Iterative design of diffractive elements made of lossy material

Verhoeven Antonie D
Iterative design of diffractive elements made of lossy materials

Antonie D. Verhoeven,1,2,* Frank Wyrowski,2 and Jari Turunen1

1Institute of Photonics, University of Eastern Finland, P.O. Box 111, FI-80101 Joensuu, Finland
2Institute of Applied Physics, Friedrich-Schiller University, Albert-Einstein-Straße 15, D-07745 Jena, Germany
*Corresponding author: antonie.verhoeven@uef.fi

Received 2 August 2017; revised 23 October 2017; accepted 13 November 2017; posted 14 November 2017 (Doc. ID 304039); published 8 December 2017

Diffractive surface relief elements made of lossy materials exhibit phase-dependent absorption, which not only reduces the efficiency but also distorts the signal if the surface profile is realized on the basis of a phase-only design. We introduce an extension of the iterative Fourier transform algorithm, which accounts for such phase-dependent absorption, and present examples of its application to the design of diffractive beam splitters. The operator required for taking absorption into account is chosen to maximize the efficiency of the found design. © 2017 Optical Society of America


https://doi.org/10.1364/JOSAA.35.000045

1. INTRODUCTION

The iterative Fourier transform algorithm (IFTA), based on the classic Gerchberg–Saxton phase retrieval method [1], is a well-established tool in the design of diffractive optics within the paraxial domain, as long as the complex amplitude transmittance approach and the thin-element approximation are applicable [2–4]. It enables efficient synthesis of diffractive elements in the form of continuous or multilevel-quantized surface relief profiles that modulate either the phase or the amplitude of the incident field [5,6]. In general, diffractive elements that modulate the phase rather than the amplitude are preferable because of their higher diffraction efficiency [7], but in certain occasions amplitude modulation is unavoidable. This is the case in spectral regions where materials that are dielectric in the visible domain become lossy [8–10] (infrared, ultraviolet, and extreme ultraviolet domains) and also in the x-ray regime where metallic materials possess complex refractive indices with absolute values close to unity [11]. In such circumstances any surface-relief-type diffractive element, even if realized in the form of a thin self-supporting membrane, will suffer from phase-dependent absorption that can affect its performance in a significant way.

We demonstrate below that if one applies a phase-only design to realize a beam splitter grating in a lossy material, the resulting array of diffraction orders exhibits a substantial degradation of uniformity. It is also shown that the choice of projection operator strongly impacts the performance of the found designs.

Here we extend the IFTA to account for phase-dependent absorption in lossy media. The basic task is described in Section 2 and the new design algorithm is presented in Section 3. In Section 4 we introduce an extension of the analytical triplicator design presented in [12] to gratings with absorption in order to evaluate the performance of the algorithm. Numerical results on the design of beam splitter gratings are provided in Section 5, and some conclusions are drawn in Section 6.

2. DIFFRACTIVE OPTICS WITH LOSSY MEDIA

A. Phase-Amplitude Constraint

Let us assume that a surface-relief-type diffractive element with a surface profile \( h(x, y) \) is made of a lossy material with complex refractive index

\[
\hat{n} = 1 + \Delta n + ik. \tag{1}
\]

Then the refractive index profile in the modulated region can be written as

\[
\hat{n}(x, y) = \begin{cases} \hat{n} & \text{if } 0 \leq z \leq h(x, y) \\ 1 & \text{if } z > h(x, y) \end{cases}. \tag{2}
\]

Assuming that the thin-element approximation is valid, the diffractive element can be characterized by a complex amplitude transmittance \( t(x, y) = |t(x, y)| \exp[i\phi(x, y)] \), where

\[
|t(x, y)| = \exp[-(2\pi/\lambda)nh(x, y)] \tag{3}
\]

is the amplitude transmittance at wavelength \( \lambda \) and

\[
\phi(x, y) = (2\pi/\lambda)\Delta nb(x, y) \tag{4}
\]

is the associated phase delay. Hence the transmission function is constrained to the set of values given by an amplitude-phase relation...

1084-7529/18/010045-10 Journal © 2018 Optical Society of America
The constraint in Eq. (5) can have a significant effect in many applications. It is therefore important to take Eq. (5) into account in the design.

**B. Task Description**

The goal is to find the transmission function \( t(x, y) \in A_c \) such that a specific signal \( U_{\text{signal}} \) is produced at the output of the system. It is assumed that the transmission function and output are connected by a linear invertible operator \( L \) so that

\[
U_{\text{out}} = L(U_{\text{in}}t), \tag{6}
\]

where \( U_{\text{in}} \) denotes the field right before the element represented by the transmission function \( t \).

As described in [13], the output of the system can be uniquely decomposed as

\[
U_{\text{out}} = aU_{\text{signal}} + U_{\text{error}} + U_{\text{freedom}}, \tag{7}
\]

where \( U_{\text{signal}} \) is defined in a specified region \( W \) (the signal window), \( U_{\text{error}} \) is the deviation from the desired signal within \( W \), and \( U_{\text{freedom}} \) denotes the field outside \( W \), as shown in Fig. 3. The goal is to find \( t(x, y) \) that maximizes \( |a| \) while minimizing \( U_{\text{error}} \).

**C. Signal Relevant Efficiency**

The recently introduced signal relevant efficiency (SRE) description [13] will be used to motivate the algorithm and to explain how it works. The SRE is defined as

\[
\eta_{\text{SRE}} = \frac{\|aU_{\text{signal}}\|^2}{\|U_{\text{in}}\|^2} = \frac{|\langle U_{\text{out}} | U_{\text{signal}} \rangle|^2}{\|U_{\text{in}}\|^2 \|U_{\text{signal}}\|^2}. \tag{8}
\]

Here the inner product and the norm are defined as \( \langle U_1 | U_2 \rangle = \int_{\mathbb{R}^2} U_1(x)U_2^*(x)dx \) and \( \|U\| = \sqrt{\langle U | U \rangle} \), respectively, the overline implies complex conjugation, and the integrations are carried out over the \( xy \) plane. In essence, the SRE measures the proportion of power in the incident field that ends up in the desired output signal.

Typically the efficiency is defined as the amount of energy that ends up in the target area over the amount of energy one started with:

\[
\eta = \frac{\|aU_{\text{signal}} + U_{\text{error}}\|^2}{\|U_{\text{in}}\|^2} = \frac{\|aU_{\text{signal}}\|^2 + \|U_{\text{error}}\|^2}{\|U_{\text{in}}\|^2}. \tag{9}
\]
The equality of the two expressions for \( \eta \) holds because, by definition, \( \langle U_{\text{signal}} | U_{\text{error}} \rangle = 0 \). Comparison of definitions (8) and (9) shows that \( \eta_{\text{SRE}} = \eta \) only if \( U_{\text{error}} = 0 \). This occurs if the obtained and desired output fields match in both phase and amplitude up to a constant. If only the amplitude of the output is specified, one can make the phases match by stating that the obtained phase is the desired phase, i.e., \( U_{\text{error}} \) is non-zero only if there is an amplitude mismatch.

**D. Signal-to-Noise Ratio**

The fidelity of the signal inside the target window \( W \) can be characterized by means of a signal-to-noise ratio (SNR). By using Eq. (7) we define the SNR as a ratio of the energies of the desired signal and the error field:

\[
\text{SNR} = \frac{\|AU_{\text{signal}}\|^2}{\|U_{\text{error}}\|^2} = \frac{\|\alpha U_{\text{signal}}\|^2}{\|U_{\text{out}} - \alpha U_{\text{signal}}\|^2_{W}}.
\] (10)

This measure of fidelity will be used along with the SRE to evaluate the results.

**E. Projection**

For a given input \( U_{\text{in}} \) and output distribution \( U_{\text{out}} \) and set of allowed transmission values \( A_{\text{c}} \), the SRE is maximized if the following transmission function is chosen [13]:

\[
\tau_{\text{opt}} = \underset{r \in A_{\text{c}}}{\text{argmin}} \lim_{r \to \infty} \left\| t - \frac{L^{-1}\{U_{\text{signal}}\}}{U_{\text{in}}} \right\|^2,
\] (11)

where \( L \) is a linear invertible propagation operator that converts the field after the transmission function into the output field. This transmission function is constructed by projecting the inverse transform of the desired field onto the constraint \( A_{\text{c}} \). If \( A_{\text{c}} \) is the phase-amplitude constraint in Eq. (5), then a given angle \( \theta \) that is projected ends up at

\[
\phi_{\text{proj}} = \begin{cases} 
\theta - \arctan(\kappa/\Delta n) & \text{if } 0 \leq \theta - \arctan(\kappa/\Delta n) \leq \theta_{M} \\
0 & \text{otherwise}
\end{cases}
\] (12)

where the constant \( \theta_{M} \) is determined numerically by solving the equation

\[
\sqrt{1 + (\kappa/\Delta n)^2 \cos[\theta_{M} + \arctan(\kappa/\Delta n)]} \exp(\theta_{M}\kappa/\Delta n) = 1.
\] (13)

A detailed derivation of this equation is given in Appendix A. If \( \kappa/\Delta n = 0 \) this equation simplifies to \( \cos(\theta_{M}) = 1 \) so that the non-trivial solution given by \( \theta_{M} = 2\pi \).

The projection method given by Eq. (12) will be referred to as direct projection when it projects points of a given angle along the shortest path possible; this is shown in Fig. 4(a). On the other hand, Fig. 4(b) shows radial projection; here the phase values after projection are identical to the values before projection.

**3. ALGORITHM**

As already indicated, the goal of the algorithm is to design a transmission function \( t(x, y) \in A_{\text{c}} \) such that the SRE is maximized while a good SNR is maintained. To satisfy these competing criteria the solution will be restricted to the respective constraints at the input and output planes by iterating between them. The goal is reached in two steps:

1. phase-only optimization to maximize SRE, and
2. phase-amplitude optimization to minimize noise.

Here, as in [13], we specify only the amplitude of \( U_{\text{signal}} \) and leave its phase free. Furthermore, we consider designs for beam splitters and diffusers and therefore only periodic diffractive elements. Normalizing the grating period to unity and assuming that the complex amplitude transmission approach is valid, we may represent the transmittance in the form

\[
t(x, y) = \sum_{(m,n)=-\infty}^{\infty} T_{mn} \exp[i2\pi(mx + ny)],
\] (14)

where

\[
T_{mn} = \iint_{0}^{1} t(x,y) \exp[-i2\pi(mx + ny)] dx dy
\] (15)

is the complex amplitude of diffraction order \((m,n)\). With unit-amplitude plane-wave illumination we then have a discrete signal, with \( W \) containing a specified set of orders. In Fig. 2, for example, \( W \) covers the \( 4 \times 4 \) array and the spots outside this array represent the field \( U_{\text{freedom}} \) in Eq. (7). The propagation operator \( L \) in Eq. (6) is taken as the Fourier transform.

**A. Phase Optimization**

The material limits the combinations of phase and amplitude that can exist in the output plane. Therefore finding the design that maximizes the desired amplitudes requires knowledge of the output phase as well. If uniformity were no criteria then this would be equivalent to just maximizing the signal relevant efficiency. This is done by only allowing the phase of the output \( T \) to change. In the absence of absorption this step would result in the same first step as employed in standard IFTA [5].

In the initialization step, the desired signal amplitudes \( D_{mn} \) are assigned to \( T_{mn} \) within \( W \) and random phases are added. Then, at each following iteration step, the output field is subjected to the amplitude constraint

![Fig. 4. Projection of angles onto the (continuous) phase-amplitude constraint for \( \kappa/\Delta n = 0.2 \). (a) Direct projection, the projection is normal to the constraint curve up to the point where the tangent line crosses the curve itself. (b) Radial projection, the projection operation affects the amplitude but not the angle.](image-url)
\[ T_{mn}^* = \begin{cases} 
D_{mn} \exp(i \arg T_{mn}) & \text{when } (m, n) \in \mathbb{W} \\
0 & \text{otherwise} 
\end{cases}, \]

where \( \arg T_{mn} \) are the phases of \( T_{mn} \) obtained at the previous iteration step (and retained by this constraint).

After applying the inverse propagation operator, the resulting transmission function is subjected to contain only the values allowed by \( A_c \). This constraint is applied by re-mapping the phase values such that SRE is maximized. Hence the constraint at the element plane may be written as

\[ t'(x, y) = \exp(i\kappa/\Delta n) \exp[(i - \kappa/\Delta n)\phi_{proj}(x, y)]. \]

The values for \( \phi_{proj} \) are given by Eq. (12), with \( \theta(x, y) \) taken as the polar angle of the complex function \( t(x, y) \) that is projected onto \( A_c \). The phase constant \( \exp(i\kappa/\Delta n) \) ensures that the global phase stays fixed when the projection results in the highest SRE.

The constraints given by Eqs. (16) and (17) are applied iteratively until the output and transmission function no longer change or if the number of iterations exceeds 100. The profile that is computed this way will have a high efficiency (limited, of course, by absorption) but typically a poor SNR due to stagnation in a local extrema. For this reason the algorithms are run multiple times in order to overcome this stagnation.

### B. Phase-Amplitude Optimization

The phase-only design is expected to lie close to the desired ideal design and is used as a starting point for the phase-amplitude optimization. This starting point typically has a high efficiency but very poor uniformity. To improve uniformity the constraints are altered to allow a non-vanishing field \( U_{\text{freedom}} \) outside the signal window:

\[ T_{mn}' = \begin{cases} 
A|D_{mn}| \exp(i \arg T_{mn}) & \text{if } (m, n) \in \mathbb{W} \\
T_{mn} & \text{otherwise} 
\end{cases}, \]

with the constant \( A \) given by

\[ A = \frac{\langle |T||D| \rangle}{\|D\|^2}. \]

The algorithm now iterates and updates the result at the input and output plane in accordance with Eqs. (17) and (18), respectively. The phase-amplitude optimization stops when the SNR exceeds the target value (here \( 10^{10} \)) or if the number of iterations exceeds 100.

A simplified flowchart of this algorithm is presented in Fig. 5.

**Fig. 5.** Flowchart of the algorithm. In both stages the material constraint is applied to maximize SRE as specified in Eq. (12). The constraint at the output is first limited to phase-only optimization [Eq. (16)] and upon stagnation replaced by phase-amplitude [Eq. (18)] to improve uniformity.

### 4. OPTIMUM TRIPLECTOR

To have an idea of how well the algorithm performs, one would ideally compare the results against a known solution. For a grating with three equal-efficiency diffraction orders (a triplicator) the best available solution is known analytically in the phase-only case and is given by [12]

\[ \phi(x) = \arctan[a \sin(2\pi x)], \]

with the constant \( a \in [0, \infty) \) such that \( |T_{-1}| = |T_0| = |T_1| \).

The derivation in [12] can be extended to include phase-dependent absorption; the details are presented in Appendix B. The solution becomes implicit and the optimum profile is

\[ \phi(x) = \begin{cases} 
0 & \text{if } x \in [-R, R] \\
\phi(x) + \phi_c & \text{if } x \in [-R, R] 
\end{cases}, \]

where \( \phi_c \) is the minimum of \( \phi(x) \), \( x \in R \), and \( \phi(x) \) is the solution of the equation

\[ (1 - \alpha)[\cosh(\alpha\phi) - a \cos(2\pi x) \sinh(\alpha\phi)] \sin \phi = (1 + \alpha)[\sinh(\alpha\phi) - a \cos(2\pi x) \cosh(\alpha\phi)] \cos \phi, \]

with the constant \( a \) chosen such that the equalities \( |T_{-1}| = |T_0| = |T_1| \) hold, and \( R \in [0, 1/4] \) is chosen such that the efficiency is maximized.

Some triplicator phase profiles given by Eq. (21) are illustrated in Fig. 6. The role of the \( R \) is clearly seen: as soon as the grating exhibits absorption, a region with zero phase delay emerges, the extent of which widens as the absorption level increases. At high absorption levels the phase profile approaches a binary shape, and hence also the amplitude profile tends toward a binary form. The condition for the uniformity of the three diffraction orders cannot be satisfied at arbitrarily high absorption levels: beyond \( \kappa/\Delta n \approx 0.43 \), Eq. (22) has no solutions. This originates from breaking the required condition \( 0 \leq \phi \leq \pi \) of the triplicator upper bound proof.

**Fig. 6.** Phase profiles of optimum triplicators under different levels of absorption: the solid line shows the optimum profile in the dielectric case (\( \kappa = 0 \)). The dashed lines illustrate the profiles at different absorption levels.
distributions. Each point is a single evaluation of the algorithm, so that Fig. 7 roughly shows the expected outcome of single runs. Clearly, the direct projection method given by Eq. (12) typically results in efficiencies close to the theoretical maximum, though occasionally it fails to give a satisfactory result. Using radial projection results in large variations in performance, and it only rarely gives an efficiency close to the optimum value, indicating that the algorithm should be run many times to obtain a reasonable result.

5. NUMERICAL RESULTS

A. On-Axis Design

If the signal window $W$ contains the zeroth diffraction order of the grating, we talk about an on-axis signal. As soon as absorption is present, the zeroth-order efficiency tends to increase, and the problem becomes increasingly severe when the number of orders $N$ within $W$ grows. The IFTA algorithm can control the zeroth-order efficiency to some extent, but the overall efficiency of the design usually suffers from such suppression.

A rough estimate of the value of $N$ at which the zeroth order becomes a problem can be obtained as follows. In order to get a rough estimate we assume that the phase values are distributed linearly in the $[0, 2\pi]$ interval. However, as we will demonstrate explicitly below, in IFTA designs this is not the case as the phase optimization step tends to avoid phase values with large absorption and favors zero phase due to the projection. Hence the simple analytical results to be given below may be considered (at least slightly) as underestimates.

The fraction $f$ of incident energy transmitted by a grating with $\phi(x) = 2\pi x$ is given by

$$\eta_0 = \left| \frac{1 - \exp(-2\pi\kappa/\Delta n)}{4\pi^2(1 + \kappa^2/\Delta n^2)} \right|^2.$$  \hspace{1cm} (24)

The zeroth order is expected to affect the design when $\eta_0 > f/N$, i.e., when

$$N > f/\eta_0.$$  \hspace{1cm} (25)

Figure 8 shows a plot of $f/\eta_0$ as a function of the absorption level. It gives an estimate for the size of the desired signal for efficiency of the zeroth order of a grating with $\phi(x) = 2\pi x$ is given by

$$\eta_0 = \left| \frac{1 - \exp(-2\pi\kappa/\Delta n)}{4\pi^2(1 + \kappa^2/\Delta n^2)} \right|^2.$$  \hspace{1cm} (24)

The zeroth order is expected to affect the design when $\eta_0 > f/N$, i.e., when

$$N > f/\eta_0.$$  \hspace{1cm} (25)

Figure 8 shows a plot of $f/\eta_0$ as a function of the absorption level. It gives an estimate for the size of the desired signal for
which the zeroth order will start to interfere with the design process. Exceeding this line (greatly) may result in a (large) sacrifice of efficiency in order to keep the zeroth order in control.

Figure 9 shows the on-axis performance of the algorithm when designing a beam splitter with an array of $5 \times 5$ equal-efficiency orders. The figure shows the efficiency and SNR when absorption is ignored during design, when radial projection is used, and when direct projection is used in the design process. Results from several random starting-phase configurations are again shown. Interestingly (and fortunately) the dependence of the efficiency on initial phase distribution is far smaller than in the case of triplicators since now the number of diffraction orders is much greater.

If absorption is ignored, the SNR is poor even at small levels of absorption, as already evidenced in Fig. 2. When $\kappa/\Delta n > 0.4$ the direct projection method can no longer keep the zeroth order under control, resulting in a significant loss in efficiency and SNR. For these high levels of absorption radial projection becomes preferable to direct projection, but all methods perform poorly due to the increasing difficulty of suppressing the zeroth order. In this regime it is beneficial to move the signal off-axis to exclude the zeroth order from $W$.

Figure 10(a) illustrates the phase distribution within one period of a $4 \times 4$ beam splitter with $128 \times 128$ sampling points at the absorption level $\kappa/\Delta n = 0.1$. Only phase values in the range $0 \leq \phi < 1.65\pi$ remain in the design because of the projection, and there are several two-dimensional regions in which the phase is zero. The distribution of phase values is shown more quantitatively in Fig. 10(b), where the phase is quantized into 100 equally spaced intervals and each bar shows the value $N$ of phase values with such an interval. Here we have considered an average of 100 (high-efficiency) designs to obtain an expectation value distribution and, for clarity, left out the value $N = 2900$ corresponding to $\phi = 0$. It visually shows how the phase values are redistributed to accommodate for both the design and the imposed constraint. The peaks around $\phi = 0$ and $\phi = \pi$ are the most efficient way to control the zeroth order without greatly sacrificing efficiency of the design.

B. Off-Axis Design

To move off-axis can be desired in situations where the zeroth diffraction order is of no interest, undesired, or if on-axis design interferes with design constraints. The latter can occur when designing a multi-level grating with a non-symmetric design, if on-axis design does not achieve the desired signal-to-noise ratio or if absorption blows up the zeroth order. These cases would result in a severe reduction in efficiency as the algorithm tries to improve uniformity in the face of these deficiencies.

The expected efficiency of an off-axis signal ($\eta_{\text{Abs}}$) can be estimated by finding the (upper bound) efficiency $\eta_{\text{noAbs}}$ for the on-axis signal when no absorption is present [14] and multiplying it with the efficiency $\eta_{\text{Carrier}}$ of the carrier grating:

$$
\eta_{\text{Abs}} = \eta_{\text{noAbs}} \times \eta_{\text{Carrier}}
$$

Fig. 10. (a) A phase map of a typical $4 \times 4$ beam splitter operating on-axis, designed by direct projection. (b) An expectation value histogram of the phase values. The value of $N$ at $\phi = 0$, which is 2900 in the shown scale, is left out for clarity.

Fig. 11. Same as Fig. 9, but for an off-axis beam splitter with the $5 \times 5$ array centered at $(m, n) = (15, 15)$. 
The carrier grating is now the single-point-signal design at the given absorption level and its efficiency is discussed in detail in [13] (see, in particular, Fig. 9 of that paper). The estimate in Eq. (26) becomes increasingly accurate when the signal is moved further away from the axis.

Figure 11 shows the performance of the algorithms for the off-axis design of a 5 × 5 beam splitter, with the signal window centered at diffraction order \((m, n) = (15, 15)\). For this task the direct projection outperforms the other methods in terms of both efficiency and SNR. From \(\kappa/\Delta n \approx 0.2\) onward the off-axis designs outperform on-axis designs for both efficiency and SNR. The improvement in efficiency comes from no longer needing to suppress the zeroth order. The results in Fig. 11 show that we can actually go somewhat above the line in Fig. 8, which predicts \(\kappa/\Delta n \approx 0.15\) for \(N = 25\). The line shows the result of Eq. (26), with the upper bound \(\eta_{\text{noAbs}} = 0.93\) multiplied by the carrier-grating efficiency \(\eta_{\text{Carrier}}\) given by the solid line in Fig. 9 of [13]. The best results given by the direct projection method exceed this line only slightly, demonstrating that Eq. (26) in indeed a good approximation for efficiencies that are attainable in off-axis design.

6. CONCLUSIONS
We have extended the iterative Fourier transform algorithm so as to enable the design of diffractive elements made of lossy materials that exhibit phase-delay-dependent absorption. This scenario is relevant in several spectral regions, especially outside the visible domain. An algorithm based on the recently introduced concept of signal relevant efficiency and its upper bound was presented and its functionality was tested against analytical results on triplicators, and for the design of both on-axis and off-axis grating beam splitters, in the presence of varying levels of absorption. We expect that the presented approach will prove useful in optical design at wavelengths ranging from the x-ray regime to the far infrared.

APPENDIX A: SIGNAL RELEVANT EFFICIENCY (SRE) PROJECTION
In this appendix we determine the SRE projection operator for a given level of absorption \(\kappa/\Delta n\). This operator is determined by the shortest projection path onto the constraint.

The constraint onto which the field at the element plane is to be projected is given by

\[
A_c(\phi) = \exp[(i - \kappa/\Delta n)\phi], \tag{A1}
\]

where \(\phi \in [0, 2\pi]\). Suppose that a distribution of angles \(\theta \in [0, 2\pi]\) that lie at \(\lim_{r \to \infty} r \exp(i\theta)\) are projected onto Eq. (A1) in order to satisfy this constraint. Then, for an unbound monotonically increasing or decreasing function, the shortest projection to that function coincides with its normal. For the constraint in Eq. (A1), the normal at point \(\phi\) goes out at angle

\[
\theta = \phi + \arctan(\kappa/\Delta n). \tag{A2}
\]

Because Eq. (A1) is bound to phase values \(\phi \in [0, 2\pi]\), an edge case arises that should be handled separately. This edge case occurs at the discontinuity at \(\phi = 0\) and, as the constraint function decreases monotonically, a finite range of \(\theta\) values will be projected onto the single point \(\phi = 0\).

To find the point at which phase values should start to be projected onto \(\phi = 0\) we introduce parameters

\[
x(\phi) = \Re\{A_c(\phi)\} = \exp(-\kappa/\Delta n)\phi \cos \phi \]
\[
y(\phi) = \Im\{A_c(\phi)\} = \exp(-\kappa/\Delta n)\phi \sin \phi, \tag{A3}
\]

which describe Eq. (A1) in Cartesian coordinates. Hence the angle of the normal at \(\phi = 0\) is given by

\[
\theta_0 = \arctan\left[\frac{x(0)}{y(0)}\right] = \arctan(\kappa/\Delta n), \tag{A4}
\]

so that for values \(0 \leq \theta < \theta_0\) the resulting projection should be \(\phi = 0\).

A second edge case is illustrated in Fig. 12, which shows the tangential line that goes through points \(A_c(\phi_M)\) and \(A_c(0)\), denoted by A and B, respectively. Any value \(\phi_M < \theta < 2\pi\) (all points at infinity with a negative imaginary value that lie to the right of the normal passing through A) should also be projected onto \(\phi = 0\).

Since \(\phi_M\) is the value of the tangent at \(A_c(\phi)\) that goes through \(A_c(0)\), it should satisfy

\[
\frac{\partial x}{\partial \phi}_{\phi_M} + X_0 = 1
\]
\[
\frac{\partial y}{\partial \phi}_{\phi_M} + Y_0 = 0, \tag{A5}
\]

where \(r\) is a running parameter and \(X_0 = \Re\{A_c(\phi_M)\}\) and \(Y_0 = \Im\{A_c(\phi_M)\}\) are the coordinates of point A in Fig. 12. These conditions lead to the expressions

Fig. 12. Inward spiral shows the constraint of Eq. (A1) with \(\kappa/\Delta n = 0.2\). The dashed line shows the tangent that goes through the point \(1 + 0i\), the line perpendicular to this represent normals to this, and the line originating from \(1 + 0i\) is inclined at an angle \(\arctan(\kappa/\Delta n)\).
\[
\exp\left(-\frac{\kappa}{\Delta n} \phi_M \right) \left(-\frac{\kappa}{\Delta n} \cos \phi_M - \sin \phi_M \right) t
+ \exp\left(-\frac{\kappa}{\Delta n} \phi_M \right) \cos \phi_M = 1,
\]
\[
\exp\left(-\frac{\kappa}{\Delta n} \phi_M \right) \left(-\frac{\kappa}{\Delta n} \sin \phi_M + \cos \phi_M \right) t
+ \exp\left(-\frac{\kappa}{\Delta n} \phi_M \right) \sin \phi_M = 0, \tag{A6}
\]
which can be combined to eliminate the parameter \(t\). Doing so, we have
\[
1 + \frac{\kappa}{\Delta n} \exp\left(\frac{\kappa}{\Delta n} \phi_M \right) \left(\cos \phi_M + \sin \phi_M \right) = 0 \tag{A7}
\]
or
\[
\sqrt{1 + (\kappa/\Delta n)^2 \cos[\phi_M + \arctan(\kappa/\Delta n)]} \exp\left(\frac{\kappa}{\Delta n} \phi_M \right) = 1. \tag{A8}
\]

This equation cannot be reduced any further and therefore \(\phi_M\) is computed numerically from this equation.

From Eq. (A2) we find that the angle of the normal of the constraint curve at \(\phi = \phi_M\) is
\[
\theta_M = \phi_M + \arctan(\kappa/\Delta n). \tag{A9}
\]

The results from Eqs. (A2), (A4), and (A9) show that the shortest projection onto Eq. (A1) is indeed given by Eq. (12).

**APPENDIX B: ANALYTICAL DESIGN OF ABSORPTIVE TRIPLECTORS**

In this appendix we determine the phase profile that results in the maximum efficiency of the triplicator (three diffraction orders with equal efficiency) when one assumes various levels of phase-dependent absorption. The derivation proceeds analogously to that presented in [12]. However, because of the presence of absorption, the derivation is a bit more elaborate.

The solution must be consistent with the phase-dependent absorption profile defined in Eq. (A1), i.e.,
\[
t(x) = \exp\left[(i - \alpha)\phi(x)\right], \tag{B1}
\]
with \(\phi \in [0, 2\pi]\). For brevity of notation, we have defined here \(\alpha = \kappa/\Delta n\). Following [12], we separate the phase function \(\phi(x)\) into even and odd parts \(\phi = \phi_e + \phi_o\), with \(\phi_o \in [-\pi, \pi]\) and suppress the \(x\)-dependence (for brevity sake this notation will be used for most of the deviation). Doing this separation gives
\[
t(x) = \exp\left[(i - \alpha)\phi_e(x)\right] \left[\cosh(\alpha\phi_o) - \sinh(\alpha\phi_o)\right]
\times \left[\cos(\phi_o) + i \sin(\phi_o)\right]. \tag{B2}
\]

The complex amplitude of \(n\)th diffraction order of this transmittance profile is given by
\[
T_n = \int_{-1/2}^{1/2} t(x) \exp(-i2\pi nx) \, dx. \tag{B3}
\]

We are interested in equalizing the efficiencies of orders \(n = 0\) and \(n = \pm 1\). In view of Eq. (B3),
\[
T_0 = \left. \int_{-\pi}^{\pi} \exp\left[(i - \alpha)\phi_e(x)\right] \left[\cosh(\alpha\phi_o) - \sinh(\alpha\phi_o)\right]
\times \left[\cos(\phi_o) + i \sin(\phi_o)\right] \, dx \right|_{\phi_e = 0} \tag{B4}
\]
and
\[
T_1 = \left. \int_{-\pi}^{\pi} \exp(-2\pi x) \, dx \right|_{\phi_e = 0} \tag{B5}
\]

In these expressions we have introduced the region \(R\) in which \(\phi_e(x) + \phi_o(x) = 0\).

Let us continue by expanding Eq. (B5) and removing integrals over odd functions. This yields
\[
T_0 = \left. \int_{-\pi}^{\pi} \exp\left[(i - \alpha)\phi_e(x)\right] \left[\cosh(\alpha\phi_o) - \sinh(\alpha\phi_o)\right]
\times \left[\cos(\phi_o) + i \sin(\phi_o)\right] \, dx \right|_{\phi_e = 0} \tag{B6}
\]
and
\[
T_1 = \left. \int_{-\pi}^{\pi} \exp(-2\pi x) \, dx \right|_{\phi_e = 0} \tag{B7}
\]

There are two ambiguities in a periodic phase profile, namely, a constant phase factor and a spatial shift. The phase ambiguity is fixed by demanding that \(T_1 = -T_{-1}\) so that Eq. (B7) takes the form
\[
T_1 = \left. \int_{-\pi}^{\pi} \exp\left[(i - \alpha)\phi_e(x)\right] \sinh(\alpha\phi_o) \, dx \right|_{\phi_e = 0} \tag{B8}
\]

The second ambiguity is fixed by demanding that the function is centered and minimal at \(x = 0\). To impose this the integration domain is set to \(x \in [-1/2, 1/2]\) and the functions in Eqs. (B5) and (B7) are shifted by \(-\pi/2\) in the \(x\)-coordinate, which yields
\[
T_0 = 2R + \left. \int_{-\pi}^{\pi} \exp\left[(i - \alpha)\phi_e(x)\right] \left[\cosh(\alpha\phi_o) - \sinh(\alpha\phi_o)\right]
\times \left[\cos(\phi_o) + i \sin(\phi_o)\right] \, dx \right|_{\phi_e = 0} \tag{B9}
\]
and
\[
T_1 = \left. \int_{-\pi}^{\pi} \exp\left[(i - \alpha)\phi_e(x)\right] \cosh(\alpha\phi_o) \, dx \right|_{\phi_e = 0} \tag{B10}
\]

Taking the absolute values of these functions shows that they are bound by
\[ |T_0| \leq 2R + 2 \int_R^{1/2} \exp(-\alpha \phi_c) \times [\cosh(\alpha \phi_o) \cos \phi_o + |\sinh(\alpha \phi_o) \sin \phi_o|] \, dx \] (B11)

and

\[ |T_1| \leq \sin(2\pi R / \pi) + 2 \int_R^{1/2} \exp(-\alpha \phi_c) \cos(2\pi x) \times [\cosh(\alpha \phi_o) \sin \phi_o + |\sinh(\alpha \phi_o) \cos \phi_o|] \, dx, \] (B12)

respectively.

### 1. Proof for Equality

We proceed to show that the \(\leq\) sign in Eqs. (B11) and (B12) can in fact become an equality sign under the right circumstances. To show this we introduce the following notation:

\[ I_1 = I_p + I_Q, \] (B13)

\[ |I_2| = |I_p| + I_Q, \] (B14)

\[ I_p = \int_0^{1/2} |P| \exp(i\phi_p) \, dx, \] (B15)

\[ I_p^2 = \int_0^{1/2} |P| \, dx, \] (B16)

where \(\phi_p\) is the phase of a function \(P(x)\). Since

\[ |I_1|^2 = |I_p|^2 + |I_Q|^2 + 2|I_p||I_Q|, \] (B17)

\[ |I_2|^2 = I_p^2 + I_Q^2 + 2|I_p||I_Q|, \] (B18)

a solution of \(|I_1|^2 = |I_2|^2\) is possible if and only if

\[ |I_p|^2 = I_p^2 \quad \text{and} \quad |I_Q|^2 = I_Q^2. \] (B19)

To discover the conditions under which these equalities can be satisfied, we express the functions \(|I_p|^2\) and \(I_p^2\) as Riemann sums with integration step size \(\delta\), so that

\[
|I_p|^2 = \delta^2 \sum_{k=1}^N |P_k|^2 + \delta^2 \sum_{k=1}^N \sum_{h \neq k} |P_k||P_h| \exp[i(\phi_{p,k} - \phi_{p,h})]
\]

\[
= \delta^2 \sum_{k=1}^N |P_k|^2 + 2\delta^2 \sum_{k=1}^N \sum_{h \neq k} |P_k||P_h| \cos(\phi_{p,k} - \phi_{p,h})
\] (B20)

and

\[
I_p^2 = \delta^2 \sum_{k=1}^N |P_k|^2 + 2\delta^2 \sum_{k=1}^N \sum_{h \neq k} |P_k||P_h|. \] (B21)

Letting \(\delta \to 0\), we see that \(|I_1|^2 = |I_2|^2\) can be true only if the conditions

\[ \cos(\phi_{p,k} - \phi_{p,h}) = 1 \quad \forall \ (h, k), \] (B22)

\[ \cos(\phi_{p,k} - \phi_{p,h}) = 1 \quad \forall \ (h, k) \] (B23)

hold. In other words, the phase functions of \(P(x)\) and \(Q(x)\) must be constant.

Applying this knowledge one finds, on substituting into Eq. (B12) the quantities

\[ P = \exp[(i - \alpha)\phi_e] \cosh(\alpha \phi_o) \cos \phi_o, \] (B24)

\[ Q = -i \exp[(i - \alpha)\phi_e] \sin(\alpha \phi_o) \sin \phi_o, \] (B25)

the phase functions

\[ \phi_p = \phi_e + \text{arg} \ |\phi_o| = \phi_e + \phi_{\cos}, \] (B26)

\[ \phi_q = \phi_e + \text{arg} \ |\sinh(\alpha \phi_o)| \sin \phi_o + 3\pi / 2 = \phi_e + \phi_{\sinh} + 3\pi / 2. \] (B27)

Therefore the statement

\[ |T_0| = 2R + \int_R^{1/2} \exp(-\alpha \phi_c) \times [\cosh(\alpha \phi_o) \cos \phi_o + |\sinh(\alpha \phi_o) \sin \phi_o|] \, dx \] (B28)

can only be true if

\[ \phi_e + \phi_{\cos} = \Phi_{c1}, \] (B29)

\[ \phi_e + \phi_{\sinh} + 3\pi / 2 = \Phi_{c2}, \] (B30)

with \(\Phi_{c1}\) and \(\Phi_{c2}\) being constants.

Likewise, if one substitutes into Eq. (B12) the quantities

\[ P = \exp[(i - \alpha)\phi_e] \cosh(\alpha \phi_o) \sin \phi_o \cos(2\pi x), \] (B31)

\[ Q = \exp[(i - \alpha)\phi_e] \sin(\alpha \phi_o) \cos \phi_o \cos(2\pi x), \] (B32)

so that

\[ \phi_p = \phi_e + \text{arg} \ |\phi_o| = \phi_e + \phi_{\sin}, \] (B33)

\[ \phi_q = \phi_e + \text{arg} \ |\sinh(\alpha \phi_o)| \cos \phi_o + \pi / 2 = \phi_e + \phi_{\sin} \cos + \pi / 2, \] (B34)

then the statement

\[ |T_1| = \sin(2\pi R / \pi) + 2 \int_R^{1/2} \exp(-\alpha \phi_c) \cos(2\pi x) \times [\cosh(\alpha \phi_o) \sin \phi_o + |\sinh(\alpha \phi_o) \cos \phi_o|] \, dx \] (B35)

can only be true if

\[ \phi_e + \phi_{\sin} = \Phi_{c3}, \] (B36)

\[ \phi_e + \phi_{\sin} \cos + \pi / 2 = \Phi_{c4}, \] (B37)

with \(\Phi_{c3}\) and \(\Phi_{c4}\) being constants.

To have an equality sign in Eq. (B12) we have to ensure that, in the integration domain \(x \in [0, 0.5]\), the following angles are equal: \(\phi_{\cos} = \phi_{\sin} = \phi_{\sinh} = \phi_{\sinh}, i.e., \) they are all either positive or negative valued; then also \(\phi_e\) will be constant.

This restriction can be reduced to requiring that \(\phi_{\cos} = \phi_{\sin}\) or more simply \(\max(\phi_e + \phi_o) - \min(\phi_e + \phi_o) \leq \pi\). If this condition is violated then Eq. (B12) will result in a strict inequality and the ideal profile cannot be retrieved via the shown proof.

### 2. Optimum Triuplicator Profile

Now that the conditions that are required to find the optimum triuplicator are known, they can be applied to find the associated phase profile.

Assuming that \(\text{arg} \ |\phi_o| = \text{arg} \ |\phi_o| = \phi_{\cos}\) and \(\phi_e = \Phi_e\), Eqs. (B11) and (B12) take the forms
and

\[ T_0 = 2R + 2 \int_R^{1/2} \exp(-\alpha f_\circ) \exp(i\phi_\circ) \times \left[ \cosh(\alpha f_\circ) \cos \phi_\circ + \sinh(\alpha f_\circ) \sin \phi_\circ \right] dx, \quad (B38) \]

and

\[ T_1 = \sin(2\pi R) / \pi + 2 \int_R^{1/2} \exp(-\alpha f_\circ) \exp(i\phi_\circ) \cos(2\pi x) \times \left[ \cosh(\alpha f_\circ) \sin \phi_\circ + \sinh(\alpha f_\circ) \cos \phi_\circ \right] dx. \quad (B39) \]

Our goal is to maximize the function \( F(\phi_\circ) = |T_0| + a |T_1| \) with \( a \) being an as-yet undetermined constant. At the maximum the function should satisfy the variational condition

\[ \delta F = \lim_{\epsilon \to 0} [F(\phi_\circ + \epsilon) - F(\phi_\circ)] = 0. \quad (B40) \]

By expanding the function

\[ F(\phi_\circ + \epsilon) \approx 2R + a \sin(2\pi R) / \pi + 2 \int_R^{1/2} \exp(-\alpha f_\circ) \times \left[ \cosh(\alpha f_\circ) \cos \phi_\circ + \sinh(\alpha f_\circ) \sin \phi_\circ \right] dx \]

\[ - 2a \int_R^{1/2} \exp(-\alpha f_\circ) \cos(2\pi x) \times \left[ \cosh(\alpha f_\circ) \sin \phi_\circ + \sinh(\alpha f_\circ) \cos \phi_\circ \right] dx \]

into a first-order Taylor series we obtain the approximation

\[ F(\phi_\circ + \epsilon) \approx 2R + a \sin(2\pi R) / \pi + 2 \int_R^{1/2} \exp(-\alpha f_\circ) \times \left[ \cosh(\alpha f_\circ) \cos \phi_\circ + \sinh(\alpha f_\circ) \sin \phi_\circ \right] dx \]

Calculating \( \delta F \) then gives

\[ \delta F = 2\epsilon \int_R^{1/2} \exp(-\alpha f_\circ) \times \left[ - \cosh(\alpha f_\circ) \sin \phi_\circ + \sinh(\alpha f_\circ) \cos \phi_\circ \right] \]

\[ + \sinh(\alpha f_\circ) \cos \phi_\circ + \alpha \cosh(\alpha f_\circ) \sin \phi_\circ \cos(2\pi x) \cosh(\alpha f_\circ) \sin \phi_\circ \cos \phi_\circ \]

\[ - \sinh(\alpha f_\circ) \sin \phi_\circ + \alpha \cosh(\alpha f_\circ) \cos \phi_\circ \cos \phi_\circ \right] dx = 0. \quad (B43) \]

Reordering this and dropping the integration yields

\[ (1 - \alpha)\cosh(\alpha f_\circ) - a \cos(2\pi x) \sinh(\alpha f_\circ) \sin \phi_\circ \]

\[ = (1 + \alpha)\sinh(\alpha f_\circ) - a \cos(2\pi x) \cosh(\alpha f_\circ) \cos \phi_\circ. \quad (B44) \]

At this point three constants need to be determined: first the region \( R \), second the phase \( \phi_c \), and finally the value of \( a \).

The phase constant \( \phi_c \) needs be fixed under the constraint that \( \phi_c + \phi_\circ \in [0, 2\pi] \) such that the efficiency is maximized. To do this \( -\phi_c \) must be the minimum value of \( \phi_\circ(x) \), \( x \in [R, 1/2] \), as any other value would reduce the efficiency. This choice will simplify the notation where the proof is valid from \( \max(\phi_c + \phi_\circ) - \min(\phi_c + \phi_\circ) \leq \pi \) to \( 0 \leq \phi(x) \leq \pi \). The phase function is therefore given by Eq. (21), i.e.,

\[ \phi(x) = \begin{cases} 0 & \text{if } x \in [-R, R] \\ \phi(x) + \phi_c & \text{if } x \in [-R, R] \end{cases} \quad (B45) \]

To obtain the solution of this equation, the value \( a \) must be chosen such that \( |T_0| = |T_1| \), and \( R \) must be chosen such that the efficiency is maximized. It can be shown that in the purely dielectric case \( a = 0 \) and \( R; \{0\} \), Eq. (B44) reduces to

\[ \sin \phi_\circ = -a \cos(2\pi x) \cos \phi_\circ, \quad (B46) \]

yielding

\[ \phi_\circ = -\arctan[a \cos(2\pi x)]. \quad (B47) \]

This is the same profile as Eq. (20) (and the same as given in [12]), be it shifted to have its minimum at zero.

**Funding.** European Union’s Seventh Framework Programme (FP7) (2007–2013) (PITN-GA-2013-608082); Academy of Finland (285280).