Upper bound of signal-relevant efficiency of constrained diffractive elements

Verhoeven AD

info:eu-repo/semantics/article
info:eu-repo/semantics/publishedVersion
© Optical Society of America
© 2016 Optical Society of America. Users may use, reuse, and build upon the article, or use the article for text or data mining, so long as such uses are for non-commercial purposes and appropriate attribution is maintained. All other rights are reserved.
http://dx.doi.org/10.1364/JOSAA.33.002425

https://erepo.uef.fi/handle/123456789/6165

Downloaded from University of Eastern Finland's eRepository
We define the signal-relevant efficiency (SRE) of a diffractive optical element as a measure of the proportion of the incident field power that ends up in the desired output signal. An upper bound for SRE is determined in the presence of arbitrary constraints imposed on the element, such as phase-dependent loss due to absorption within the microstructure and quantization of the surface profile. We apply the theory to the important class of diffractive elements that contain only one desired diffraction order (such as diffractive lenses) and derive the surface profile that provides the highest efficiency allowed by the constraints.

© 2016 Optical Society of America


http://doi.org/10.1364/JOSAA.33.002425

1. INTRODUCTION

The upper bound \( \eta_u \) of diffraction efficiency [1,2] is one of the key concepts in the paraxial theory of diffractive optics [3,4]. It can be easily calculated for any fixed complex-amplitude signal, without designing the element that generates the signal, and can never be exceeded by the efficiency \( \eta \) of an actually designed element. Apart from the one-point signal, \( \eta_u < 100\% \), and \( \eta = \eta_u \) only for one-point and two-point signals. The upper bound is a useful concept also if only the intensity of the signal is fixed and its phase is free: the design is good only if \( \eta \) is close to \( \eta_u \) computed for the resulting complex-amplitude distribution of the signal. Methods for finding upper bounds have been presented for discrete [5] and continuous [6] signals in the case of this phase freedom being available. Other attempts have also been developed to define the upper bound [7,8]. Finally, it should be noted that the statements made above apply only to scalar signals. In the (paraxial) electromagnetic case, where the element modulates the polarization state of incident light and the state of polarization of the signal is free, \( \eta = \eta_u = 100\% \) is possible for many but not all signals [9].

In this paper, we present a generalization of the upper bound theorem [1] in a form that is applicable to thin paraxial-domain diffractive elements with arbitrary constraints and output fields, within scalar theory. The discussion is based on [10]. In particular, we apply the theory to elements that exhibit phase-dependent absorption and produce single- and two-point signals. Examples of such elements include transmission-type diffraction gratings and diffractive lenses with the modulated region made of a lossy material. In the special case of single- and two-point signals, i.e., interested in single or two diffraction order output, our approach actually allows not only the determination of the upper bound of efficiency but also the design of a surface profile that produces the highest possible efficiency under the given constraints.

We begin with the problem description in Section 2, where we also define the general concept of signal-relevant efficiency (SRE). In Section 3 we proceed to derive the new upper bound theorem for SRE. The theory is then applied to elements with phase-dependent loss, discussed in Section 4 on a general level. In Section 5 we apply the new upper bound theory to elements that generate a single-point signal. Finally, some conclusions are drawn in Section 6.

2. PROBLEM DESCRIPTION AND KEY DEFINITIONS

For the task ahead we require that two assumptions hold and consider these for transmission-type diffraction elements. The first assumption is that the thin-element approximation can be applied and that the scalar theory is adequate to describe the element, i.e., no space-variant polarization modulation takes place. If we denote by \( U_{in}(x, y) \) the field that illuminates the element with transmission function \( t(x, y) \), the field after the element is given by

\[
U_t(x, y) = t(x, y) U_{in}(x, y),
\]

where \( t(x, y) \) is the transmission function of the element.
where for small angles the transmission function \( t(x, y) \) can be written as function of the height profile \( h(x, y) \), the wavelength \( \lambda \), and the refractive index \( n \) that is used:

\[
\begin{align*}
  t(x, y) &= \exp \left[ \frac{2\pi(n - 1)}{\lambda} h(x, y) \right].
\end{align*}
\] (2)

The second assumption is that the field at the output plane (with coordinates \( x \) and \( y \)) can be obtained by applying a linear invertible operator \( L \) to the field after the element:

\[
U_{\text{out}}(x,y) = L(U_i(x,y)).
\] (3)

The general system that is used in Eq. (3) is shown in Fig. 1. The form of the linear operator \( L \) depends on the physical situation: \( L \) may be, e.g., a Fourier transform, a Fresnel propagation operator, or the Collins operator describing a paraxial lens system [11].

Let us consider optical fields within the Hilbert space \( \mathcal{L}^2_2(\mathbb{R}^2) \) of functions \( U : \mathbb{R}^2 \to \mathbb{C} \) with finite energy. We thereby define the usual inner product as:

\[
\langle U_1|U_2 \rangle := \int_\mathbb{R}^2 U_1(x)\overline{U_2(x)}\,dx,
\] (4)

with the corresponding norm

\[
\|U\| = \sqrt{\langle U|U \rangle}.
\] (5)

Our basic task is to find a diffractive element that would produce the specific output field distribution \( U_{\text{desired}} \) that lies in a confined region in the \((\tilde{x}, \tilde{y})\) space, called the signal window \( W \). Depending on the constraints imposed on the element, some sacrifices must nearly always be made. These constraints may arise from a variety of reasons, the most notable being fabrication restrictions imposed on the height profile. However, regardless of the imposed constraints, one can compose the output field as a combination of the desired field \( U_{\text{desired}} \) with some error field \( U_{\text{error}} \), along with a field outside the signal window, termed \( U_{\text{freedom}} \). Let \( \langle U_{\text{desired}} \rangle = \mathbb{C}U_{\text{desired}} \) denote the one-dimensional subspace spanned by the desired field \( U_{\text{desired}} \). This subspace is defined by orthogonal projection to its complement \( \langle U_{\text{desired}} \rangle^\perp \) of all functions \( U \in \mathcal{L}^2_2(\mathbb{R}^2) \) with \( \langle U|U_{\text{desired}} \rangle = 0 \). The subspace \( \langle U_{\text{desired}} \rangle^\perp \) may again be divided into the subspaces \( \mathcal{H}_{\text{error}} \) of functions identical to zero outside the signal window and \( \mathcal{H}_{\text{freedom}} \) of functions identical to zero within the signal window. Thus, the Hilbert space \( \mathcal{L}^2_2(\mathbb{R}^2) \) may be divided into these orthogonal subspaces:

\[
\mathcal{L}^2_2(\mathbb{R}^2) = \langle U_{\text{desired}} \rangle \oplus \mathcal{H}_{\text{error}} \oplus \mathcal{H}_{\text{freedom}}
\] (6)

and the desired output field can be uniquely decomposed into

\[
U_{\text{out}} = \alpha U_{\text{desired}} + U_{\text{error}} + U_{\text{freedom}},
\] (7)

where \( U_{\text{error}} \in \mathcal{H}_{\text{error}}, U_{\text{freedom}} \in \mathcal{H}_{\text{freedom}} \) and \( \alpha U_{\text{desired}} \) is the orthogonal projection of \( U_{\text{out}} \) onto the subspace spanned by \( U_{\text{desired}} \) with

\[
\alpha = \frac{\langle U_{\text{out}}|U_{\text{desired}} \rangle}{\langle U_{\text{desired}}|U_{\text{desired}} \rangle} = \frac{\langle U_{\text{out}}|L^{-1}U_{\text{desired}} \rangle}{\langle U_{\text{desired}}|U_{\text{desired}} \rangle}.
\] (9)

This identity is the result from projections being invariant under linear operators in Hilbert space.

We define the signal-relevant efficiency \( \eta_{\text{SRE}} \) to be the proportion of the incident field power that ends up in the desired output signal:

\[
\eta_{\text{SRE}} := \frac{\|\alpha U_{\text{desired}}\|^2}{\|U_{\text{in}}\|^2} = \frac{\|U_{\text{out}}|U_{\text{desired}}\|^2}{\|U_{\text{in}}\|^2\|U_{\text{desired}}\|^2}.
\] (10)

The SRE, as defined here, can take any (real) value between zero and unity. It can reach values close to unity if and only if

\[
\frac{\|U_{\text{out}}\|^2}{\|U_{\text{in}}\|^2} = \frac{\|U_{\text{desired}}\|^2}{\|U_{\text{in}}\|^2}.
\] (12)

It is this representation that we will use to determine how close one can get to a desired field distribution, i.e., to determine the upper bound for the SRE.

### 3. Upper Bound Theory

It is assumed that the transmission function \( t \) is limited to an (arbitrary) set of values \( A \), so that \( t(x, y) \in A, \forall (x, y) \). The task is to find the SRE upper bound for a given desired field \( U_{\text{desired}} \) under this constraint for \( t \).

From Eqs. (9) and (11) it follows that the upper bound is given by the output field \( U_{\text{out}} \) that will have largest projection onto the subspace spanned by the desired field. This particular output field will be denoted by \( U_{\text{opt}} \) and its projection is abstractly shown in Fig. 2. Therefore, by definition the upper bound for the SRE is given by

\[
\eta_{\text{SRE}}^\text{max} := \frac{\|U_{\text{opt}}|U_{\text{desired}}\|^2}{\|U_{\text{in}}\|^2\|U_{\text{desired}}\|^2} = \frac{\|L(U_{\text{in}}|t_{\text{opt}})|U_{\text{desired}}\|^2}{\|U_{\text{in}}\|^2\|U_{\text{desired}}\|^2}.
\] (13)

The projection of the output field \( U_{\text{out}} \) onto \( U_{\text{desired}} \) is independent on the norm of the latter, including the limit

\[
\lim_{r \to \infty} r|U_{\text{desired}}|^2.
\] (15)
The field with largest projection onto $U_{\text{desired}}$ coincides with the one closest to Eq. (15) as graphically shown in Fig. 3:

$$t_{\text{opt}} := \arg \max_{r \in A_c} \frac{\| (L_t U_{\text{in}}) | U_{\text{desired}} \|}{\| U_{\text{in}} \|}^2 \| U_{\text{desired}} \|$$  \hfill (16)

$$= \arg \min_{r \in A_c} \lim_{r \to \infty} \| L_t U_{\text{in}} - r U_{\text{desired}} \|^2. \hfill (17)$$

This equation cannot be directly evaluated without running through all possible permutations of $t \in A_c$. Using that the smallest norm is also the smallest after applying the linear operator, Eq. (17) can be rewritten as

$$t_{\text{opt}} = \arg \min_{r \in A_c} \lim_{r \to \infty} \| t - r L_t U_{\text{in}}^{-1} U_{\text{desired}} \|^2. \hfill (18)$$

From where the ideal transmission function $t_{\text{ideal}}$ will be defined as

$$t_{\text{ideal}} := \frac{L_t U_{\text{in}}^{-1} U_{\text{desired}}}{U_{\text{in}}}, \hfill (19)$$

so that Eq. (18) can be reduced to

$$t_{\text{opt}} = \arg \min_{r \in A_c} \lim_{r \to \infty} \| t - r t_{\text{ideal}} \|^2. \hfill (20)$$

This equation can be evaluated point wise, by taking an arbitrary position $(x, y)$ and searching the constrained transmittance value within $A_c$ that lies closest to $t_{\text{ideal}}$. Inserting the resulting transmission function into Eq. (14) yields the SRE upper bound.

It should be noted that the desired output field typically has a constant phase factor $\phi_r \in [0, 2\pi)$ that can be chosen arbitrarily, i.e., $e^{i\phi_r} U_{\text{desired}}$. For the upper bound the choice of $\phi_r$ does only matter if neither $A_c$ nor the phase values of $t_{\text{ideal}}$ form a rotationally symmetric pattern. In that case the upper bound for the SRE should be evaluated for all $\phi_r$ so that the largest resulting SRE will represent the upper bound.

A flowchart to find the upper bound is shown in Fig. 4. To summarize, first the desired signal distribution is defined in the output plane, along with the linear invertible operator that propagates the field after the transmission function to the output plane. The second step is to define the constraints $A_c$ imposed on the transmission function and compute $t_{\text{ideal}}$ with Eq. (19). The last step is to compute the “optimal” transmission function with Eq. (20) so that the upper bound of the signal relevant efficiency can be determined by inserting $t_{\text{opt}}$ into Eq. (14).

In the specific case that the input is a plane wave ($U_{\text{in}}(x, y) = 1$), the material is phase only ($A_c : \arg t_{\text{ideal}} = 1$) and the linear operator is the Fourier transform, Eqs. (19) and (20) can be rewritten as

$$t_{\text{ideal}} = \mathcal{F}^{-1} U_{\text{desired}}, \quad t_{\text{opt}} = e^{i\arg t_{\text{ideal}}}. \hfill (21)$$

Combining this with the identity given by Eq. (9), Eq. (14) can be reduced to

$$\eta_{\text{SRE}}^\text{max} = \frac{\langle e^{i\arg t_{\text{ideal}}} \rangle | t_{\text{ideal}} |^2}{\| t_{\text{ideal}} \|^2} \hfill (22)$$

$$= \frac{\langle | t_{\text{ideal}} |^2 \rangle}{\| t_{\text{ideal}} \|^2}, \hfill (23)$$

with $\langle | t_{\text{ideal}} | \rangle = \int_\mathbb{R} | t_{\text{ideal}} | dx$ denoting the expectation value of $| t_{\text{ideal}} |$. The resulting upper bound is the same as presented in [1].
4. ELEMENTS WITH PHASE-DEPENDENT LOSS

As an example of constraints \( A_e \) we consider a diffractive element with a periodic surface profile \( b(x, y) \) and refractive-index distribution

\[
\hat{n}(x, y) = \begin{cases} 
\hat{n} & 0 \leq z \leq b(x, y) \\
1 & \text{else}
\end{cases},
\]

where \( \hat{n} \) is given by

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

The transmission function now takes the form

\[
t(x, y) = |t(x, y)| \exp[i\varphi(x, y)],
\]

where the amplitude modulation \(|t|\) and phase delay \(\varphi\) induced by the element are given by

\[
|t(x, y)| = \exp[-(2\pi/\lambda)kb(x, y)],
\]

and

\[
\varphi(x, y) = (2\pi/\lambda)\Delta nb(x, y),
\]

respectively, and \(\lambda\) is the wavelength of light.

The amplitude modulation and the phase delay are now coupled and we have the constraint

\[
A_e(\varphi) = \exp[-(\kappa/\Delta n)\varphi].
\]

Figure 5 illustrates Eq. (29) in the complex plane for various levels of absorption given by the parameter \(\kappa/\Delta n\). If \(\kappa = 0\) there is no absorption and one is limited to phase-only values on the black circle in Fig. 5. If the material absorbs \((\kappa > 0)\), Eq. (29) represents an inward spiral, on which the values of \(t(x, y)\) at any given point \((x, y)\) must be confined to.

5. EXAMPLE: LOSSY SINGLE-ORDER ELEMENTS

Let us now consider thin two-dimensionally periodic diffractive elements (gratings) illuminated by a unit-amplitude plane wave \(U_{in}(x, y) = 1\). Assuming that the output plane is at infinity and considering the paraxial case, we have the following direct and inverse relations between the field immediately after the element and in the Fourier domain:

\[
U_{out}(m, n) = \int \int t(x, y)e^{i2\pi(mx + ny)} \, dx \, dy.
\]

and

\[
U_t(x, y) = t(x, y) = \sum_{(m, n) = -\infty}^{\infty} U_{out}(m, n)e^{-i2\pi(mx + ny)}.
\]

The coefficients \(U_{out}(m, n)\) represent the complex amplitudes of diffraction orders \((m, n)\), and the efficiencies of these orders are given by

\[
\eta_{(m,n)} = |U_{out}(m, n)|^2.
\]

Here we have normalized (without losing generality) the periods in \(x\) and \(y\) directions equal to unity.

We assume in particular that the desired output field consists of only a single diffraction order:

\[
U_{\text{desired}}(m, n) = \delta_{m-n, 1}
\]

where \(\delta\) is the Kronecker delta symbol and \((\hat{n}, \hat{n})\) denote the indices of the desired order. Since the signal window \(W\) now contains only a single point, the signal-to-noise ratio loses its meaning and the transmission function that gives the maximum SRE is also the one that gives the highest \(\eta_{(1,0)}\). With this in mind we can use Eq. (31) to calculate the Fourier inverse of Eq. (33) and insert the result into Eq. (18) to find the transmission function that maximizes the SRE:

\[
\eta_{\text{SRE}}^{\max} = \frac{|\{t_{\text{opt}}(x, y)\}|^2}{\|\{\delta_{m-1, n-1}\}\|^2}.
\]

The transmission function \(t_{\text{opt}}\) that will match the SRE upper bound can be obtained by

\[
t_{\text{opt}} = \arg\min_{t \in A_e} \|t - r e^{-i2\pi(mx + ny)}\|^2.
\]

For every point \((x, y)\) on the circle with an infinite radius \(r \to \infty\), one needs to find the value in \(A_e\) that lies closest to that point. This will determine \(t_{\text{opt}}\) at this particular point. The transmission profile that is constructed in this way will yield the largest possible SRE (and actual efficiency).

In the examples below we look for grating profiles that maximize the efficiency of diffraction order \((m, n) = (1, 0)\); hence \(t_{\text{opt}}(x, y) = t_{\text{opt}}(x)\).

Figure 6 shows a visualization of how \(t_{\text{opt}}\) is obtained for the constraint Eq. (29) with \(\varphi \in [0, 2\pi]\) and \(\kappa/\Delta n = 0.2\). In this figure the black lines connect the circle at infinity to the closest point that lies closest to it in \(A_e\). Note that the two black lines right next to the “gap” are running (almost) parallel, which means that they arrive at almost the same angle from the distribution at infinity. There are no lines in between because there is no point in the transmission constraint \(A_e\) that lies close enough to connect to the distribution at infinity. As a result, the optimum phase profile has no values in the range \(1.5\pi \leq \varphi < 2\pi\), where the absorption is highest; these are replaced by zero phase and thereby a fully transparent “slit” in the lossy grating results. Inserting the found \(t_{\text{opt}}\) into Eq. (34) yields a \(\eta_{\text{SRE}}^{\max} = \eta_{(0,0)} = 0.44\) in this particular case.

We also considered other constraints, such as a phase-only element with \(A_e\) constrained by \(|t(x, y)| = 1\), a quantized phase element with \(A_e = [1, i, -1, -i]\), and a binary-amplitude
The construction of $t_{\text{opt}}$ in these cases is illustrated in Fig. 7. Let us first look at the phase-only constraint, which leads to the well-known result of 100% efficiency with a triangular profile. From the SRE computation point of view, this profile is obtained because for a point at infinity under a certain angle, the closest point in the constraint set lies at the same angle.

When the constraint is that of a four-level quantized phase, $A_c = [1, i, -1, -i]$, the results in the middle of Fig. 7 are obtained. The gaps that are visible in the top middle plot are a mathematical consequence of the projection from infinity. The result is the staircase profile shown in the middle below plot. The binary-amplitude constraint is shown on the right; the well-known binary profile with 50% fill factor is obtained from the SRE considerations.

When the phase-only constraint of the transmission function is quantized to allow $Z$ equally spaced levels, the highest possible efficiency of (off-axis) diffractive elements can be approximated as

$$\eta_{\text{quant}} \approx \text{sinc}(1/Z)^2 \times \eta.$$  \hfill (36)

where $\eta$ is the efficiency before quantization. Figure 8 compares the results of this formula for the blazed grating with $\eta = 1$, with the efficiencies obtained for the single-point signal by the SRE upper bound theory, and an excellent match is obtained.

Finally, we return to the case of phase-dependent loss and consider systematically various levels of absorption $\kappa/\Delta n$. In Fig. 9, the computed results show the efficiency of the gratings obtained by maximizing the SRE. The blazed profile represented the effect of phase-dependent absorption in the efficiency of a blazed grating with a linear phase in the range $[0, 2\pi]$. The efficiency of such a grating approaches zero at high levels of absorption, and becomes inferior to the efficiency (10.13%) of a binary-amplitude grating with a 50% fill factor when $\kappa/\Delta n > 0.48$, where the blazed and binary profile have equal efficiency. The efficiencies obtained by the SRE theory are above those of either the triangular grating or the binary amplitude grating for all finite levels of absorption.

Figure 10(a) shows some of the phase profiles optimized by the SRE theory. At small levels of absorption the results are of the type already seen in Fig. 6, and the width of the transparent gap increases with $\kappa/\Delta n$. With higher levels of absorption the optimum profile starts to lose its sawtooth character and finally approaches the binary case (leading to a binary-amplitude transmission function) in the limit of high absorption.
6. CONCLUSIONS AND OUTLOOK

We have introduced the concept of signal-relevant efficiency and presented a procedure to determine it for an arbitrary signal and any constraint set for the complex transmission properties of the associated diffractive element. In particular, we applied the theory to gratings with a single diffraction order, establishing that the SRE upper bound theory provides the well-known results for several standard constraints. In addition, we demonstrated that the SRE concept can be used to derive optimal grating profiles in the presence of phase-dependent loss.

The upper bound theory provides significant insight into diffractive optical element (DOE) design and can be used to reduce the complexity of these designs. We will investigate and discuss the DOE design for DOEs with phase and amplitude constraints in forthcoming papers.

**Funding.** Seventh Framework Programme (FP7) (PITN-GA-2013-608082).

**Acknowledgment.** The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union’s FP7 under REA Grant Agreement PITN-GA-2013-608082.

**REFERENCES AND NOTES**

8. In Ref. [7] it is claimed that the upper bound formula presented in Ref. [1] is incorrect. However, this claim is wrong and the dimensions in Eq. (19) of Ref. [1] are correct. Moreover, the mathematical form of the definition given by Eq. (19) in Ref. [1] ensures that the upper bound of efficiency is always less than 100%.