Theory of Strain Relaxation for Epitaxial Layers Grown on Substrate of a Finite Dimension

F. Y. Huang*
IBM Microelectronics, 1580 Route 52, Hopewell Junction, New York 12533-6531
(Received 10 January 2000)

We present an equilibrium theory for strain relaxation in epitaxial layers grown on substrates of a finite dimension. The conventional dislocation model is refined to take account of the multiple reflection of image dislocations. The effect of strain transfer and dilution due to finite vertical and lateral dimensions of the substrate is also considered. The critical thickness has been obtained based on an energy balance approach. Detailed numerical analysis with primary experiments for the SiGe alloy system is also provided.

PACS numbers: 62.20.Dc, 62.20.Fe, 68.55.Jk, 68.60.Bs

Epitaxial layers grown on lattice mismatched substrates, such as SiGe alloy on Si, have found a wide variety of applications. The large lattice mismatch of about 4% between germanium and silicon, however, has limited the growth of high-quality SiGe alloys to within a certain thickness, the so-called critical thickness, beyond which misfit dislocations start to generate [1]. To circumvent this limitation, a novel approach via substrate engineering (i.e., tailoring the substrate to form a finite dimension in the vertical and/or lateral directions) has been proposed to transfer or dilute the misfit strain [2,3].

Theoretically, many different models have been established to predict the critical thickness for strained layers. Among the most celebrated ones are the Matthews-Blakeslee [1] (MB) and People-Bean [4] (PB) models for the equilibrium theory, and the Dodson and Tsao model for the kinetic theory [5]. Because of its clarity and reasonably good agreement with experiments, the equilibrium theory, especially the PB model, has been widely accepted.

However, the conventional dislocation model for epitaxial layers on a bulk substrate runs into serious difficulties when the substrate has finite dimension. This is particularly true when the thickness of the epilayer is comparable to either the substrate thickness (as in a *compliant* substrate) or the lateral dimensions of the substrate (as in a *mesa* structure). In addition to the lack of a dislocation model, there has been no treatment which can thoroughly describe the elastic strain in such systems. Establishing a rigorous theoretical model for strain relaxation in these systems is not only of scientific interest, it also has practical value. Technology advancement in patterning the mesa and thinning the substrate thickness down to a few hundred angstroms results in the need to analyze situations in which the epitaxial layer is of comparable thickness to the substrate dimension [2,6].

In this paper, we provide a theoretical analysis for dislocations and elastic strain in epitaxial layers grown on substrates of a finite dimension. The conventional dislocation model is refined to accommodate the multiple image dislocations generated from multiple surface boundaries. Strain transfer and dilution due to a finite substrate dimension have also been considered. The critical thickness is derived within the equilibrium theory, and detailed numerical analysis is provided for the SiGe alloy system. Recent experiments on compliant substrates are also discussed.

Figure 1 illustrates the structure of an epitaxial layer of thickness h_1 grown on a thin-film compliant substrate of thickness h_2. For a mesa substrate, the length and the width of the mesa are $2l$ and $2w$, respectively.

We first discuss the limitation of the conventional dislocation model when applied to a finite-sized substrate. For an epitaxial layer (*epilayer*) grown on a bulk substrate, the lattice constant a_1 will match that of the substrate a_s, provided the epilayer is coherently strained. The misfit is $f = (a_1 - a_s)/a_s$ and the strain energy per unit area is $E = Bf^2h_1$. Here, B is a material constant and given by $2\mu(1 + \nu)/(1 - \nu)$, with μ and ν the shear modulus and Poisson's ratio of the epilayer, respectively. For an isotropic medium with straight dislocations lying on the interface plane between the epilayer and the substrate, the energy per unit length for a dislocation is [7]

$$E_{\mathrm{dis}} = \frac{\mu b^2(1 - \nu\cos^2\theta)}{4\pi(1 - \nu)} \ln(\alpha h/b), \qquad (1)$$

where h, the outer radius associated with a dislocation line, is truncated at the surface of the epilayer with $h = h_1$, and

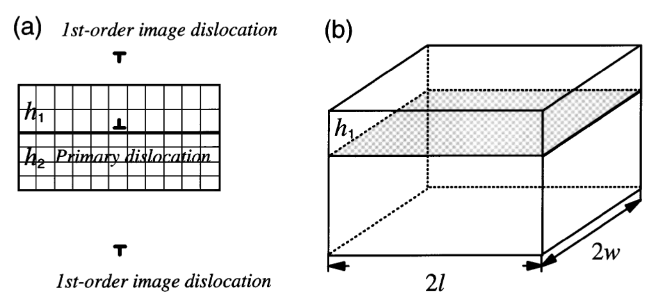

FIG. 1. (a) Schematic of an epilayer situated on a compliant substrate. The primary dislocation is lying on the interface between the epilayer and the substrate. The first-order image dislocations generated by the top surface of the epilayer and the bottom surface of the substrate are illustrated. (b) Mesa substrate of finite lateral dimension.

 0031-9007/00/85(4)/784(4)$15.00

b is the magnitude of the Burger's vector. The prefactor α is related to the uncertainty in the cutoff of the inner radius b and the outer radius, and is normally chosen between 1 and 4.

To apply Eq. (1) to the situation of a thin-film substrate, we simply replace h by the distance from the dislocation line to the nearest free surface, which is either the epilayer thickness or the substrate thickness, assuming that the dislocation spacing is much larger than the thickness of either film. The dislocation energy will then be equivalent to the case of an epilayer situated on a bulk substrate, a clearly nonrigorous result. The error arises from ignoring the surface boundary conditions. For an epilayer on a bulk substrate, the top surface is the only surface requiring the stress-free boundary condition. Thus, it is a common practice to ignore the trivial image term associated with this single interface [8]. However, in the case of a finite-sized substrate where multiple boundaries exist (analogous to a multilayer structure [9]), the contribution from all images is nontrivial and must be considered.

For a dislocation lying on the interface between an epilayer and a compliant substrate, there is another first-order image generated by the bottom of the substrate, in addition to the primary dislocation and the first-order image from the epilayer top surface. This, in turn, will generate a second-order image by the top surface, and so forth. The same situation will also apply to the top image. The total dislocation energy, therefore, must include the original as well as all the image dislocations generated by the top and bottom surfaces, and can be evaluated based on the superposition principle using the expression

$$2C_0 \ln(h_r/b) = C_0 \ln(h_1/b) + C_0 \ln(h_2/b) - \sum(\text{image}) . \tag{2}$$

Here, the constant C_0 represents $\mu b^2(1 - \nu \cos^2\theta)/[4\pi(1 - \nu)]$, with both layers assumed to have the same constant for simplicity. The last term in Eq. (2) is given by

$$\sum(\text{image}) = C_0 \ln\left(\prod_{m=1}^{n} \frac{(mH + h_1) \times (mH + h_2)}{[(m - 1)H + h_1] \times [(m - 1)H + h_2]}\right) = C_0 \ln\left(\frac{(nH + h_1)(nH + h_2)}{h_1 h_2}\right), \tag{3}$$

where $H = h_1 + h_2$, and n is the order of the image dislocations to be considered. It can be seen that when the number of terms in the sum approaches infinity the energy from all the image dislocations is *divergent*. To avoid the logarithmic divergence, we must truncate the infinite sum. This approach is similar to the treatment used in the original dislocation theory through the introduction of the cutoff radius [10].

When only the first-order image dislocations are considered, the total dislocation energy becomes

$$E_{\text{dis}} = C_0 \ln\left(h_1 h_2/\{[(H + h_1)(H + h_2)]^{1/2} b\}\right), \tag{4}$$

which can be expressed as $C_0 \ln(h_r/b)$ with $h_r = h_1 h_2/[(H + h_1)(H + h_2)]^{1/2}$. The dislocation on a compliant substrate can then be viewed as if on a "bulk" substrate, but with the renormalized thickness h_r replacing the actual thickness h_1 or h_2. Under the special condition that one of the films is much thicker than the other one, the above expression can be simplified to $1/h_r \approx (1/h_1 + 1/h_2)$, which can be further reduced to the identical result as for the case of a bulk substrate when one of the films approaches infinite thickness. When the epilayer has a thickness comparable to the substrate, the conventional model gives a dislocation energy $E_{\text{dis}} = C_0 \ln(h_1/b) = C_0 \ln(h_2/b)$. Our result, on the other hand, yields $E_{\text{dis}} = C_0 \ln[h_1/(3b)]$. As compared to the conventional treatment [2,6], in which only two extreme conditions have been separately considered with either the epilayer or the substrate thickness incorporated in the dislocation, the present model can evaluate the dependence of the dislocation energy as a function of both films.

Having obtained the dislocation energy, we now turn to discussing the elastic strain in a bilayer structure. If two freestanding thin films of lattice constant a_1 and a_2 are brought together in close proximity at thermal equilibrium, they will reach a common lattice constant a_0 and experience a misfit as if they were situated on a "virtual" substrate of a lattice constant a_0. The misfit is given by $f_1 = (a_1 - a_0)/a_0$ and $f_2 = (a_2 - a_0)/a_0$ for the epilayer and thin film substrate, respectively. The total energy per unit area related to the elastic strain of both films is $E_s = B_1 f_1^2 h_1 + B_2 f_2^2 h_2$. At thermal equilibrium, the total elastic strain energy of the bilayer structure will be minimized (i.e., $\partial E/\partial a_0 = 0$), which yields a common lattice constant of $a_0 = (B_1 h_1/a_1 + B_2 h_2/a_2)/(B_1 h_1/a_1^2 + B_2 h_2/a_2^2)$. When $B_1 = B_2$ and a_1 is close to a_2, the above expression can be reduced to the average lattice constant weighed by the film thickness $a_0 = (h_1 a_1 + h_2 a_2)/(h_1 + h_2)$ commonly used for multilayers [1].

The misfit strain in the epilayer is given by $\varepsilon_1 = f_0/\gamma_1$, with $f_0 = (a_1 - a_2)/a_2$ and the strain dilution factor $\gamma_1 = [1 + (B_1 h_1 a_1^2)/(B_2 h_2 a_2^2)]$. A similar expression holds for the substrate film with $\varepsilon_2 = f_0/\gamma_2$. It can be seen that, as the substrate thickness decreases, γ_1 increases and ε_1 decreases, implying strain transfer from the epilayer to the substrate film. The total elastic strain energy is given by

$$E_s = B_1 f_0^2 h_1/\gamma_1^2 + B_2 f_0^2 h_2/\gamma_2^2 . \tag{5}$$

Within the equilibrium theory, there are two different criteria to define the critical thickness for epilayers on a bulk substrate. The MB model requires that the total energy from the elastic strain and dislocations be at its minimum [1], while the PB model assumes that the energy from elastic strain is equal to that from dislocations [4]. Because of its better fit with experiment, we shall adopt the approximate PB model to define the critical thickness. Based on Eqs. (4) and (5), we obtain the critical thickness for the epilayer as a function of the substrate thickness:

$$B_1 f_0^2 h_{1c}/\gamma_1^2 + B_2 f_0^2 h_2/\gamma_2^2 = \left(\frac{C_0}{2\sqrt{2}\, a(x)}\right) \ln(h_r/b)\,. \tag{6}$$

For the SiGe alloy, $a(x)$ is usually chosen as the mean value of 5.54 Å. It can be verified that when the substrate thickness approaches infinity, $\gamma_1 = 1$ and $\gamma_2 = \infty$, Eq. (6) then yields the conventional critical thickness on a bulk substrate with $h_c = 13.3/x^2 \ln(h_c/b)$ for screw dislocations, where x is the germanium mole fraction [4]. A detailed numerical analysis will be given later in conjunction with the mesa substrate.

For substrates with a finite lateral dimension, we first analyze the elastic strain for an epilayer of a lateral length $2l$ and infinite width. The stress distribution can be described in a similar manner as a finite bimaterial assembly [2]. The strain energy density per unit volume is given by the energy density for an infinite slab, $B_1 f_0^2$, modulated by a distribution factor $\omega(y,z)$, with

$$\omega(y,z) = \left(1 - \frac{\cosh(ky)}{\cosh(kl)}\right)^2 \exp(-\pi z/l)\,. \tag{7}$$

In Eq. (7), k is the interfacial compliance parameter, and can be estimated by $\zeta(\nu)/h_e$, with $\zeta(\nu)$ about 0.93 for SiGe alloys, and the effective thickness $h_e \equiv \int_0^h \omega(y,z)\, dz|_{y=0} \equiv h/[\phi(l/h)]^2$. The stress reduction factor $\phi(l/h)$ is given by [2]

$$\phi(l/h) = [(1 - \mathrm{sech}kl)^2 (1 - e^{-\pi h/l}) l/(\pi h)]^{-1/2}. \tag{8}$$

This factor contributes to a reduction in the effective strain, $f = f_0/\phi$. In the previous treatment [2], however, only vertical stress distribution is considered, and the maximum stress at $y = 0$ for the y dependence is assumed. The distribution in the y direction should also contribute to the stress reduction by a factor of $\psi(l/h)$, with the effective length of the ridge, $l_e \equiv \int_0^l \omega(y,z)\, dy|_{z=0} \equiv l/[\psi(l/h)]^2$, and

$$\psi(l/h) = \left[\left(1 - \frac{\tanh kl}{kl}\right) + \frac{[1 + (\sinh 2kl)/(2kl)]}{1 + \cosh 2kl}\right]^{-1/2}. \tag{9}$$

The strain energy per unit area is therefore given by $E_s = B_1 f_0^2 h_1/[\phi(l/h)\psi(l/h)]^2$. If the ridge is further confined in the other lateral direction with a finite width w to form a mesa structure, the stress reduction factor will contain another term φ^2, which has the same form as Eq. (9) with l replaced by w.

For a mesa structure without substrate compliance, the energy balance criterion yields a critical thickness:

$$B_1 f_0^2 h_{1c}/(\psi^2\phi^2\varphi^2) = \left(\frac{C_0}{2^{1/2} a(x)}\right) \ln\left(\frac{h_r}{b}\right), \tag{10}$$

where the renormalized thickness h_r can be considered in a similar manner as in a compliant substrate. It can be seen that Eq. (10) reduces to the previous result with $\phi = \varphi = 1$, and h_r is replaced by h_1 or l as one of the dimensions is much larger than the other [2].

For a mesa structure with substrate compliance, the critical thickness is given by

$$B_1 f_0^2 h_{1c}/\Gamma_1^2 + B_2 f_0^2 h_2/\Gamma_2^2 = \left(\frac{C_0}{2^{1/2} a(x)}\right) \ln(h_r/b)\,. \tag{11}$$

Here, $\Gamma_1 = \gamma_1 \psi_1 \phi_1 \varphi_1$ for the epilayer, and a similar expression holds for the substrate thin film Γ_2. For practical applications, the substrate can be thinned down to about a few hundred angstroms, while the mesa size has to be kept at least on the thousand angstroms scale for active devices. The renormalized layer thickness h_r, therefore, will be dominated by the substrate thickness as given by Eq. (4), when the first-order images are considered.

Using the material constant for SiGe with $\nu = 0.28$ and $b = 4$ Å, the critical thickness of the SiGe alloy as a function of the germanium content is shown in Fig. 2 for different substrate thickness and mesa size. As can be seen from Fig. 2(b), there exists a threshold value for the Si substrate

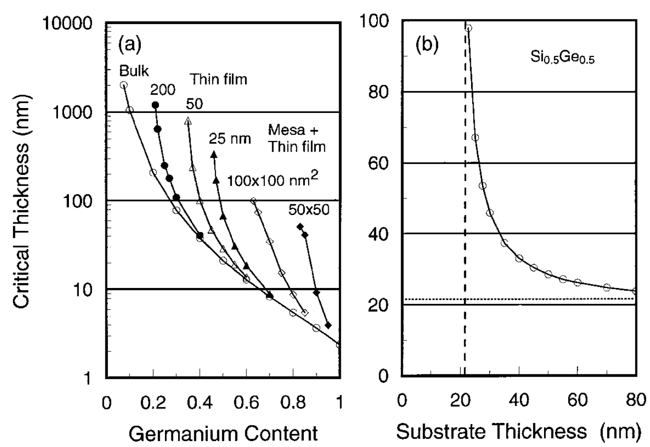

FIG. 2. (a) Critical thickness of SiGe alloy as a function of germanium content on different substrates (from left to right): bulk substrate, compliant substrate with a 2000, 500, and 250 Å Si film, and mesa substrate of 0.1×0.1 μm^2 and 0.05×0.05 μm^2 situated on a 250 Å Si thin film. (b) Critical thickness of $Si_{0.5}Ge_{0.5}$ as a function of the Si substrate thickness. As the Si substrate thickness decreases, the critical thickness of SiGe increases continuously above the conventional value on a bulk substrate. There exists a threshold value for the Si substrate (~200 Å), below which $Si_{0.5}Ge_{0.5}$ of an infinite thickness can be grown without dislocations, which has been confirmed by experiment.

thickness, below which a SiGe alloy of any thickness can be grown without generating misfit dislocations.

Practically, the mesa structure can be formed by microfabrication, while the compliant substrate can be realized by using the silicon-on-insulator (SOI) structure with subsequent thinning of the top Si film. In a SOI substrate, the buried oxide will become viscous and reflow upon high temperature anneal allowing for the strain transfer [3,11]. However, the high temperature treatment ($\sim$1100 °C) will introduce dislocations, inhomogeneity, and germanium segregation, and, hence, degrade the SiGe quality. This problem can be solved by implanting boron onto the buried oxide to form boron-silicate-glass (BSG) SOI. The lower viscous temperature of BSG will lead to a lower anneal temperature. In a recent experiment on an extremely thin ($\sim$200 Å) Si thin film situated on a boron-implanted SOI substrate, almost complete strain relaxation ($\sim$95%) for the 1200 Å $Si_{0.7}Ge_{0.3}$ alloy has been obtained, and near-band-gap photoluminescence indicating the high quality of the SiGe alloy has been observed [12]. Future investigation on the critical thickness will be needed to make a detail comparison with the present theory.

In summary, we provide a theoretical model for strain and dislocations in epitaxial layers grown on substrates of a finite dimension, which extends the scope of the existing theories beyond the conventional limitation of a bulk substrate. Numerical analysis is carried out for the SiGe alloy system and compared to the conventional results based on a bulk substrate. Primary experiments for SiGe alloy grown on extremely thin Si films are also discussed.

Helpful discussions and critical review of the manuscript from Dr. David R. Greenberg and Kathy T. Schonenberg of IBM are acknowledged. I would also like to thank Dr. B. Meyerson, Dr. S. Iyer, and Dr. S. Subbanna of IBM for continuous support and encouragement.

*Electronic address: huangf@us.ibm.com

[1] J. W. Matthews and A. E. Blakeslee, J. Cryst. Growth **27**, 118 (1974); **32**, 265 (1976).
[2] S. Luryi and E. Suhir, Appl. Phys. Lett. **49**, 140 (1986).
[3] A. R. Powell, S. S. Iyer, and F. K. Legoues, Appl. Phys. Lett. **64**, 1856 (1994).
[4] R. People and J. Bean, Appl. Phys. Lett. **47**, 322 (1985).
[5] J. Y. Tsao, B. W. Dodson, S. T. Picraux, and D. M. Cornelison, Phys. Rev. Lett. **59**, 2455 (1987).
[6] F. Y. Huang and K. L. Wang, Philos. Mag. Lett. **72**, 231 (1995).
[7] J. P. Hirth and J. Lothe, *Theory of Dislocations* (McGraw-Hill, New York, 1985).
[8] The image dislocation in a bulk substrate contributes a constant term $-\mu b^2(1 - \nu \cos^2\theta)/[4\pi(1 - \nu)]$ to the dislocation energy, which can be incorporated into the logarithmic prefactor with $\alpha_0 = \alpha/e$ in Eq. (1).
[9] J. Y. Shen, J. P. Hirth, F. W. Zok, and J. A. Heathcote, Scr. Mater. **38**, 15 (1997).
[10] Analogous to the prefactor α in the original dislocation energy, the truncation of the infinite series of the image dislocations will contribute to another uncertainty factor.
[11] M. A. Chu, M. O. Tanner, F. Y. Huang, K. L. Wang, G. Chu, and M. S. Goorsky, J. Cryst. Growth **175**, 1278 (1997).
[12] F. Y. Huang, M. A. Chu, K. L. Wang, G. U'Ren, and M. S. Goorsky, Appl. Phys. Lett. **76**, 2680 (2000).