[논 문] 한국재료학회지
Korean Journal of Materials Research
Vol. 11, No. 7 (2001)

高變形된 異種 에피층에서 응력 집중이 결정결함 생성에 미치는 영향

김 삼 동 · 이 진 구
동국대학교 전자공학과

Stress Concentration Effects on the Nucleation of the Structural Defects in Highly Strained Heteroepitaxial Layers

Sam-Dong Kim and Jin-Koo Rhee
Department of Electronic Engineering, Dongguk University, 3-26 Pildong, Joonggu, Seoul 100-715, Korea

(2001년 3월 16일 받음, 2001년 7월 4일 최종수정본 받음)

초 록 본 연구에서는 고변형된 이종 에피층에서 두 가지 종류의 반원 전위 루프 (60° 및 쌍격자 전위)의 생성 속도를 예측하는 모델을 제안한다. 모델링 시, 에피층 표면에서 발생하는 결함과 이곳에 집중되는 응력 효과를 고려하였으며, Matthew의 식을 발전시켜 에피층 두께에 따른 잔류 변형율을 변수로 사용하였다. 모델링을 통한 계산 결과에 의하면, 응력 집중 현상은 고변형된 이종 에피층에서 전위 및 결정 결함 현상을 설명하는 데 매우 중요하였다. 또한, 본 연구를 통하여, 응력 집중 현상이 에피층 성장 초기에 생성되는 전위 형태를 결정하는 주요한 인자 중 하나임을 단면 투과 전자 현미경 결과와의 비교를 통해 확인할 수 있었다.

Abstract We carried out the kinetic model calculations in order to estimate the nucleation rates for two kinds of half-loop dislocations in highly strained hetero-epitaxial growths; 60° dislocations and twinning dislocations. The surface defects and the stress concentration effects were considered in this model, and the remaining elastic strain of the epilayers with increasing film thickness was taken into account by using the modified Matthews' relation. The calculations showed that the stress concentration effect at surface imperfections is very important for describing the defect generation in highly mismatched epitaxial growth. This work also showed that the stress concentration effect determined the type of dislocation nucleating dominantly at early growth stages in accordance with our XTEM (cross-section transmission electron microscopy) defect observation.

Key words : highly strained hetero-epitaxy, dislocation, nucleation, kinetics, stress-concentration, TEM

1. Introduction

Many recent advances have been aimed for reducing the defect density of highly strained InGa(Al)As/GaAs, GaAs/Si, and SiGe/Si hetero-epitaxial layers by using a variety of treatments and growth schemes, such as post-annealing, SLS (strained layer superlattice) growth and compositionally graded InGaAs buffer growth. However, other evidences have been reported which shows that the surface smoothness significantly affects the defect density of the highly strained heteroepitaxy.[1] This suggests that defect nucleation is related to the epilayer surface condition in the early stage of growth. Although considerable theoretical work has been performed to understand the dislocation nucleation mechanism in the coherent strained system, they have focused mainly on a calculation of the equilibrium critical thickness, such as the studies by van der Merwe[2], Matthews and Blakeslee[3] and Fritz.[4] The prediction of the critical thickness is very important because the performance of some modern devices, such as QW (quantum well) devices and high speed FET (field effect transistor), is very sensitive to the density of structural defects and the resulting electronic defects. However, as the mismatch between the epilayers and the substrates becomes larger, dislocation formation can not follow the equilibrium mechanics. Moreover, since the highly strained epilayer nucleates and grows in a 3D mode, surface imperfections of the epilayer surface must be considered.

A kinetic model of the half-loop nucleation was suggested by Dodson[5] for the calculation of equilibrium critical thickness. We develop this model for the highly strained hetero-epitaixal system in order to investigate the role of the surface roughness and to examine the mechanism of stacking fault formation. This study includes theoretical estimations of the nucleation rates for the 60° perfect dislocations and the pure-edge partial dislocations. We examine how the stress concentration due to the surface roughness of the initial epilayer

affects the nucleation kinetics for these two kinds of dislocations. The stability condition for stacking faults is also examined through a simple model that is based upon the difference in work-effect between the glide of separate partial dislocation and the recombination process of two partial dislocations to form a perfect dislocation.

2. Energy balance in dislocation nucleation in highly strained epilayers

As was suggested by the previous dislocation nucleation mechanisms, the major dislocation nucleation sources for the highly strained epitaxial systems, such as GaAs/Si and $In_xGa_{1-x}As/GaAs$, are most probably provided from the top surface of the epilayers.[6] The substrates can provide some portion of them, however most of misfit dislocations nucleate at the top surface for producing the high density of threading dislocations. If we consider the growing top surface of the epilayer as an imaginary reservoir for an infinite number of dislocations, then the half-loop dislocation nucleation from the top surface becomes one of the most probable mechanisms. In this model, we consider the surface imperfections as heterogeneous nucleation sites, and estimated the nucleation efficiency during the growth.

Based upon continuum elasticity, we can express the total energy change, ΔE_{total}, due to a half-loop nucleation as:

$$\Delta E_{total} = \Delta E_{loop} - \Delta E_{work} + \Delta E_{fault} \quad (1)$$

where ΔE_{loop} is the half-loop energy, ΔE_{work} is the amount of work done by the expansion of the half-loop and ΔE_{fault} is the stacking fault energy created by the expansion of an imperfect dislocation. Each energy contribution in equation 1 is expressed in the following equations. The half-loop energy consists of the half-loop strain energy, E_d, and the energy change due to the surface step annihilation:

$$\Delta E_{loop} = \Delta E_d - 2\rho\gamma_{111} b \sin\alpha \quad (2\text{-a})$$

$$\Delta E_d = \frac{\mu b^2(-\nu)}{8(1-\nu)} \rho [1 + \delta \ln(\rho/b)] \quad (2\text{-b})$$

where the elastic properties of the epilayer are assumed to be isotropic, μ is the shear modulus of the film, b is the magnitude of the Burgers vector, ν is the Poisson ratio, ρ is the radius of half-loop, γ_{111}[7] is the {111} surface energy, and α is the angle between the

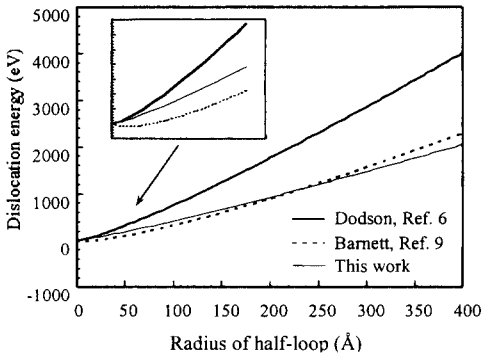


Fig. 1. Plots of the dislocation energy versus the half-loop radius. Three different expressions are shown for the elastic energy of half-loop dislocations.

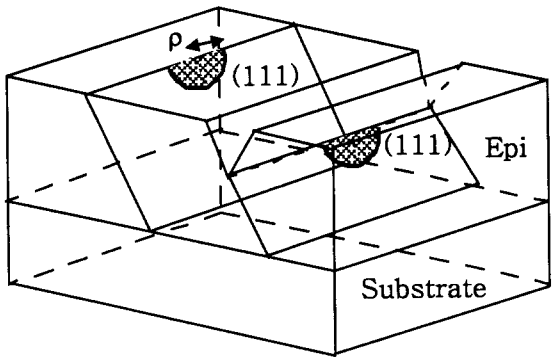


Fig. 2. A schematic illustration of two kinds of the dislocation nucleations; homogeneous and heterogeneous nucleations

Burgers vector direction and the line vector direction of the dislocation. Dodson[5] estimated the strain energy for the half-loop dislocation as simply a half of the circular loop dislocation in the bulk crystal. Compared to a more careful estimation of the half-loop energy by Barnett[8], Dodson's half-loop energy is an over-estimate, as shown in figure 1. Barnett's estimation may be correct, but it has negative values in the wider range of small half-loop radius (1~16Å), as shown in the zoomed view of figure 1. Therefore, this expression is not proper in estimating the half-loop energy at small radius. We use the correction factor, δ (=0.403), in a modified Dodson's expression. This not only allows Dodson's estimate to fit the Barnett's curve over the entire range of half-loop radius but also reduces the range of negative ΔE_d at small radius. Dodson also thought that the nucleation of a dislocation on the flat surface creates a new surface step. However, the nucleation at a surface ledge, which eliminates a surface step, must be energetically favored. This type of heterogeneous nucleation for the dislocation nucleation is described in a schematic of the figure 2. The effect of heterogeneous

nucleation is expressed by the second negative term in equation 2-a. The amount of work done by the expansion of the half-loop is given by:

$$\Delta E_{work} = \int \tau b dS = \pi \mu b K \varepsilon \rho^2 (1+\nu)(1-\nu)^{-1} \cos\phi \cos\lambda \quad (3)$$

where τ is the shear stress acting on the dislocation loop, K is a stress concentration factor which reflects the ratio between the built-up local stress at the surface notch and remote stress in the bulk region of the film, ε is the remaining elastic strain in the film, ϕ is the angle between the surface normal and the normal to the slip plane and λ is the angle between the surface normal and the slip direction. If the half-loop is an imperfect dislocation, we have an additional energy increase, such as:

$$\Delta E_{stacking\ fault} = \pi \rho^2 \gamma_f / 2 \quad (4)$$

where γ_f[9] is the stacking fault energy.

3. Nucleation kinetics for two kinds of dislocations

Figure 3 shows the total energy change due to the dislocation nucleation versus the radius of the half-loop for two different types of the dislocations; 60° and pure-edge partial (twinning dislocation) in a typical system of GaAs/Si. In figure 3, we arbitrarily take the remaining compressive strain in the GaAs layer as 0.03, which is only ~25% smaller than the initial misfit strain of 0.04, and use two stress concentration factors of K=1 and 5. The calculations are shown in figure 3, and these exhibit two results. First, the half-loop nucleation from the stress concentrated region (K=5) has a much lower activation energy, E*, than that from a flat surface (K=1). Thus, the nucleation rate, which is proportional to exp(-E*/kT), in a stress-concentrated region is much faster than in the flat surface. For example, the junction of island coalescence can provide these sites at the early stages of the epilaxial growths. If we consider 60° dislocation nucleation at an ideal flat surface, ~100eV of E* at r=70 Å half loop radius. This is a very high activation energy, and it leads to an very small nucleation rate compared to the real case. Second, the nucleation rate of the pure-edge partial dislocations creating stacking faults in the epilayer is much faster than that of the perfect 60° dislocation in the early stage of the strain relaxation (ε=0.03). This comes from the fact that, when the strain of the film is not sufficiently relaxed, the effect of total energy increase, which is due to the stacking fault formation caused by partial dislocation expansion, is much smaller than the energy required for half-loop expansion.

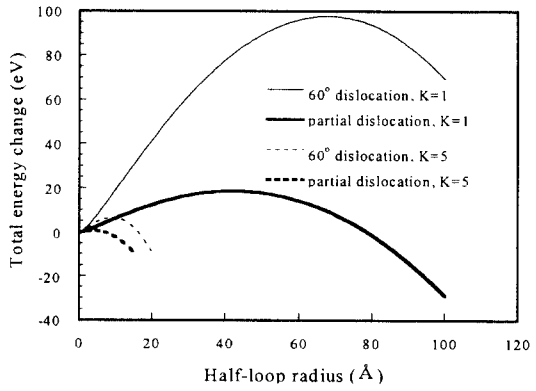


Fig. 3. Plots of the calculated energy change vs the half-loop radius for the dislocation nucleations at two different surface conditions (K=1 and 5)

In order to examine which type of dislocation plays a major role in relaxing the misfit strain at the early growth stage, it is important to calculate the activation energy variation with film thickness. For this calculation, we need a relationship between the remaining elastic strain in the film and the film thickness over the entire range of the epilayer layer thickness (1000Å for this work). Therefore, we can monitor theoretically the change of nucleation rate for both the twinning dislocations and the 60° dislocations during the growth. If we follow the continuum treatment of Matthews, we can derive a following expression for the planar elastic strain remaining in the film:

$$\varepsilon = \sqrt{6}\,\tau_f / M_{100} + \left[\frac{b}{\pi(1-\nu)}\left(\frac{\mu_f \mu_s}{\mu_f + \mu_s}\right)\right] \Big/ \left[M_{100}\frac{h}{\ln(\beta h/b)}\right] \quad (5)$$

where τ_f is the friction stress, M_{100} is the biaxial elastic modulus for the (100) film, μ_s is the shear modulus of the substrate, h is the film thickness, and β(=0.701) is a constant. The friction stress, τ_f, can be scaled as $M_{f/m}$, where m is an empirical constant. We utilized equation 5 to obtain a fit to our experimental data that were produced by the laser wafer curvature measurements. Figure 4 shows that the curve with a negligible contribution of friction stress ($m=10^5$) agrees best with the experimental data. The friction stress is regarded as an initial stress which is required to initiate the dislocation motion under a given stress state. The higher friction stress is present in a crystal, the wider

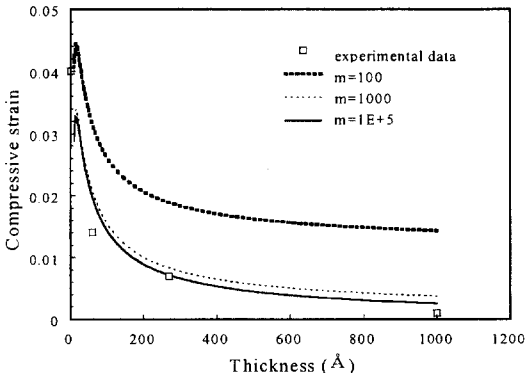


Fig. 4. Plots of the residual strain vs the epilayer thickness for GaAs/Si hetero-epitaxy. Measured strain of the film is plotted with open squares.

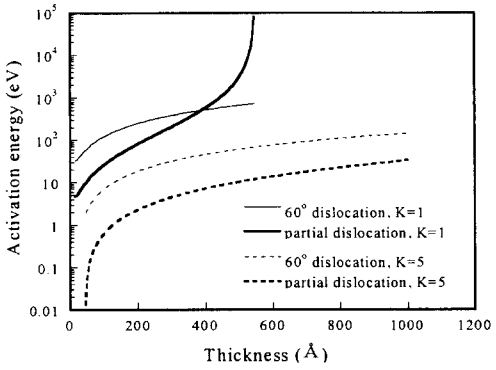


Fig. 5. Plots of the activation energy versus the epilayer thickness for half-loop dislocation nucleations in GaAs/Si hetero-epitaxy. Activation energies were calculated for the two kinds of dislocations (60° and pure-edge partial) at two different surface conditions (K=1 and 5).

range of elastic strain allows the metastable state for the dislocation nucleation.[10]

As shown in figure 3, both the perfect 60° dislocation and the twinning dislocation have each critical radii of half-loops, ρ_c, and activation energies, E*, for their nucleation. The critical radii of these half-loops are calculated by setting the derivative of energy with respect to ρ to zero in equation 2-a. The activation energy is obtained from equation 6 using this critical radius, ρ_c.

$$E^* = \Delta E_{loop}(\rho_c) - \Delta E_{work}(\rho_c) + \Delta E_{stacking\ fault}(\rho_c) \quad (6)$$

Perfect 60° dislocations do not require the third term, $\Delta E_{stacking\ fault}(\rho_c)$, since it does not produce the stacking fault. The calculated activation energies for two kinds of the dislocations versus the epilayer thickness are shown in figure 5. We used the fitted data of the remaining elastic strain, ε, versus the GaAs film thickness in figure 4 for this calculation. As shown in figure 5, the half-loop nucleation is strongly preferred at stress concentrated regions. If the source of the stress concentration comes from the surface-notch with a depth, C and tip radius, R, the stress concentration factor, K, is expressed as $2(C/R)^{1/2}$.[11] The sharpest surface notch formed by the initial island coalescence can have a radius at the notch tip, which is comparable with the lattice parameter of epilayer crystal. For example, if two 150Å tall GaAs islands coalesce, a local stress of five times larger than the applied remote stress is built up at the junction. This stress concentration (K=5) can have much lower activation energy than at the flat surface (K=1), as shown in figure. 5. Impurity clusters on the film can be the stress concentration sites for activating heterogeneous nucleation, as discussed by Dodson.[6]

4. Stacking fault formation at surface defects

As Pirouz suggested[12], the stacking faults can be formed simply by the wrong arrangement of atoms on (111) facets of the 3D islands in the early stage of growth. However, our calculated results suggests that the slips of pure-edge partial dislocations play a major role in stacking fault formation and relax misfit strain at early stages of the growth. In the stress concentrated region (K=5), the calculations show smaller activation energy for the nucleation of pure-edge partial dislocation than for perfect 60° dislocation over the entire 1000Å InGaAs layer thickness. On the flat surface (K=1), the partial dislocations nucleate faster than the 60° dislocations at a thickness less than 380Å. At larger thickness, the 60° dislocations will nucleate faster. This transition at film thickness of 380Å can be explained as follows. When the remaining elastic strain is large with small film thickness, the stacking fault energy is negligibly small compared to the amount of work done by any type of dislocations, thus the pure-edge partial dislocations nucleates faster due to their larger work-effect. After the film thickness reaches this transitional point, the stacking fault energy becomes more significant than before. This is because the amount of energy released from dislocation glides decreases with the film thickness. Above this thickness, the 60° dislocations start to nucleate faster than the pure edge partial dislocations.

The transition thickness has the following significance. Below a thickness of the strained epilayer, huge misfit stress drives faster nucleation of the partial dislocations, specially at the surface imperfections. Above

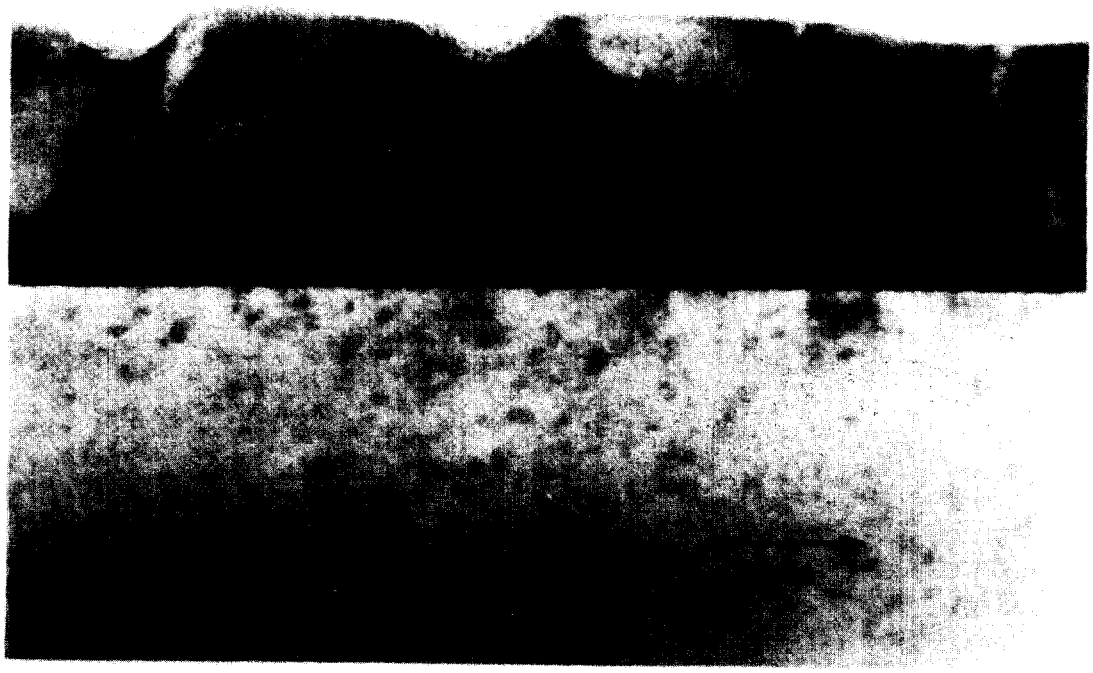


Fig. 6. Cross-section TEM images of the 1000Å thick GaAs/Si sample. A series of stacking faults are observed at the surface notches, but very few in the flat regions.

the transition thickness, the epilayer already relaxed much amount of the misfit strain. Therefore, the 60° dislocations start to nucleate dominantly during the further growth. However, at stress concentrated regions, the nucleation of pure-edge partial dislocations and the stacking fault formation is dominant although the epilayer thickness becomes larger than the transitional thickness. This phenomenon clearly coincides with our observation in the cross-section TEM micrograph of figure 6. 1000Å thick GaAs/Si samples were grown by MBE (molecular beam epitaxy) at 300℃ and at a growth rate of 0.3㎛/hr. A concentration of stacking faults is clearly observed at the surface notches, but very few in the flat regions. The morphology of these stacked stacking faults also indicates that they are produced from the nucleations of partial dislocations. If we consider these surface notches as the junctions of island coalescences, the "Λ" shapes of these stacked stacking faults could not arise from the mis-stacking of atoms on the island (111) facets but from the nucleation by partial dislocations.

5. Stacking fault stability

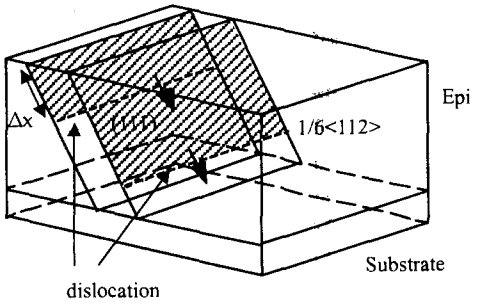
In the previous section, we discussed the nucleation kinetics for two kinds of the dislocations. Although nucleation kinetics determines which kind of dislocation nucleates faster, it does not describe the stability of the resulting stacking faults after the nucleations. We postulate the two simple situations in figure 7, where two successive dislocations glide into the fcc crystals. Figure 7 (a) shows the glide of a pure-edge partial dislocation by an infinitesimal distance, Δx, into the crystal where the same kind of dislocation is present on the adjacent glide plane. The energy increment, ΔE_{sf}, produced by this process is given by:

$$\Delta E_{sf} = \gamma_f \Delta x - \tau b_p \Delta x \qquad (7)$$

where τ is the resolved shear stress applied to a dislocation and b_p is the magnitude of the Burgers vector, 1/6 [112], for a pure-edge partial dislocation. Therefore, we obtain the driving force for successive stacking fault formation, F_{sf}, which is:

$$F_{sf} = -\frac{\Delta E_{sf}}{\Delta x} = -\gamma_f + \tau b_p \qquad (8)$$

Figure 7 (b) shows the glide of a 30° partial dislocation with an infinitesimal distance, Δx, on a slip plane which is swept by a pure-edge partial dislocation. The glide of the following 30° dislocation eliminates the stacking fault formed by the leading pure-edge partial dislocation. The incremental energy, ΔE_{60}, and the driving force, F_{60}, for this process are given by:



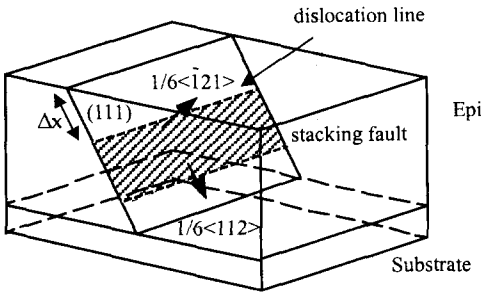


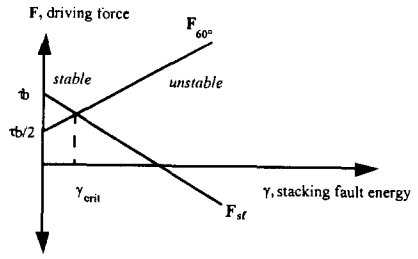
Fig. 7. Schematic illustrations of (a) the successive stacking fault formations due to glide of two edge partial dislocations and (b) the stacking fault annihilations involving the recombination of edge and 30° partial dislocations in a strained fcc crystal.

$$\Delta E_{60} = -\gamma_f \Delta x - \tau b_{30} \Delta x/2 \quad (9)$$

$$F_{60} = -\frac{\Delta E_{60}}{\Delta x} = \gamma_f + \tau b_w/2 \quad (10)$$

where b_{30} is the magnitude of Burgers vector, 1/6[-121], for the 30° partial dislocation. If this process has a larger driving force than that of the process described in figure 7(a), the stacking fault produced by pure-edge partial dislocation will be eliminated by the recombination of the two partial dislocations, such as 1/6 [112] + 1/6[-121] = 1/2[011]. This reaction creates a perfect 60° dislocation. The magnitudes of the driving force for these two processes depend on both the crystal stacking fault energy, γ_f, and the applied shear stress, τ.

Figure 8(a) and (b) show the schematic plots of the driving forces for these two processes versus the stacking fault energy and the applied shear stress. Figure 8 (a) agrees with the common sense argument that the stacking fault becomes less stable for the crystals with higher stacking fault energy. However, the stability of the stacking fault also depends on the shear stress applied to crystal, as shown in figure 8(b). There exists a critical shear stress, τ_{crit}, above which the stacking fault is stable in a crystal with stacking fault energy, γ_f. This critical shear stress can be expressed as $4\gamma_f/b$, where b



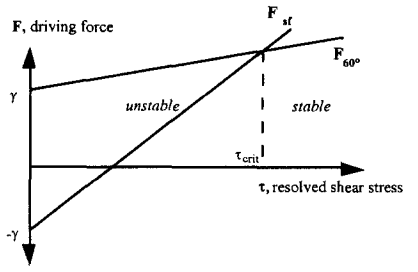


Fig. 8. Schematic plots of the driving forces for the stable stacking fault formation versus (a) the crystal stacking fault energy and (b) the resolved shear stress applied to an edge partial dislocation.

is the magnitude of Burgers vector for the partial dislocation. If we substitute a value of γ_f for GaAs crystal, we obtain a value of ~0.9 GPa at this critical shear stress. The unrelaxed resolved shear stress of τ applied to the dislocation is given by:

$$\tau = \cos\phi\cos\lambda M_{100}\varepsilon \quad (11)$$

where ϕ is the angle between the surface normal and the normal to the slip plane, λ is the angle between the surface normal and the slip direction, M_{100} is the biaxial elastic modulus for the (100) epilayer film and e is the remaining elastic strain in the film. Since the unrelaxed misfit strain between the epilayer and the substrate is 0.04 for the case of GaAs/Si system, the resolved shear stress is 2.3 GPa. Therefore, stacking faults formed during the initial growth stage of this system are located within the stable region shown in figure 8(b). If we substitute t in equation(11) for the critical shear stress, 0.9 GPa, we have 0.0156 for the remaining elastic strain, ε. This means that stacking faults produced by pure-edge partial dislocations are not eliminated by recombination until 61 % of the initial misfit strain in the GaAs film is relieved. If we assume that the elastic strain versus GaAs film thickness follows the curve fitted in figure 2, the stability of stacking faults is maintained until a 100Å thick GaAs film is grown. This sta-

ble thickness is, however, larger at the surface notch because it increases with the magnitude of stress concentration factor. A surface notch gives a stress concentration factor of 5, then the unrelaxed shear stress becomes 11.5 GPa within this region. Therefore, the stacking fault needs a 92 % relaxation of initial misfit strain before it becomes unstable. This yields a thickness of a 800 Å thickness for a GaAs film. The stacking fault nucleated from the highly stress concentrated region can be stable up to much larger film thickness.

Although this model is too simple to include all details of the strain relaxation phenomena in highly strained hetero-epitaxial growth, it explains why the high density of stacking faults is observed only within the limited area close to the interface region of the epilayers. The high density of stacking faults at the surface grooves, as observed in the TEM micrograph, is supposed to be related to this phenomenon.

6. Conclusions

We carried out kinetic model calculations for the dislocation nucleations in highly strained hetero-epitaxial systems and investigated the role of pure-edge dislocations in stacking fault formation. The activation energies for the nucleations of 60° and pure-edge partial dislocations were calculated with the epilayer thickness. At stress concentrated regions, the calculated results showed that the pure-edge partial dislocations nucleate much faster. The nucleations at flat surfaces proceed in two steps; the faster nucleation of the pure-edge partial dislocations at early growth stage, and the faster nucleation of the perfect dislocations after the transition film thickness. The GaAs/Si heteroepitaxial layers were examined by the XTEM and showed the concentrations of the stacking faults at surface notches produced by the partial dislocations.

We proposed a stability model for the stacking faults in the strained fcc crystals which suggests that the stability of stacking faults depends strongly on the surface roughness and the stress concentration. The stress concentration contributes not only to the nucleation rate of stacking faults, but also to the stable film thickness of the stable stacking fault formation.

Acknowledgements

This work is supported by KOSEF under the ERC program through the Millimeter-wave INnovation Technology research center at Dongguk University.

References

1. J. E. Palmer, G. Burns, C. G. Fonstad, and C. V. Thompson, Appl. Phys. Lett. **55**, 990 (1989).
2. J. H. van der Merwe, J. Appl. Phys. **34**, 123 (1963).
3. J. W. Matthews and A. E. Blakeslee, J. Cryst. Growth **27**, 118 (1974).
4. I. J. Fritz, S. T. Picraux, L. R. Dawson, T. J. Drummond, W. D. Laidig, and N. G. Anderson, Appl. Phys. Lett. **46**, 967 (1985).
5. B. W. Dodson, Phys. Rev. (B) **35**, 5558 (1987).
6. B. W. Dodson, Appl. Phys. Lett. **53**, 394 (1988).
7. D. K. Choi et al., Journal of Crystal Growth **85**, 9 (1987).
8. S. D. Gavazza and D. M. Barnett, Journal of Mechanics and Physics of Solids **24**, 171 (1976).
9. H. Gottschalk et al., Phys. Stat. Sol. (a) **45**, 207 (1978).
10. J. Y. Tsao and B. W. Dodson, Appl. Phys. Lett. **53**, 848 (1988).
11. G. E. Dieter, Mechanical Metallurgy (McGraw-Hill, 1976).
12. P. Pirouz, F. Ernst and T. T.Cheng, Heteroepitaxy on Silicon: Fundamentals, Structure and Devices (MRS, Pittsburgh PE., USA, 1988).