Theory of Fiber Optical Bragg Grating - Revisited

H. Tai

Langley Research Center, NASA, Hampton VA 23681

ABSTRACT

The reflected signature of an optical fiber Bragg grating is analyzed using the transfer function method. This approach is capable to cast all relevant quantities into proper places and provides a better physical understanding. The relationship between reflected signal, number of periods, index of refraction, and reflected wave phase is elucidated. The condition for which the maximum reflectivity is achieved is fully examined. We also have derived an expression to predict the reflectivity minima accurately when the reflected wave is detuned. Furthermore, using the segmented potential approach, this model can handle arbitrary index of refraction profiles and compare the strength of optical reflectivity of different profiles. The condition of a non-uniform grating is also addressed.

Keywords: Bragg grating, reflectivity, period, segmented potential, non-uniform grating

1. INTRODUCTION

Fiber Bragg gratings represent an important element in the emerging fields of optical communications and optical sensing. Despite its vast usefulness, the device is comparatively simple. Typically a dielectric cylinder of index \( n_1 \), usually referred as core, is surrounded by a concentric dielectric cylinder of index \( n_2 \). The two refractive indices obey the relation \( n_1 > n_2 \). In such an arrangement, since the field decays exponentially inside region of index \( n_2 \), practically no field exists outside of region 2. In its simplest form a fiber Bragg grating consists of a periodic modulation of the refractive index in a core of a single mode optical fiber, where the phase fronts are perpendicular to the fiber’s longitudinal axis and with grating planes having a constant period. Because of its intrinsic simple physical nature, the theory behind Bragg grating is equally simple. Light, guided along the core of an optical fiber, is scattered by each grating plane. If the Bragg condition is satisfied, the contributions of reflected light from each grating plane add constructively in the backward direction to form a back reflected peak with center wavelength defined by the grating period. A considerable amount of theoretical work \(^1\text{-}^{11}\) has been reported with various approaches giving reasonable results in predicting the reflectivity as a function of wavelength. The theoretical works fall into two categories, the matrix transfer function method and the coupled mode theory. In the matrix transfer function, the Bragg grating is simulated as an alternating stratified medium having index of refraction of \( n \) and \( n+\Delta n \) and weak guiding approximation is assumed, i.e., the difference of index refraction between the core and cladding is ignored, which some times is called the scalar wave approximation. In this approximation, the modes are transverse. When the grating is periodic, a closed solution for the reflectivity as a function of scanning wavelength is obtained. However when the grating periodicity is slightly off as happens in a non-uniform strain measurement, the reflected spectrum can only be obtained numerically\(^ {12} \). The coupled mode method assumes the grating serves as a perturbation which couples power between forward and backward moving modes with a \( x \) (optical axis) dependence amplitude. This method also assumes that the amplitudes do not change abruptly and the second derivative of amplitudes with respect to \( x \) are dropped resulting in the amplitudes being represented by a first order linear differential equation with constant coefficients. Both approaches provide reasonable results.

2. EVALUATION OF REFLECTIVITY USING TRANSFER MATRIX METHOD

Following references\(^ {13}\text{-}^{14}\) closely, we outline the simple results of the expressions for light propagating in a Bragg grating. Under the assumption of a scalar wave approximation (weakly guiding approximation) the electric and magnetic field amplitudes are assumed to be perpendicular to the propagation direction and both satisfy the simple plane wave equation. When the plane wave travels through regions of different index of refraction, the wave number adopts the local index and part of the wave gets reflected back with different amplitude. A typical square index profile for a Bragg grating with period \( s \) is shown in Fig.1.
Consider a plane wave polarized in the y direction from a medium of index \( n \) and is incident normally to a film of thickness of \( 2a \), of index \( n' \) and emerging to medium of \( n \). We may interpret this arrangement as equivalent to an optical index potential.

\[
E = \begin{cases} 
\hat{y}E_1^+ e^{+ikx} + \hat{y}E_1^- e^{-ikx}, & x < -a \\
\hat{y}E_2^+ e^{+ik'x} + \hat{y}E_2^- e^{-ik'x}, & -a < x < a \\
\hat{y}E_3^+ e^{+ikx} + \hat{y}E_3^- e^{-ikx}, & x > a
\end{cases}
\]  

(1)

where \( \hat{y} \) stands for a unit vector in the y direction and the quantity \( E_1^+ e^{ikx} \) represents a plane wave propagating in the positive x direction with an amplitude \( E_1^+ \); the quantity \( E_1^- e^{-ikx} \) represents a plane wave propagating in the negative x direction with an amplitude \( E_1^- \) etc.

\[ k = (\omega/c)n = k_0n \quad ; \quad k' = (\omega/c)n' = k_0n' \]  

(2)

and \( k_0 = \omega/c = 2\pi/\lambda \) represents the free space wave number. The corresponding magnetic fields \( H \) can be evaluated as

\[
H = k \times E / \omega \mu
\]  

(3)

\[
H = \begin{cases} 
\hat{z} \frac{n}{c\mu_0} E_1^+ e^{ikx} - \hat{z} \frac{n}{c\mu_0} E_1^- e^{-ikx}, & x < -a \\
\hat{z} \frac{n'}{c\mu_0} E_2^+ e^{ik'x} - \hat{z} \frac{n'}{c\mu_0} E_2^- e^{-ik'x}, & -a < x < a \\
\hat{z} \frac{n}{c\mu_0} E_3^+ e^{ikx} - \hat{z} \frac{n}{c\mu_0} E_3^- e^{-ikx}, & x > a
\end{cases}
\]  

(4)
where \( \omega \) is the frequency and \( \mu_0 \) is the magnetic permeability of the free space. Since both \( E \) and \( H \) represent tangential components, they must be continuous at the boundaries \(-a\) and \( a\). By eliminating \( E^+_2\) and \( E^-_2\), \( E^+_1 \) and \( E^-_1 \) can be expressed in terms of \( E^+_3 \) and \( E^-_3 \) as follows

\[
\begin{pmatrix} E^+_1 \\ E^-_1 \\ \end{pmatrix} = \mathbf{m} \begin{pmatrix} E^+_3 \\ E^-_3 \\ \end{pmatrix}
\]

(5)

\[
\mathbf{m} = \begin{pmatrix} (\cos 2k'a - i\varepsilon, \sin 2k'a)e^{2ika} & i\varepsilon_+ \sin 2k'a \\ -i\varepsilon_+ \sin 2k'a & (\cos 2k'a + i\varepsilon, \sin 2k'a)e^{-2ika} \end{pmatrix} \equiv \begin{pmatrix} w & z \\ z^* & w^* \end{pmatrix}
\]

(6)

\[
\varepsilon_\pm = \frac{1}{2} (\eta \pm \frac{1}{\eta}), \quad \eta = \frac{n}{n'}
\]

(7)

where \( \mathbf{m} \) is referred to as the transfer matrix for single potential. We may construct the transfer matrix for the whole array by duplicating the basic potential in Eq.(6) \( N \) times at regular interval \( s \) with \( s > 2a \).

The electric field between the potentials (where the index is \( n \)) can be written

\[
E_m(x) = A_m e^{ik(x - ms)} + B_m e^{-ik(x - ms)} \quad \text{where} \quad (m-1)s + a < x < ms - a,
\]

\[0 < m < N\]

(8)

To keep track of the subscript index, we have changed the notation from \( E^+_i \) and \( E^-_i \) to \( A_m \) and \( B_m \). Again, \( A_m \) and \( B_m \) represent the electric field amplitude of the wave moving to the right and left, respectively.

For the \( m \)th potential we can write as in Eq. (6) and obtain

\[
\begin{pmatrix} A_m \\ B_m \end{pmatrix} = \begin{pmatrix} (\cos 2k'a - i\varepsilon, \sin 2k'a)e^{2ika} & i\varepsilon_+ \sin 2k'a \\ -i\varepsilon_+ \sin 2k'a & (\cos 2k'a + i\varepsilon, \sin 2k'a)e^{-2ika} \end{pmatrix} \begin{pmatrix} A_{m+1}e^{-iks} \\ B_{m+1}e^{iks} \end{pmatrix} = \mathbf{P} \begin{pmatrix} A_{m+1} \\ B_{m+1} \end{pmatrix}
\]

where \( \mathbf{P} = \mathbf{m} \begin{pmatrix} e^{-iks} & 0 \\ 0 & e^{iks} \end{pmatrix} = \begin{pmatrix} (\cos 2k'a - i\varepsilon, \sin 2k'a)e^{2ika}e^{-iks} & i\varepsilon_+ \sin 2k'a e^{-iks} \\ -i\varepsilon_+ \sin 2k'a e^{-iks} & (\cos 2k'a + i\varepsilon, \sin 2k'a)e^{-2ika}e^{iks} \end{pmatrix} \)

(9)

\( \mathbf{P} \) as well as \( \mathbf{m} \) is uni-modular, i.e., \( \det \mathbf{P} = 1 \). Using Eq. (9) recursively,

\[
\begin{pmatrix} A_0 \\ B_0 \end{pmatrix} = \mathbf{P}^N \begin{pmatrix} A_N \\ B_N \end{pmatrix}
\]

(10)

Now the task reduces to the evaluation of \( \mathbf{P}^N \). Among many methods reported, the most elegant is to use the Cayley-Hamilton theorem\(^\text{15}\) to obtain the expression of \( \mathbf{P}^N \).

\[
\mathbf{P}^N = \begin{pmatrix} (\cos k'r - i\varepsilon, \sin k'r)e^{iks}U_{N+1}(\xi) - U_{N-2}(\xi) & i\varepsilon_+ \sin k'rU_{N+1}(\xi)e^{iks} \\ -i\varepsilon_+ \sin k'rU_{N-1}(\xi)e^{-iks} & (\cos k'r + i\varepsilon, \sin k'r)e^{-iks}e^{iks}U_{N-1}(\xi) - U_{N-3}(\xi) \end{pmatrix}
\]
where
\[ \xi = \frac{1}{2} Tr(\mathbf{P}) = \frac{1}{2} \left( (\cos k'r - i\varepsilon_\gamma k'r)e^{ikl} + (\cos k'r + i\varepsilon_\gamma k'r)e^{-ikl} \right) \] (11)

We have changed the notation from $2a$ to $r$, i.e., the width of the potential (layer). $l$ is the width between the potentials and $s=r+l$ is the period as shown in Fig.1. $U_N$ is the $N$th Chebychev polynomial of the second kind.\(^{16}\)

2a Maximum Reflectivity (tuned)

From Eq.(10), we can easily determine the transmission coefficient (transmittivity) and reflectivity for $N$ potentials. For example, the transmission coefficient is

\[ T_N = \frac{1}{1+\left| i\varepsilon_\gamma \sin k'r \right| U_{N-1}(\xi)} \] (12)

Assuming there is no absorption, the reflectivity $R_N$ is

\[ R_N = 1 - T_N. \] (13)

Where
\[ \xi = \cos(k'r) \cos(\ell) - \varepsilon_\gamma \sin(k'r) \sin(\ell) \] (14)

and
\[ U_N(\xi) = \frac{\sin(N+1)\gamma}{\sin \gamma} \] (15)

\[ \gamma \equiv \cos^{-1} \xi \]

$T_N$ can also be expressed as

\[ T_N = \left| 1 + i\varepsilon_\gamma \sin k'r \right|^2 \left[ \frac{\sin N\gamma}{\sin \gamma} \right]^2 \] (16)

In practice, the difference $\Delta n$ is minute which implies that $\varepsilon_\gamma \approx 1$. We can further write Eq. (14) as

\[ \xi = \cos(k'r + kl) \] (17)

The numerical result shows this assumption is good, especially when tuned, i.e., $\xi = \cos(k'r + kl) = -1$. The quantity of interest to us is $R_N$ and in fact we would like to have $R_N$ approach $1$. In order that the value of $T_N$ approaches zero, i.e., the value of the denominator of $T_N$ has to approach infinity which implies that in Eq.(12) the value of $U_N(\xi)$ has to be large. This means $\gamma$ has to be zero or a multiple of $\pi$ and $\xi$ takes the value of $-1$ or $+1$ at which $U_N(1) = N+1$, $U_N(-1) = (-1)^N(N+1)$. The quantity inside the absolute value is finite and it takes the maximum value when $k'r$ becomes an odd multiple of $\pi/2$. Therefore, we can conclude that for the reflectivity to reach maximum value, in other words “tuned”, the optical path length in the $n'$ region (index of refraction higher) has to be close to odd multiple of $\pi/2$ and so is the optical path length in the $n$ region (index unchanged), but the sum of the optical path length in the whole period has to be a multiple of $\pi$. The number of periods $N$ happens to be a multiplicative factor. When $\gamma$ approaches a multiple of $\pi$, we find $T_N$ approaches zero as

\[ T_N \approx \frac{CN^{-2}}{1} \] (18)
\[ R_N = 1 - C N^2 \]

where \( C \) is just some proportional constant.

To reach maximum reflectivity,
\[ \begin{align*}
  k'r &= n'r2\pi / \lambda_0 = \pi / 2 \rightarrow r = \frac{\lambda_0}{4n'} \\
  kl &= nl2\pi / \lambda_0 = \pi / 2 \rightarrow l = \frac{\lambda_0}{4n}
\end{align*} \]

The period \( s = r + l = \frac{\lambda_0}{4} \left( \frac{1}{n'} + \frac{1}{n} \right) = \frac{\lambda_0}{2n_{\text{eff}}} \) \( \quad (19) \)

where \( \lambda_0 \) is the "tuned" wavelength and \( n_{\text{eff}} = \frac{2n'n}{n+n'} \)
as commonly used in the literature.

2b. Detuned reflectivity

When the scanning laser wavelength range is out of the “tuned” wavelength which is determined by the period \( s \), the reflectivity graph commonly shows damped oscillatory behavior with a series of outlying minima as a function of wavelength (shown in Fig. 2). The reflectivity graph is symmetrically displayed along both sides of the tuned wavelength \( \lambda_0 \).

In this method some explanation can be easily provided.

For \( R_N \) to be approaching zero, \( T_N \) has to be approaching 1. Since \( r \) and \( l \) have been fixed, we can rewrite Eq. (19) as
\[ \begin{align*}
  \tilde{k'r} &= \frac{\pi}{2} \left( \frac{\lambda_0}{\lambda} \right) = \frac{\pi}{2} \rho \\
  \tilde{kl} &= \frac{\pi}{2} \left( \frac{\lambda_0}{\lambda} \right) = \frac{\pi}{2} \rho
\end{align*} \]

First consider \( \lambda \) is greater than \( \lambda_0 \), i.e., \( \rho \) is less than 1 but greater than 0.

In the denominator of Eq. (12) \( \varepsilon \) is small in the order of \( \Delta n \), however is finite, and the other two terms contain all the relevant information. As \( \lambda \) departs from \( \lambda_0 \), i.e., \( \rho \) departs from 1 and approaches to 0; the term \( \sin k'r \) monotonically changes from 1 to 0 which gives the damping effect. And \( \xi \) changes from -1 to +1. For the domain -1 to +1, the Chebychev polynomial \( U_{N-1}(\xi) \) has \( N-1 \) zeros which are located according to this formulas \( \quad (21) \)

\[ \xi_{m-1}^{-N} = \cos \frac{m}{N} \pi, \quad m=1\cdots N-1 \]

We rewrite Eq. (14) as
\[ \xi = \cos(k'r)\cos(kl) - \varepsilon \sin(k'r)\sin(kl) \]
\[ = \cos(k'r)\cos(kl) - (1+\varepsilon-1)\sin(k'r)\sin(kl) \]
\[ = \cos(k'r+kl) - \frac{\Delta n^2}{2n(n+\Delta n)} \sin(k'r)\sin(kl) \] \( \quad (22) \)

With Eq. (21) and simple algebra, we obtain
\[ \lambda_m = \frac{\pi}{2} \frac{\lambda_0}{\cos^{-1} \left( \frac{B + \cos \left( \frac{m\pi}{N} \right)}{1+B} \right)} \] \( \quad (23) \)

where \( B \equiv 1+\frac{\Delta n^2}{2n(n+\Delta n)} \)
What will happen if \( \lambda \) departs from \( \lambda_0 \) and moving toward the left, i.e., toward shorter wavelength. As \( \rho \) increases from 1 to 2, the reflectivity is the mirror reflection image around the axis \( \rho = 1 \). Therefore, in principle, the reflectivity curve as a function of reduced parameter \( \rho \) has a period of \( 2\rho \); that is the reflectivity curve repeats itself every \( 2\rho \).

Fig. 2  a typical reflectivity vs wave length curve, \( \lambda_0 = 1550 \) nm, N=3000

3. IMPERFECT SQUARE INDEX PROFILE
If the profile is not perfectly square but is still periodic, we can rely on a numerical procedure to arrive at the value of \( m \) and subsequently \( P_N \). As shown in Fig. 3, an arbitrary profile can be segmented into \( M \) equally spaced square profiles of width \( w \) and continuity conditions applied to obtain \( m \).

Fig. 3 Imperfect index profile

The reflectivity of a perfect square profile and an imperfect profile were calculated using the following parameters: \( n = 1.45, \Delta n = 10^{-3}, n' = n + \Delta n, \lambda_0 = 1550 \) nm, grating period \( s = \lambda_0 / 2n_{\text{eff}} \), and the scanning wavelength ranging from 1546 to 1552 nm. A typical graph is shown as in Fig. 2. The peak of the reflected spectrum is located at 1550 nm as expected and the reflectivity registered at 0.93808 for a perfect square profile with 3000 periods. As shown in Fig. 3, the \( M \) value is chosen to be 8, i.e., the profile is segmented to be 8 equal parts of rectangular profile and the height of each rectangle, say \( n \), is greater than \( n \) but less than \( n' \). This represents the profile with a round off shoulder. Table 1 lists the segmented profiles and the index of refraction \( n \) at each station is uniformly distributed. We termed imperfect profile #1 and #2 and
in particular, the perfect profile can be represented in this format. Table 2 gives a comparison of these three profiles with total number of gratings of 3000, 2000, 1000 and 800. The values listed represent the reflectivity at the peak (at \(1550\) nm). For imperfect profile #2 the peak value of reflectivity has slightly shifted to \(\lambda_0=1549.8\) nm.

### Table 1 segmented index representation for profile

<table>
<thead>
<tr>
<th></th>
<th>Perfect profile</th>
<th>Imperfect profile #1</th>
<th>Imperfect profile #2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.451</td>
<td>1.4507098</td>
<td>1.45001332</td>
<td></td>
</tr>
<tr>
<td>1.451</td>
<td>1.4508549</td>
<td>1.45005685</td>
<td></td>
</tr>
<tr>
<td>1.451</td>
<td>1.451</td>
<td>1.4507098</td>
<td></td>
</tr>
<tr>
<td>1.451</td>
<td>1.451</td>
<td>1.451</td>
<td></td>
</tr>
<tr>
<td>1.451</td>
<td>1.451</td>
<td>1.451</td>
<td></td>
</tr>
<tr>
<td>1.451</td>
<td>1.451</td>
<td>1.4507098</td>
<td></td>
</tr>
<tr>
<td>1.451</td>
<td>1.4508549</td>
<td>1.45005685</td>
<td></td>
</tr>
<tr>
<td>1.451</td>
<td>1.4507098</td>
<td>1.45001332</td>
<td></td>
</tr>
</tbody>
</table>

### Table 2

Comparison of peak value of reflectivity at \(\lambda_0=1550\) nm for three profiles

<table>
<thead>
<tr>
<th>Total # of gratings</th>
<th>Imperfect profile #2</th>
<th>Percentage change from perfect profile</th>
<th>Imperfect profile #1</th>
<th>Percentage change from perfect profile</th>
<th>Perfect profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>800</td>
<td>0.151765</td>
<td>39.7</td>
<td>0.238144</td>
<td>5.68</td>
<td>0.251673</td>
</tr>
<tr>
<td>1000</td>
<td>0.223794</td>
<td>37.3</td>
<td>0.339586</td>
<td>5.16</td>
<td>0.357135</td>
</tr>
<tr>
<td>2000</td>
<td>0.596927</td>
<td>23.0</td>
<td>0.756083</td>
<td>2.50</td>
<td>0.775615</td>
</tr>
<tr>
<td>3000</td>
<td>0.831459</td>
<td>11.4</td>
<td>0.928576</td>
<td>1.00</td>
<td>0.938080</td>
</tr>
</tbody>
</table>

### 4. NON-UNIFORM GRATING

Non-uniform grating can exist either by design or arise from circumstance, such as in a structure non-uniform strain measurement. Since the “period” has lost its meaning, a compact expression for reflectivity can not be obtained. On the other hand, one can easily carry out the transfer function manipulation grating by grating by numerical procedure. As previously reported, a pre-described non-uniform strain can be suggested, including linear, sinusoidal and random distributions. Interesting results are obtained depending upon the regime of strain. The simulations indicate that for non-uniform strain, the multi-peak spectra occur when the strain has reached the order of \(10^{-2}\) or greater. If the strain is small, say less than \(10^{-3}\), non-uniformity is not an important issue; all the reflected spectra would give a sharp peak and uniquely determine the strain. However when the strain increases to the order of \(10^{-2}\), the spectrum is broadened and splits into multiple peaks. Finally, when the strain increases beyond \(10^{-2}\) for a non-uniform grating, the reflected signals can be completely lost, which has been observed in some experiments. These phenomena can be understood in a qualitative sense, i.e., each grating plane defines a reflected and selected wavelength. And when the grating distance is constant, all the reflected waves contribute constructively, creating a strong peak uniquely defining the grating distance. But when the successive grating distances are off slightly, each distance selects a slightly shifted wavelength. Therefore each back-scattered wave contributes non-coherently and multi-peak spectra are produced.

### 5. CONCLUSION

We have revisited the transfer matrix formulation once again, but also have gained some new understanding which was not well known before: (1) The transfer matrix is easy to implement and involves less approximations and in our opinion is more physically transparent than the coupled mode theory. (2) From the mathematical point of view, we understand...
how the maximum reflectivity is obtained; it is not the individual layer that produces the phase change of $\pi/2$, equivalent to say the individual layer width equal to $\lambda/4n$ or $\lambda/4n'$. The important fact is the sum of the phase changes in the individual double layer has to be $\pi$ or a multiple of $\pi$, and each layer can be slightly off from $\pi/2$. The reflectivity will decrease if the phase shift $k'r$ is moving away from odd multiple of $\pi/2$ because of the quantity $\sin(k'r)$ in Eq. (16). To gain maximum reflectivity, each layer must have a phase path length of odd multiples of $\pi/2$ which has some practical implication, namely we can inscribe less points on the fiber to achieve the same results. For example, in the $n$ region we can set $l=3\lambda_0/4n$. This has been demonstrated in numerical simulation. (4) We have gained the understanding of the role of number of double layers $N$. The reflectivity approaches 1 in the fashion shown in Eq. (18) for large $N$, which is in contrast to results shown in Ref.18 where the maximum reflectivity is proportional to $N$, when $N$ is small. (5) Eq.(23) shows how we can compute the full width half maximum of the major peak and also the distance between a pair of out lying minima. (6) The transfer function method can study and evaluate the relative efficiency of various index profiles by using the segmented potential treatment. (7) For non-uniform grating, a numerical procedure can be applied with interesting results.

6. REFERENCES