Heteroepitaxy on compliant substrates: relaxation of misfit stress at low critical film thickness

G. Kästner*
Max Planck Institute of Microstructure Physics, 06120 Halle (Saale), Germany

Received 26 February 2002, revised 30 September 2002, accepted 2 October 2002
Published online 24 January 2003

PACS 68.35.Gy, 68.55.Ac, 68.60.Bs

In spite of outstanding experimental success in reducing the density of threading dislocations (TDs) within thin films grown heteroepitaxially on using a compliant substrate instead of a conventional bulk substrate, theories of strain relief on compliant substrates are often based on inadequate theorems like the elastic strain-partitioning and the free-slipping hypothesis. Instead, the present paper modifies the well accepted dislocation model of Matthews by supposing that the misfit dislocations (MDs) relax or dissipate owing to the presence of a weak layer in compliant substrates. This relaxation reduces the drag force on the TDs and therefore the critical film thickness as estimated in a quantitative way. The consequences include less interaction and multiplication of the dislocations, reduced TD density, more homogeneous film structure, and lower final equilibrium strain.

1 Introduction

Heteroepitaxially grown thin films usually suffer from a high number (area density) of threading dislocations (TDs) adverse to the desired film properties, above all at semiconductor and optoelectronic applications. The TDs can form as growth defects and by the heteroepitaxial misfit strain. In this paper, we consider epitaxial films (layers) of uniform thickness where the misfit strain can activate TDs to slip on inclined slip planes in order to relax this strain. The basic mechanism known from Matthews [1, 2] is illustrated in Fig. 1a. A moving TD trails two line defects: a slip step along the film surface and a misfit dislocation (MD) along the interface. This MD displays the degree of local relaxation, determined by the

Fig. 1 Cross-sectional schematics of a heteroepitaxial film F grown either a) on a bulk substrate B or b) on a compliant substrate (consisting of a crystalline template layer T separated by a weak interfacial layer from a bulk substrate B). A threading dislocation TD moves towards right in its slip plane. Along its trace (indicated by broken or solid lines), the TD forms a slip step on the film surface and either a misfit dislocation MD or a strain-relaxed trace between T and B in the case (a) or (b), respectively.

*e-mail: kaestner@mpi-halle.de
film-parallel component of the Burgers vector. According to this mechanism, no relaxation can occur without moving a TD.

A lot of attempts were made to reduce the excessive TD density by means of controlling this relaxation mechanism, e.g. [3–5]. Particular progress has been achieved, however, on replacing the conventional bulk substrate by a “compliant substrate” as repeatedly reviewed [6–9]. According to Fig. 1b, this compliant substrate consists, on principle, of a thin single-crystalline template layer T which is “weakly” bonded on a bulk substrate B. The “weak” interface or interfacial layer between T and B is crucial. It has been formed in various ways, including an amorphous layer [10–13], a twist-bonded interface [8, 14], and in some respect also an ion-implanted layer [15] as well as porous material [16, 17]. In all such cases, considerable reduction of the TD density or improved strain relief was reported. In spite of this experimental success, many theories on strain relief by a means of a compliant substrate suffer from inadequate theorems and concepts. These theories seem to be fixed by the aim of increasing the critical film thickness for the onset of plastic relaxation. Below an opposite approach will be explained.

2 Basic theorems

The basic question is whether a compliant substrate allows the film to relax elastically or plastically. Lo [18] introduced the idea of elastic relief, supposing that the heteroepitaxial strain will be relaxed elastically by a lateral force balance between the film F and the template T. This hypothesis, known as strain partitioning, is based on the idea that the template as a whole can contract or expand by slip along the weak interface (“free-slipping hypothesis”). Although the latter hypothesis became quite common, it has been rejected by several authors [9, 13, 19–22] for three reasons: First, real “weak” interfaces have a considerable sliding resistance [6] which appears much too high [13, 21, 22]. Second, the whole area of the film would have to slip in order to allow for strain partitioning because the resulting elastic extension or contraction of the film as a whole necessarily changes its macroscopic lateral dimension [22]. This elementary fact is commonly overlooked when adhering to the schematic drawings far from real scale. To elastically relax a film of 100 mm diameter by 1%, it has to slip 0.5 mm at its rim. No experimental indication of this became known. Third, the initiating shear stress [21] acting along the weak interface is confined to the very rim of the film within few micrometers [22]. For these reasons, the free-slipping hypothesis might be valid for films up to mesoscopic lateral dimensions in the µm range [23, 24] but it appears basically misleading in the case of real wafer-sized film areas.

Instead, it is much more reasonable to consider plastic relief of the heteroepitaxial film in common terms of plastic slip [22]. That means the film has to relax by slip on its crystallographic slip planes where the slip direction is oblique to the film plane for reasons of the maximum Schmid factor. (The latter is zero for slip planes parallel or perpendicular to the film plane, e.g. in the case of GaN films [25]. In such a case, however, relaxation may occur by climbing of the TDs instead of slip.) In our case of plastic slip, we have necessarily to deal with threading dislocations moving on these oblique slip planes. Therefore we have no reason to reject the elementary Matthews model of Fig. 1a, but it has to be modified according to Fig. 1b as explained below.

3 Relaxation of misfit dislocations

In our view, it is the essential feature of a compliant substrate that the misfit dislocation formed behind a moving TD can disappear by local strain relaxation owing to the “weak layer”. Two steps are included: First, the initial dislocation will move through the template T towards the weak layer since the latter exerts an attracting image force [11, 21, 26]. This force is the stronger the thinner the template and the better the weak layer relaxes strain [20, 21]. A supporting force will arise if the template T is pseudomorphically strained as the film F, e.g. in the case of [12]. Second, on arriving at the weak interlayer, the dislocation can relax or dissipate in various ways, depending in the nature of this layer. The following possibilities will be distinguished:

(i) Viscous dissipation [21], e.g. by an interlayer of a low-melting metal [27] or low-melting glass [13, 20, 21]. The geometry of this relaxation can be closely compared to the case when a dislocation slips
off a free surface: The stress field of the dislocation disappears by forming stress-free displacements, consisting of two components: The in-plane displacement is responsible for locally relaxing the misfit while the orthogonal component forms an atomic slip step on the surface (or the interface in our case). If a viscous layer of suitable thickness is present, principally the same relaxation happens but with a time- and temperature-dependent delay. Quantitative calculations have shown that a viscous borophosphosilicate glass layer can dissipate an MD within seconds at 800 °C [21].

(ii) Plastic relaxation by means of diffusion: As known easily diffusing species generally tend to relax strain fields, thus forming a Cottrell atmosphere around a dislocation, in our case relaxing the MDs. For example, “weak” interfaces formed by wafer bonding may include diffusing species as well as imperfections, e.g. dislocations, which support diffusion.

(iii) Elastic relaxation due to a low elastic constant (strength) of the “weak” layer. This might be valid for a layer of low-temperature deposited silicon containing a high density of vacancy clusters [17].

(iv) Disappearance of the MDs along a certain length, possibly in two ways: First, by absorption into microcavities, e.g. those of a porous interlayer [16, 17] or those formed by ion implantation. In the latter case [15], “nanocavities” of surface-parallel platelet shape were observed below and within the epitaxial interface. An MD absorbed into cavities reduces its effective length and hence its average line energy. Second, disappearance of an MD by attractive reaction with other dislocations, e.g. those formed by twisted wafer bonding [8, 14] as discussed in [28, 29], or by reaction with dislocation loops formed by ion implantation [15]. Thus, an MD may disappear partially or even totally along a certain length.

Irrespective of these various mechanisms – which of course may act together in various cases – the resulting relaxed state instead of the former MD is symbolized in Fig. 1b by a broken line along the weak interlayer. As shown below this kind of relaxation affects the force and energy balance and hence the critical thickness in comparison with Fig. 1a.

4 Critical thickness

Figure 1a, representing the well accepted model of Matthews [1, 2, 30], illustrates the force balance acting on a moving TD. This TD experiences at least two static drag forces: one acting along the film surface due to forming the slip step and the other one acting along the interface due to forming the MD. The sum of these drag forces has to be exceeded by the Peach-Koehler driving force [3] imposed by the heteroepitaxial stress. Since it acts along the whole length of the TD, it increases with the film thickness \( h \). Therefore a critical thickness \( h_c \) is required to start the relaxing slip. (The same \( h_c \) results from the alternative approach of minimizing the total energy of the system as shown in the more elaborate theory [30].)

If, instead, a compliant substrate is used which is able to relax the line energy (energy per unit length) of an MD – whatever the details might be – the drag force exerted on a moving TD is reduced. This reduces the necessary driving force and, consequently, the critical film thickness \( h_c \). This conclusion, pointed out in a preceding paper [22], is in contrast to enhanced \( h_c \) derived from current theories of a compliant substrate which start from the misleading hypotheses of strain-partitioning and free slipping.

Since qualitative arguments only were given in [22], a quantitative calculation will be presented below. It is simplified as follows: No specification is made about the crystallographic slip system and the nature of the weak interface. The elastic constants, including the shear modulus \( \mu \) of F, T, and B will be taken as identical. As usual we consider the force equilibrium [1, 2, 30] on the line defects (TD, MD, relaxed MD). If \( h_t \) and \( e_t \) as well as \( h_i \) and \( e_i \) denote the thickness and the biaxial strain of the film and the template, respectively, these quantities determine the driving force exerted on a TD, in our approximation as

\[
F_{\text{TD}} = 2\mu b (e_1 h_1 + e_2 h_2) ,
\]

where \( b \) is the effective component of the Burgers vector.

To give an idea of the opposing drag forces, we suppose the relaxed MD to be a slip step as in the case of viscous relaxation. This step is supposed to create a new interface of width \( b \) and interfacial energy \( \sigma \).
between \( T \) and the weak layer. Hence, the step energy per unit length (line energy) is \( F_s = \sigma_s b \). Analogously to the case of a free surface \([1, 31]\), we take roughly \( \sigma_s = \mu b/10 \), hence

\[
F_s = \frac{b^3 \mu}{10}. \tag{2}
\]

The same line energy will be used to represent the drag of the usual slip step created on the film surface. Thus, the total drag force is \( 2F_s \), when neglecting the additional but frictional Peierls force. Hence, a static force equilibrium is taken as \( F_{TD} = 2F_s \), i.e. \( \varepsilon_f h_f + \varepsilon_t h_t = b/10 \). Solving for \( h_f = h_s \) yields the critical film thickness of this model

\[
h_s = \frac{b}{10} - \frac{\varepsilon_f h_f}{\varepsilon_t}. \tag{3}
\]

For comparison, the known formula of the critical film thickness \( h_c \) on a usual bulk substrate will be derived based on essentially the same approximations: The drag force (line tension) of an MD is written as \( F_{MD} = b^2 (\mu/10) \ln (2h/b) \) on taking \( b/2 \) as the core radius of the dislocation \([30, 32]\). From the force balance \( F_{TD} = F_{MD} + F_s \) where \( F_s \) is the drag of the surface step and where \( F_{TD} \) of (1) is used on supposing a strain-free template \( \varepsilon_t = 0 \), the critical thickness follows implicitly:

\[
h_c = \frac{b \ln (2h_c/b) + 1}{20 \varepsilon_t}. \tag{4}
\]

Using (4) and (3), the ratio of the critical film thickness on a bulk and on a compliant substrate with \( \varepsilon_t = 0 \) is

\[
h_f/h_s = \frac{\ln (2h_c/b) + 1}{2}, \tag{5}
\]

where the ratio of the drag forces \( F_{MD}/F_s = \ln (2h_c/b) \). Numerical examples are given in Table 1. Remarkably, \( h_c \) agrees with more elaborate data (Fig. 9 of [30]) within 10%.

In this example, the use of a compliant instead of a bulk substrate reduces the critical thickness by a factor of 2 to 3.5. Evidently this factor depends on the degree of the reduced drag force and thus on the nature of the weak interlayer. The better this layer can relax a MD, the lower the critical thickness of the compliant substrate is expected. For example, as we treated the interfacial slip step like a surface step, its “surface energy” \( \sigma_s \) and hence \( F_s \) of Eq. (2) may be in fact lower if the wetting by the viscous interfacial matter reduces this surface energy towards a lower interfacial energy. If – instead of viscous material – the “weak” interface contains dislocations able to attract and annihilate an approaching MD, this attractive force instead of a drag would support the driving force. Thus, the critical thickness may become extremely low. Apparently any given system requires a more specific treatment, eventually including the kinetic and dynamic behaviour of the dislocation elements involved.

<table>
<thead>
<tr>
<th>( \varepsilon_t ) (%)</th>
<th>( h_s ) (nm)</th>
<th>( h_f ) (nm)</th>
<th>( h_f/h_s )</th>
<th>( F_{MD}/F_s )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.17</td>
<td>23</td>
<td>81</td>
<td>3.5</td>
<td>6</td>
</tr>
<tr>
<td>0.41</td>
<td>9.9</td>
<td>30</td>
<td>3.0</td>
<td>5</td>
</tr>
<tr>
<td>0.92</td>
<td>4.4</td>
<td>11</td>
<td>2.5</td>
<td>4</td>
</tr>
<tr>
<td>2.0</td>
<td>2.0</td>
<td>4.0</td>
<td>2.0</td>
<td>3</td>
</tr>
</tbody>
</table>
5 Consequences of a reduced critical thickness

Some basic consequences of a reduced critical thickness will be briefly outlined.

5.1 Mean free slip path of TDs

As known the slip path of the moving TDs is limited, in general, by mutual impeding or reaction of the TDs. If two TDs move along intersecting slip planes, the probability of TD intersection increases with the film thickness. As a simple general approximation, Matthews [33] takes the minimum passing distance (corresponding to the range of interaction) of two TDs equal to the film thickness \( h \). In this way, he defines a mean free slip path of the TDs

\[
\lambda = \frac{1}{h \rho},
\]

(6)

at a given area density \( \rho \) of the TDs. The lower this density and the lower the film thickness at the onset of slip – i.e. the lower the critical thickness – the larger is the mean free slip path.

5.2 Removal of TDs by slip off the film

In a favourable case of low film thickness, a low density of TDs initially present may escape by slip until the rim of the film if their mean slip path becomes as large as the wafer diameter. Examples satisfying a mean slip path (6) as large as 100 mm are listed in Table 2. These data outline approximate conditions where TDs can slip without mutual interaction and leave the film.

As a result of this slip, a certain relaxation occurs: If TDs of any density \( \rho \) slip a mean path \( \lambda \), the resulting strain (relaxation) is given by the known [34] relationship

\[
\varepsilon_r = \frac{b \lambda \rho}{2},
\]

(7)

where \( \rho/2 \) accounts for one of the biaxial film directions and \( b \) is the relaxing component of the Burgers vector. Examples of \( \varepsilon_r \) corresponding to the data of Table 2 are given in its third column. It is seen that relatively low TD densities cover the range of common relaxations \( \varepsilon_r \), apparently because of the large slip path supposed. Vice versa, \( \varepsilon_r \) of Table 2 can be interpreted as the minimum epitaxial strain which is necessary to eliminate TDs of a density \( \rho \) by slipping a path of 100 mm. (These simple approximations do not include, e.g., the final equilibrium strain which remains by minimizing the total strain energy [30]).

It will be noted that Eq. (7) is equivalent with the commonly used relationship \( \varepsilon_r = \frac{b \lambda d_m}{2} \) where \( d_m \) is the mean distance of the MDs. The reciprocal \( 1/d_m = \rho \lambda \) represents the linear dislocation density obtained by projecting the area density \( \rho \) along the slip path \( \lambda \).

| Table 2 | Relaxation \( \varepsilon_r \) of (7) achieved in films of thickness \( h \) if TDs of density \( \rho \) and \( b = 0.4 \) nm slip a path (6) of \( \lambda = 100 \) mm. |
|-------|---------|--------|
| \( h \) (nm) | \( \rho \) (cm\(^{-2}\)) | \( \varepsilon_r \) (%) |
| 1000 | \( 1 \times 10^3 \) | 0.2 |
| 100 | \( 1 \times 10^4 \) | 2 |
| 10 | \( 1 \times 10^5 \) | 20 |
5.3 Reduction of common TD densities

The given TD density of heteroepitaxial films is frequently around $10^6$ to $10^{10}$ cm$^{-2}$, i.e. much higher than that of Table 2. What is the advantage of film growth on a compliant substrate in view of lowering a high density of TDs?

As seen from both, Eqs. (6) and (7), a high density $\rho$ is basically connected to a slip path much shorter than the wafer diameter. Therefore the moving TDs will frequently interact, including mutual blocking, reaction, and annihilation. (A deliberately high and homogeneous TD density, giving a high probability of mutual annihilation, is achieved in the ion implantation experiments [15].) To discuss the literature on TD interaction [e.g. 34, 35] would, however, exceed the scope of this paper. To give an idea only, we refer to a recent theory of Romanov et al. [35] on TD reduction in strained layers growing on an ordinary bulk substrate. The authors treat the interaction of TDs and MDs by differential equations. For the case of no blocking of TDs by MDs (a condition well justified as we consider a compliant substrate where the MDs relax or even disappear) the authors obtain an analytical solution $\rho = \rho_0 \exp \left(\frac{4r_A}{h}\right) \exp \left(-\frac{4r_A}{h_c}\right)$ where $\rho_0$ is the initial, starting TD density at the film thickness $h = h_c$ and where $r_A$ is the radius of TD interaction, taken as constant. This formula indicates that for a given final film thickness $h$, the TD density $\rho$ decays exponentially on lowering the critical thickness $h_c$. A modification can be made according to the Matthews approach [33] where the range of interaction is taken equal to the film thickness: For this purpose, we take $nr_A = h$ where $n$ is a geometry factor of order of magnitude 1. Thus, the differential equation (17a) of [35] changes slightly, resulting in a power law $\rho = \rho_0(h_c/h)^n$. This result also means that a lower $h_c$ decreases the final density of TDs at a given film thickness $h$. Both results indicate the advantage of using a compliant substrate. To provide an underlying more general idea, we suppose that a film of given final thickness can reduce its TD density the more the earlier a suitable dislocation mechanism can start, i.e. the lower its critical thickness. In addition, less blocking of TDs by MDs is expected if the MDs relax.

5.4 Homogeneity of relaxation

There is a further advantage of a low critical thickness: It will avoid multiplication of TDs at sources which usually start to operate at a considerably higher film thickness [36], forming an inhomogeneous TD density and the known cross-hatch surface pattern. Hence, a low $h_c$ might be advantageous to obtain more uniform relaxation.

5.5 Final degree of relaxation

As known the final equilibrium state of relaxation is determined by minimizing the total energy. The latter decreases by relaxing the misfit strain but increases by forming the MDs [30]. Therefore relaxation of the MDs principally improves the overall relaxation of the misfit, i.e. the remaining equilibrium strain is lowered.

6 Summarizing conclusions

The relief of heteroepitaxial strain in films grown on a compliant substrate can be basically understood by modifying the known Matthews model. To our view, it is essential that the weak interfacial layer of a compliant substrate (whatever its nature) is enabled to relax the strain field (line tension) of the misfit dislocations. The aim is to reduce the drag force at the moving TDs (irrespective whether they slip or climb). Thus, the critical thickness for the onset of plastic relaxation will be reduced. This conclusion is in contrast to current theories which predict an enhanced critical thickness on using a compliant substrate.
There are various favourable consequences of a reduced critical thickness:
– A low $h_c$ reduces the probability of interaction between TDs. Thus, a larger mean free slip path of the TDs is achieved.
– If the slip path is comparable to the wafer diameter (Table 2), a considerable magnitude of epitaxial strain can be relaxed on removing of a low TD density by slip off the film.
– If as usual the slip path is much lower so that mutual interaction of the TDs predominates, it is a complex and specific matter to explain how the usually high TD density can be reduced by means of reducing $h_c$. An indication of this by means of analytical formulae can be derived from a recent theory of dislocation interaction in films growing on common substrates [35]. We suppose that a film of given final thickness can reduce its TD density the more the earlier a suitable dislocation mechanism can start, i.e. the lower its critical thickness.
– Early relaxation at a lowered $h_c$ will avoid TD multiplication at sources which usually need a much higher film thickness to operate. Hence a low $h_c$ is favourable to relax a film more uniformly than in the case of TD multiplication.
– A more complete degree of relaxation is possible in the final stage of minimized total strain energy if the MDs relax their energy.

In order to experimentally form a compliant substrate, it seems essential to employ any means which allows to relax the line energy of the misfit dislocations. The conventional technique to obtain a compliant substrate means to form a “weak” template/substrate interlayer, e.g. by wafer bonding and thinning the template or, as proposed more recently, by internal oxidation before epitaxial growth [12]. Alternatively, one may omit the template and introduce a “weak” interlayer subsequently, e.g. by ion implantation [15] after epitaxial growth.

Acknowledgement

The author thanks U. Gösele for stimulating discussions.

References