A New Type of Coupled Wave Theory Capable of Analytically Describing Diffraction in Polychromatic Gratings and Holograms

David Brotherton-Ratcliffe
Geola Technologies Ltd, Sussex Innovation Centre, Science Park Square, Falmer, East Sussex, BN1 9SB, United Kingdom
dbr@geola.co.uk

Abstract. A new type of coupled wave theory is described which is capable, in a very natural way, of analytically describing polychromatic gratings. In contrast to the well known and extremely successful coupled wave theory of Kogelnik, the new theory is based on a differential formulation of the process of Fresnel reflection within the grating. The fundamental coupled wave equations, which are an exact solution of Maxwell’s equations for the case of the un-slanted reflection grating, can be analytically solved with minimal approximation. The equations may also be solved in a rotated frame of reference to provide useful formulae for the diffractive efficiency of the general polychromatic slanted grating in three dimensions. The new theory is compared with Kogelnik’s theory where extremely good agreement is found for most cases. The theory has also been compared to a rigorous computational chain matrix simulation of the un-slanted grating with excellent agreement for cases typical to display holography. In contrast, Kogelnik’s theory shows small discrepancies away from Bragg resonance. The new coupled wave theory may easily be extended to an N-coupled wave theory for the case of the multiplexed polychromatic grating and indeed for the purposes of analytically describing diffraction in the colour hologram. In the simple case of a monochromatic spatially-multiplexed grating at Bragg resonance the theory is in exact agreement with the predictions of conventional N-coupled wave theory.

1. Introduction
Kogelnik’s coupled wave theory [1] of the volume holographic grating has been very successful. Published in 1969 it still constitutes today a useful tool for the rapid computation of the diffractive response of general sinusoidal volume phase or mixed volume phase-amplitude gratings. However Kogelnik’s theory is not rigorous. To arrive at the fundamental set of first-order coupled partial differential equations starting from Maxwell’s equations, no fewer than five major assumptions must be made. Various authors have offered alternative models [e.g.2-6] or have extended Kogelnik’s work [e.g.6,7].

Moharam and Gaylord [8] reported in 1981 a rigorous coupled wave theory. This theory provided an accurate computational solution to the problem of diffraction from planar gratings. It was shown that Kogelnik’s analytic theory [1] provided a rather good description of the un-slanted reflection grating but that at high index modulations it overestimated somewhat the diffraction
efficiency of the slanted grating. Differences in the transmission grating were shown to appear at a rather lower modulation.

Gratings produced using multilayer dielectric stacks have been traditionally analysed in a rather different manner to the analysis methods used to describe volume holographic gratings. Here Maxwell's equations are solved by analysing the individual Fresnel reflections from each layer in the stack. In 1937 Rouard [9] introduced a method by which the complex electric field amplitude reflected by a general stack could be calculated. In 1950 Abeles [10,11] introduced his highly successful chain matrix formulism which allowed a more efficient calculation of the reflective properties of a stratified medium.

Moharam and Gaylord [12] used the chain matrix model in 1982 to computationally analyse the un-slanted volume reflection holographic grating. The results were compared with both rigorous coupled wave theory [8] and with Kogelnik's theory [1]. Ning [13] also developed a polychromatic coupled wave theory of the volume reflection phase grating and compared numerical calculations of this theory with corresponding chain matrix calculations for the un-slanted case. In addition Diehl and George [14] compared the chain matrix approach to a method based on a solution of the Hill's matrix. Finally Ludman [15] has derived approximate estimations for the bandwidth and diffraction efficiency of the thick slanted transmission phase grating using the concept of Fresnel reflections.

In this paper we present a differential representation of the chain matrix model which is capable of accurately describing the general volume holographic grating (we refer for convenience to this model as the PSM model - short for ‘Parallel Stacked Mirrors’). The model effectively constitutes a new type of coupled-wave theory capable of describing volume phase and mixed phase-amplitude gratings of both the transmission and reflection type. Rather than seeking an approximate solution to the Helmholtz equation as Kogelnik's theory does, the PSM model is based upon a detailed analytic description of the process of Fresnel reflection within the grating.

The PSM theory produces a mixture of analytical expressions and simple differential equations which can be used to effectively characterise the general multi-colour holographic grating. In the limit of zero slant and normal incidence, the basic differential formulation is rigorous. Many of the single-colour formulae which are derived from this formulation give expressions identical or extremely close to Kogelnik's. Like Kogelnik's theory the PSM model can be generalised to an $N$-coupled wave theory. In the simple case of the monochromatic spatially-multiplexed volume reflection grating at Bragg resonance the generalised PSM equations are identical to the $N$-coupled wave equations of Solymar and Cooke [6-7]. For a full and concise mathematical exposition of the PSM theory the reader is referred to [16-18]. Here we shall restrict ourselves to a brief and simplified introduction to PSM using the least possible mathematics.

2. A Brief and Very Simplified Review of Kogelnik's Theory

The starting point for Kogelnik's theory is the Helmholtz equation. This is an equation which describes how light propagates within a medium. Mathematically, and in its most simple form, the Helmholtz equation may be written as

$$\frac{d^2 u}{dy^2} - \gamma^2 u = 0$$  \hspace{1cm} (1)

Here the variable $u$ measures the electric field of the light wave in a given direction, $y$ measures the distance along the path of the light wave and $\gamma$ is a variable which describes the medium and the wavelength of light travelling through that medium. In words, this equation simply states that the rate
at which the rate of change of the electric field changes is proportional to the electric field itself. When the parameter $\gamma$ is a constant - for example in air - then the Helmholtz equation predicts simple oscillatory or wave motion. In optics this is what we call a plane wave. The same equation arises throughout physics whenever wave motion is present.

Kogelnik's theory actually requires a slightly more complicated version of the Helmholtz equation. This is known as the two-dimensional wave equation

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} - \gamma^2 u = 0$$

The only difference here is that our wave is now not limited to travelling in a single straight line but can have a slightly more complex behaviour within the $(x, y)$ plane.

Kogelnik's coupled wave theory starts off by assuming that the propagation parameter, $\gamma$ in equation (2) is composed of two parts. The first and largest part just corresponds to a constant index of refraction. But added to this is a small oscillatory part which models the holographic grating.

Next Kogelnik makes the assumption that an impinging monochromatic plane light wave which illuminates the grating evenly will propagate in a rather special way within the grating. Specifically it is assumed that there will always be two waves within the grating - the driving wave, $R$ and a diffracted wave, $S$. Both these waves are mathematically written as standard plane waves. But their amplitude is assumed to vary only with their depth of penetration within the grating (see Fig.1).

![Fig.1 Kogelnik's Coupled Wave Theory assumes that two waves only propagate within the grating: The reference or R wave and the diffracted signal wave or S wave. Both waves are assumed to have the mathematical form of plane waves but with amplitudes which vary exclusively in the y dimension. The amplitude of both waves in the x direction is assumed to be constant. The example here is of a reflection grating.](image)

By substituting the assumed form of the index distribution of the grating (or equivalently, $\gamma$) and the assumed form for the $R$ and $S$ electric fields into equation (2), Kogelnik derived a relatively complicated equation relating the electric fields within the grating. This equation cannot itself be solved however and needs further simplification.
The term $\gamma^2 u$ in equation (2) couples the oscillatory form of the grating (described by $\gamma$) with the oscillatory form of the two propagating waves. The first assumption is therefore to simply ignore any coupling which produces higher modes of oscillation. But this still leaves an equation containing several types of oscillating modes. Kogelnik therefore argues that each type of mode (there are two of them) must satisfy its own equation independently. This process then leads to two second order differential equations for $R$ and $S$. Finally Kogelnik argues that since the amplitude of $R$ and $S$ must be expected to vary slowly with grating depth their second order rate of change must be small and so should be ignored. This then leads to two first order differential equations which at Bragg resonance define $R$ and $S$ within the grating, given the details of the illuminating wave.

Away from Bragg resonance, Kogelnik needs to make one further somewhat special assumption concerning the propagation wavevectors of the $R$ and $S$ waves within the grating. Specifically he assumes that the $R$ wave propagates "normally" - that is that it propagates in the same direction and with the same free-space wavelength as it does outside the grating (note that the average index inside and outside the grating are the same so Snell's law doesn't bend the ray). However he is obliged to make the assumption that the $S$ wave propagates in a rather special fashion in that in general its free-space wavelength is not equal to the free-space wavelength of the driving $R$ wave and also its direction is different from the ray direction at Bragg resonance.

One can perhaps appreciate that Kogelnik's model, despite its great practical success in providing useful formulae for evaluating the diffractive efficiency of gratings, doesn't therefore constitute a completely transparent model of what is actually occurring inside a volume grating during the process of diffractive replay. It was with this thought in mind that I became motivated to develop an alternative coupled wave description of diffraction which is based on the concept of Fresnel reflection.

3. A Coupled-Wave Theory based on Fresnel Reflection - The "PSM" Model

When a plane light wave propagates across the boundary of two regions of constant but differing refractive index, a portion of the light wave is reflected and a portion is transmitted. In the case of a purely real index there is no absorption and the transmitted and reflected energies add up to the incident energy. The laws describing how the amplitude of the transmitted and reflected waves depends on the indices of the two regions are known as Fresnel's laws. In the case of normal incidence they are trivially simple. If the wave passes from index $n_1$ to index $n_2$, then the ratio of the reflected to incident amplitude is simply given by

$$ r = \frac{n_2 - n_1}{n_2 + n_1} \tag{3} $$

Likewise the ratio of the transmitted amplitude to the incident amplitude is given by

$$ t = \frac{2 n_1}{n_2 + n_1} \tag{4} $$

The PSM theory models a volume grating as many thin parallel stacked layers, each with a slightly different index. Overall the grating looks just like a normal grating except that under the microscope the index does not vary smoothly from one depth to another but rather it makes small jumps. However, by making the distance between such jumps vanishingly small, we won't know the difference!

When an illuminating wave strikes a jump between one index and the next, we can use equations (3) and (4) to write down the amount of light transmitted and reflected. Of course these equations are only valid for the case of normal incidence but we can generalise them easily enough to incidence at any angle. Since the index layers are assumed flat and parallel in the simplest PSM model
we can also use the well-known law that the angle of incidence is equal to the angle of reflection. If we then call our (monochromatic) illuminating wave, \( R \), it is immediately obvious that at each index jump, \( R \) is slightly depleted and a new wave, which we can call \( S \), is created by Fresnel reflection.

3.1. The Simplest Possible PSM Theory - The Normal-Incidence Volume Reflection Phase Grating

We will now present the simplest possible mathematical derivation of the simplest example of a PSM coupled-wave theory. This is the case of the sinusoidal un-slanted volume reflection phase grating under illumination at normal incidence. Fig.2 shows a diagram of the set-up. We start by modelling the grating as a stack of parallel slices, each having a slightly differing refractive index. At this stage we don't make any assumption about the form of the index distribution other than it being composed of slices! Note that each slice will have a small but finite thickness and in the other two dimensions it will have the form of an infinite rectangular plane.

![Fig.2](image)

**Fig.2** The simplest PSM model: The un-slanted volume reflection phase grating illuminated at normal incidence.

We now assume that a monochromatic plane wave illuminates the grating. Mathematically we state this by the equation

\[
R_R^{in} = e^{i\beta y}
\]

where

\[
\beta = \frac{2\pi n_0}{\lambda_c}
\]

Here \( n_0 \) is the average value of the index inside and outside the grating and the parameter \( \lambda_c \) represents the free-space wavelength of the wave. If we now label each one of our "slices" by the integer, \( J \) and let the index of the \( J \)th slice be \( n_J \), then we can use the Fresnel relations to write down equations for how the \( S \) wave is generated by the \( R \) wave and how the \( S \) wave gives rise in turn to a reverse contribution to the \( R \) wave. Mathematically this can be written down as

\[
R_J = 2e^{i\beta n_J y/n_0} \left\{ \frac{n_{J-1}}{n_J + n_{J-1}} R_{J-1} + e^{i\beta n_J y/n_0} \left\{ \frac{n_{J-1} - n_J}{n_J + n_{J-1}} \right\} S_J \right\} \]

\[
S_J = 2e^{i\beta n_J y/n_0} \left\{ \frac{n_{J+1}}{n_J + n_{J+1}} S_{J+1} + e^{i\beta n_J y/n_0} \left\{ \frac{n_{J+1} - n_J}{n_J + n_{J+1}} \right\} R_J \right\}
\]

This is actually very simple! The first equation just describes how the \( R \) wave at the \( J \)th position is generated by the transmission of the \( R \) wave at the \( J-1 \) position and by the reflection of the \( S \) wave at
the Jth position. Likewise the second equation defines how the S wave at the Jth position is formed by the transmission of the S wave at the J+1 position and by the reflection of the R wave at the Jth position. You will recognise no doubt the terms in curly brackets as just being the Fresnel amplitude transmission and reflection expressions. The exponential terms take care of the phase of the waves. As each wave propagates so its phase changes and this must be accounted for properly.

The next stage is to use the laws of calculus and to let the thickness of the slices, δy tend to zero. This allows us to convert these "discrete difference" equations into the form of differential equations:

\[
\frac{dR}{dy} = \frac{R}{2} (2i\beta - \frac{1}{n_0} - \frac{1}{n} \frac{dn}{dy}) - \frac{1}{n} \frac{dn}{dy} S
\]

\[
\frac{dS}{dy} = -\frac{S}{2} (\frac{1}{n} \frac{dn}{dy} + 2i\beta - \frac{1}{n_0} \frac{dn}{dy}) - \frac{1}{n} \frac{dn}{dy} R
\]

Here \( n_J \) has been replaced by the continuous index variable, \( n \). These are the PSM coupled wave equations for a normal incidence un-slanted reflection volume phase grating. Unlike Kogelnik's coupled wave equations they are actually exact. Or in other words they constitute a different way of writing down the Helmholtz equation!

In order to solve the PSM coupled-wave equations we need to define the grating in terms of an index profile. To investigate the behaviour of the simple sinusoidal grating we use the following definition

\[
n = n_0 + n_1 \cos(\frac{4\pi n_0}{\lambda_r} y) = n_0 + \frac{n_1}{2} (e^{i\frac{\delta n}{\lambda_r}} + e^{-i\frac{\delta n}{\lambda_r}})
\]

This just says that the index profile of the grating is made up of two parts - a constant part, \( n_0 \) and a small sinusoidally varying part, \( n_1 \). We then make a mathematical transformation

\[
R \rightarrow R'(y) e^{i\beta y} ; \quad S \rightarrow S'(y) e^{-i\beta y}
\]

Due to the symmetry of the problem at hand and the assumed form of the illuminating wave, (5) this is only a mathematical tool - it is not an assumption. In fact the only assumption that we need to make is that the primed terms are slowly varying compared to the exponential. And this will nearly always be true. Under this one assumption, we can write

\[
\langle R' \rangle = R ; \quad \langle S' \rangle = S
\]

where the operator \( \langle \cdot \rangle \) takes an average over several cycles of the wave. Then simple substitution of (9) and (10) in (8) yields the following much simplified coupled wave equations:

\[
\frac{dR}{dy} = -i\alpha S e^{2i\beta y (a-1)} ; \quad \frac{dS}{dy} = i\alpha R e^{-2i\beta y (a-1)}
\]

where

\[
\alpha = \frac{\lambda_r}{\lambda}
\]

is the ratio of the playback to recording wavelengths. If we choose appropriate boundary conditions for the reflection grating (\( R(0) = 1; S(0) = 0 \)) these equations actually have analytic solutions which are characterised by a diffractive response which is almost exactly the same as Kogelnik's theory. We can use a mathematical trick which makes this rather more obvious. This is the introduction of the pseudo-field in place of the real electric field, \( S \). In fact, this is what Kogelnik does implicitly. You will remember that he assumed that the \( S \) field propagated in a very special manner...
The pseudo-field is defined as
\[ \hat{S} = S e^{2i\beta y(\alpha^{-1})} \] (14)

If we substitute this expression into (12) then we arrive at Kogelnik's coupled wave equations:
\[ c_R \frac{dR}{dy} = -i\kappa \hat{S} ; \quad c_S \frac{d\hat{S}}{dy} = -i\vartheta \hat{S} - i\kappa R \] (15)

However the constants \( c_R, c_S \) and \( \vartheta \) are slightly different from Kogelnik's values. The PSM theory gives
\[ c_R = \frac{1}{\alpha} ; \quad c_S = -\frac{1}{\alpha} ; \quad \vartheta = 2\frac{\beta}{\alpha} (1 - \alpha) \] (16)

whereas Kogelnik's theory gives
\[ c_R = 1 ; \quad c_S = (2\alpha - 1) ; \quad \vartheta = 2\alpha \beta (1 - \alpha) \] (17)

As usual Kogenik's constant is defined as
\[ \kappa = \frac{\pi n_1}{\lambda_e} \] (18)

One immediate advantage of the PSM theory is that it can be extended trivially to cover multi-colour gratings. We simply replace (9) with a sum of different recorded index profiles, one for each colour. Mathematically we write this as
\[ n = n_0 + n_x \cos(2\alpha \beta y) + n_y \cos(2\alpha \beta y) + ... \]

\[ = n_0 + n_x \frac{1}{2} \left( e^{2i\beta x y} + e^{-2i\beta x y} \right) + n_y \frac{1}{2} \left( e^{2i\beta y z} + e^{-2i\beta y z} \right) + ... \] (19)

Then (12) just becomes
\[ \frac{dR}{dy} = -\sum_{j=1}^{N} i\kappa_j \alpha_j e^{2i\beta y(\alpha_j^{-1})} ; \quad \frac{d\hat{S}}{dy} = R \sum_{j=1}^{N} i\kappa_j \alpha_j e^{-2i\beta y(\alpha_j^{-1})} \] (20)

where the integer index, \( j \) sums the colours. As long as the colours are not too close together these equations give the following useful expression for the diffraction efficiency of the polychromatic un-slanted volume phase grating
\[ \eta = \sum_{j=1}^{N} \frac{\alpha_j^2 \kappa_j^2}{\beta^2 (1 - \alpha_j)^2 + \left( \alpha_j^2 \kappa_j^2 - \beta^2 (1 - \alpha_j)^2 \right) \coth^2 \left( \frac{d \sqrt{\alpha_j^2 \kappa_j^2 - \beta^2 (1 - \alpha_j)^2}}{2} \right)} \] (21)

3.2 Angled Incidence, Slanted Gratings and Holograms

The PSM theory can easily be generalised to the case of a tilted illumination beam [16,18]. The equations can then be solved in a tilted frame of reference to give analytic expressions for the diffraction efficiency of the polychromatic slanted reflection volume phase grating (for both \( \sigma \) and \( \pi \)-polarisations). The final results are very similar to Kogelnik's but the PSM model can usually provide a somewhat superior estimation of the off-Bragg behaviour - see for instance Fig.3. By superposing many differently tilted gratings the PSM equations can be extended to an analytic description of the spatially multiplexed grating and by employing an infinity of such gratings, to the case of the volume hologram [17,18].

The extended PSM theory directly shows that the lossless polychromatic reflection volume phase hologram is theoretically capable of perfect diffractive replay at each and every wavelength. In other words, if such a hologram is co-illuminated by \( P \) co-propagating reference plane waves, each of a different wavelength, then the hologram is capable of producing a perfect diffractive response to each of these \( P \) waves simultaneously. The all-important parameter here is the effective index
modulation achievable in a given material at a given wavelength. If this is low then a thicker grating will be needed to achieve a bright hologram. The PSM model can be further extended to model lossy holograms by consideration of a complex index. In this case the strategy of simply making a reflection hologram thicker to produce a brighter diffractive response does not work as the further the reference wave penetrates into the hologram the more it is absorbed. This leads to the lossy polychromatic reflection hologram being characterised by a maximum attainable diffractive response which is dependent on the maximum achievable modulation in a given material. As such the PSM theory shows clearly that both the achievable index modulation and the intrinsic loss of a given photosensitive material must be expected to constitute the key parameters in determining the maximum practical diffractive efficiency of full-colour reflection holograms.

4. References

[1] H.Kogelnik, “Coupled Wave Theory for Thick Hologram Gratings”, Bell Syst. Tech. J. 48, 2909-47 (1969).
[2] E.J Saccocio, "Application of the dynamical theory of X-ray diffraction to holography", J.Appl. Phys. 38, 3994-8 (1967).
[3] V.V Aristov and V.Sh. Sheekhtman, "Properties of three-dimensional holograms", Sov. Phys. Uspechi 14, 263-77 (1971).
[4] C.J.R Sheppard, "The application of the dynamical theory of X-ray diffraction to thick hologram gratings", Int. J.Electronics 41, 365-73 (1976).
[5] R.S. Chu and J.A.Kong, "Modal Theory of spatially periodic media", Trans IEEE MTT-25, 18-24 (1977).
[6] L.Solymar,"Two-Dimensional N-Coupled-Wave Theory for Volume Holograms”, Opt.Comm. 23, 199-202 (1977).
[7] D.J.Cooke and L.Solymar, “Comparison of Two-Wave Geometrical Optics and N-Wave Theories for Volume Phase Holograms”, J.Opt.Soc.Am.70, 1631 (1980).
[8] M.G.Moharam and T.K Gaylord, "Rigorous Coupled Wave Analysis of Planar Grating Diffraction", JOSA, Vol. 71, Issue 7, pp. 811-818 (1981).
[9] M. P. Rouard, “Etudes des propriétés optiques des lames métalliques très minces,” Ann. Phys. (Paris) Ser. II 7, 291–384 (1937).
[10] F. Abeles, “Recherches sur la propagation des ondes électromagnétiques sinusoïdales dans les milieux stratifiés. Application aux couches minces,” Ann. Phys. (Paris) 5, 596–640(1950).
[11] R. Jacobson, “Light reflection from films of continuously varying refractive index”, Progr in Optics, E. Wolf, ed. North-Holland, Amsterdam, Vol. 5, Chap. 5, pp. 247–286.(1966).
[12] M. G. Moharam and T. K. Gaylord, “Chain-matrix analysis of arbitrary-thickness dielectric reflection gratings,” J. Opt. Soc.Am. 72, 187–190 (1982).
[13] X. Ning, “Analysis of multiplexed-reflection holographic gratings,” J. Opt. Soc. Am. A 7, 1436–1440 (1990).
[14] D.W. Diehl and N.George, "Analysis of Multitone Holographic Interference Filters by Use of a Sparse Hill Matrix Method," Appl. Opt. 43, 88-96 (2004).
[15] J.E.Ludman, “Approximate bandwidth and diffraction efficiency in thick holograms" Am. J. Phys., 50, 244, (1982).
[16] D.Brotherton-Ratcliffe, "A Treatment of the General Volume Holographic Grating as an Array of Parallel Stacked Mirrors", Submitted 24th December 2011 and Accepted for publication 14th May 2012 to J.Modern Optics.
[17] D.Brotherton-Ratcliffe, "An Analytical Treatment of the Polychromatic Spatially Multiplexed Volume Holographic Grating", Submitted 18th May 2012 to Applied Optics.
[18] H.Bjelkhagen and D.Brotherton-Ratcliffe, "Ultra-Realistic Imaging - Advanced Techniques in Analogue and Digital Colour Holography", CRC Press, Taylor and Francis Group, ISBN: 978-1-4398-2799-4, December (2012).