft, Netherlands: Uitgeverij TNO.
- Chimento, P. F., 't Hooft, G. W., & Eliel, E. R. (2010). Plasmonic tomography of optical vortices. *Optics Letters, 35,* 3775–3777.
- Chimento, P. F., Kuzmin, N. V., Bosman, J., Alkemade, P. F. A., 't Hooft, G. W., & Eliel, E. R. (2011). A subwavelength slit as a quarter-wave retarder. *Optics Express, 19,* 24219–24227.
- Chimento, P. F., Alkemade, P. F. A., 't Hooft, G. W., & Eliel, E. R. (2012). Optical angular momentum conversion in a nanoslit. *Optics Letters, 37,* 4946–4948.
- Chimento, P. F., 't Hooft, G. W., & Eliel, E. R. (2013). When the dip doesn't tell the whole story: interpreting the surface plasmon resonance in lossy metals. Submitted to *Optics Express*.
- Chimento, P. F., 't Hooft, G. W., & Eliel, E. R. Anomalous dispersion of surface plasmons. In preparation.
- Chimento, P. F., 't Hooft, G. W., & Eliel, E. R. Enhancing the anomalous surface plasmon dispersion in aluminum. In preparation.

Acknowledgements

I WOULD MOST LIKE TO THANK the students that I had the pleasure of mentoring: Carolina Rendón Barraza, Johan Bosman, Mark Bogers, David Kok, and Tobias de Jong. They all contributed in important ways, even though the project that Carolina, David, and Tobias worked on did not make it into the publishable stage because of time constraints.

One is not allowed any more to thank one's coworkers indiscriminately, but some people deserve a mention for their contributions beyond those of the co-authors on my papers. Wolfgang Löffler's expertise is woven all throughout this book; he was always ready to bounce ideas off and share lab tips. Michiel de Dood took a special interest in the aluminum project (chapters 6 and 7) and our discussions were invaluable in understanding the solid-state physics involved. Daan Boltje put time into preparing the Kretschmann prisms used in chapter 6.

The work described in chapter 7 involved cryostats and liquid nitrogen, something I had had little experience with when I started. Jelmer Renema helped to close this experience gap, and assisted with the COMSOL heat flow simulations. Mirthe Bergman, Arjen Geluk, and others in the Fine Mechanics Department worked on the cryostat that I used and made sure it was simple, easy, and leak-free.

Philippe Lalanne, professor at the Institut d'Optique, CNRS, was willing to share the Gaussian quadrature code from their paper[35] which I adapted for chapter 2. Speaking of sharing computer code, I relied heavily on open source software almost from the start of this research. NumPy and SciPy[36] did all the number crunching. I made all the graphs in this book with Matplotlib[37] and the diagrams with Inkscape. I used DataThief III[38] to digitize printed specs of anti-reflection coatings.

[35] Lalanne et al., 2006.

[36] Jones, Oliphant, and Peterson, 2001.

[37] Hunter, 2007.

[38] Tummers, 2006.