

Papers

A Two-Dimensional Nonisothermal Finite Element Simulation of Laser Diodes

Gen-Lin Tan, Naveen Bewtra, Keith Lee, and J. M. Xu, *Senior Member, IEEE*

Abstract—A new fully self-consistent nonisothermal, two-dimensional model of a semiconductor laser device is presented. The model consists of the simultaneous solution of the electrical equations (Poisson's and electron and hole continuity equations), along with the wave equation, photon rate equation, and thermal conduction equation. An analysis is presented for an AlGaAs–GaAs ridge laser diode structure using this model as a representative example. The results agree well with available experimental data. A comparison of the results between isothermal and nonisothermal simulations shows that the nonisothermal case has a higher threshold current and lower quantum efficiency than the idealized isothermal model as expected. The result was found to depend critically on the thermal exchange boundary condition of the simulated device, demonstrating the importance of considering thermal exchange in the design of laser diodes.

NOMENCLATURE

B	Radiative recombination coefficient.
c_0	Speed of light in vacuum.
C_p	Specific heat in constant pressure.
\bar{E}	Optical field.
E_g	Energy bandgap.
E_p	Optical field for p th mode.
G	Carrier generation rate.
G_{av}	Impact ionization generation rate.
G_p	Mode gain for p th mode.
g	Local gain.
g_c, g_r	Degeneracy factor for conduction and valence band.
h	Planck constant.
J_n, J_p	Current density vectors for electron and hole.
k	Boltzmann constant.
k_0	Wavenumber in free space.
L_c	Cavity length.
μ_{\min}	Minimum mobility.
μ_{\max}	Maximum mobility.

N_{ref}	Reference concentration.
N_A, N_D	Acceptor and donor impurity concentrations.
N_c, N_v	Effective densities of states in conduction and valence bands.
n	Electron concentration.
\bar{n}	Complex refractive index.
n_{eff}	Effective refractive index.
n_i	Intrinsic carrier concentration.
n_i^Γ	Electron concentration in Γ conduction band.
P	Light output.
p	Hole concentration.
q	Electron charge.
R	Carrier recombination.
R_{aug}	Auger recombination rate.
R_f, R_r	Front and rear facet reflectivities.
R_{sp}	Spontaneous emission recombination rate.
R_{SRH}	Shockley–Read–Hall recombination rate.
R_{st}	Stimulated emission recombination rate.
S_p	Photon density for p th mode.
T	Lattice temperature.
V_T	Thermal voltage ($= kT/q$).
β	Propagation constant.
V_n, V_p	Band parameter for electrons and holes.
ϵ	Dielectric constant.
λ_0	Emission wavelength.
μ_n, μ_p	Electron and hole mobilities.
τ_{ph}	Photon lifetime.
Φ_n, Φ_p	Slotboom variables.
ϕ_n, ϕ_p	Quasi-Fermi potentials for electron and hole.
χ	Electron affinity.
κ	Thermal conductivity.
Ψ	Electrostatic potential.
$\Delta E_D, \Delta E_A$	Energy level of donor and acceptor relative to conduction and valence band edge.

Manuscript received March 2, 1992; revised July 24, 1992. This work was supported in part by the Natural Science and Engineering Research Council of Canada, Ontario Laser and Lightwave Research Center, and Bell-Northern Research, Ltd.

The authors are with the Department of Electrical Engineering, University of Toronto, Toronto, Canada M5S 1A4.

IEEE Log Number 9206523.

I. INTRODUCTION

OVER the past decade, semiconductor laser diodes have become an increasingly interesting subject of research in the field of optoelectronics. Along with the high performance of recent lasers comes an increase in

structural complexity. Accurate device modeling can be used to assist in the optimization and development process and reduce the cost in experimental implementation of new device structures and in optimization.

There has been several works on two-dimensional (2-D) simulation of laser diodes [1]–[6]. Some of the references [1]–[2] perform only a quasi-2-D analysis, whereas others [3]–[4] incorporate a 2-D finite difference analysis to solve the electrical equations and a quasi-2-D effective index method to solve the wave equation. In [5], the problem is approached two dimensionally, however, the electrical equations are solved using a finite difference method while the wave equation is solved using a finite element method. Although this model uses two 2-D approaches, the methods of solution are not consistent. In [1]–[5], the effect of heat generation and dissipation which is present in the actual device is neglected, in other words, a constant temperature over entire device (isothermal condition) is assumed in these models. In [6], there is a brief report on a supercomputer based 2-D laser simulator which also took into account the energy transfer equation based on the finite difference method.

In this paper, we present a fully self-consistent non-isothermal, 2-D finite element model of laser diodes. The model includes the self-consistent solution of electrical, optical, and thermal conduction equations.

The electric part of the model consists of the solution of the Poisson and the electron and hole Continuity equations. The model also incorporates the complications of Fermi–Dirac statistics, position-dependent band structure and incomplete ionization.

The optical part of the model consists of the solution of the wave equation and the photon rate equation. The 2-D wave equation is solved using the finite element method making it consistent with the method used to solve the device equations [7].

The thermal part of the model consists of solving the thermal conduction equation. It incorporates various lattice-temperature dependent physical parameters, in particular, it includes temperature dependent band structure and carrier statistics.

The model has been implemented as a general purpose device simulator named FELES-1 (finite element light-emitter simulator, version 1.0). The simulations are carried out on DEC3100 workstations.

The laser model outlined above is presented in Section II. A brief discussion of the finite element solution technique is given in Section III and in Section IV, the model is applied to an AlGaAs–GaAs ridge waveguide quantum-well (QW) laser diode. A comparison between the isothermal and nonisothermal results of the model are given. It is found that even under ideal thermal exchange boundary conditions, the internal heating and local variation of temperature is significant enough to affect device performance. It is also found that the peak of the temperature profile does not necessarily locate at the center of the active region and is affected primarily by the local resistivities and current density as well as boundary conditions.

The effects of different thermal exchange boundary conditions on the performance of a laser device are examined. The results demonstrate that the thermal exchange rates at the boundaries could significantly alter the performance of laser diodes. It was also found that improving the thermal exchange condition through incorporating different semiconductor substrates such as silicon can improve the laser performance.

II. FORMULATION

In this section, we will first briefly describe the conventional semiconductor device equations which have been incorporated in various 2-D models of electronic devices including our own general-purpose semiconductor device simulator (GPSDA-1) [7]. The additions made to include the wave and photon rate equations in order to model photon generation and propagation will be discussed next. And finally, the thermal model used will be presented with a discussion of the thermal conduction equation.

A. Device Equations

The device equations which govern the electrical behavior are the following:

1) *The Poisson equation:*

$$\nabla \cdot \epsilon \nabla \psi = -q(p - n + N_D^+ - N_A^-) \quad (1)$$

where N_D^+ , N_A^- are ionized donor and acceptor concentration with

$$N_D^+ = N_D \frac{N_c/g_c \exp(-\Delta E_D/k_B T)}{N_c/g_c \exp(-\Delta E_D/k_B T) + n} \quad (2)$$

$$N_A^- = N_A \frac{N_v/g_v \exp(-\Delta E_A/k_B T)}{N_v/g_v \exp(-\Delta E_A/k_B T) + p} \quad (3)$$

2) *The Continuity equations:*

$$\frac{\partial n}{\partial t} = \frac{1}{q} \nabla \cdot J_n + (G - R) \quad (4)$$

$$\frac{\partial p}{\partial t} = -\frac{1}{q} \nabla \cdot J_p + (G - R). \quad (5)$$

The current density is composed of a drift component and a diffusion component of current. Under nonisothermal conditions, an additional thermal diffusion current arises from temperature gradients [8], [9]. The current densities are then expressed by

$$\begin{aligned} J_n &= -q\mu_n \nabla \psi + qD_n \nabla n + qD_n^T \nabla T \\ &= -q\mu_n n \nabla \phi_n + \alpha_n q\mu_n n V_T \frac{\nabla T}{T} \end{aligned} \quad (6)$$

$$\begin{aligned} J_p &= -q\mu_p \nabla \psi - qD_p \nabla p - qD_p^T \nabla T \\ &= -q\mu_p p \nabla \phi_p - \alpha_p q\mu_p p V_T \frac{\nabla T}{T} \end{aligned} \quad (7)$$

where D_n^T and D_p^T are electron and hole thermal diffusion coefficients, respectively. If we assume that the temperature dependence of mobility is given by $\mu = CT^{\alpha-1}$ [9], then $D^T = \alpha\mu k_B/q$. The carrier concentrations, as a func-

tion of the quasi-Fermi levels, are given by:

$$n = n_{ir} e^{(\Psi - \phi_n + V_n)/V_T} \quad (8)$$

$$p = n_{ir} e^{(\Psi - \phi_p + V_p)/V_T} \quad (9)$$

where V_n and V_p are the band parameters, which take into account the variation of band edge with doping and composition, the shape of energy band, and the Fermi statistics [10]. In FELES-1, a position dependent lattice temperature $T(x, y)$ is used in the expressions above rather than a constant temperature as used in previous models.

The dependence of the carrier mobility μ on the total doping concentration N_T [11], the composition x [3] and the electrical field F is modeled as:

$$\mu(N_T, T) = \left[\frac{\mu_{\min}(N_T - 10^6 N_{\text{ref}})}{N_T + 10^6 N_{\text{ref}}} + \frac{\mu_{\min} + \mu_{\max}}{[1 + N_T/N_{\text{ref}}]^\alpha} \right] \frac{300}{T} f(x) \quad (10)$$

where the composition dependent $f(x)$ is the same as that in [3]. The electron mobility's dependence on electric field is expressed by

$$\mu_n(N_T, T, F) = \frac{\mu_n(N_T, T) + V_{\text{nsat}}/F(F/F_0)^4}{1 + (F/F_0)^4}. \quad (11)$$

This model takes into account both the negative differential mobility and velocity saturation in the high-field region, unlike the previous models [3]–[5], where the mobilities were computed only as a function of the doping densities and the compositions of the materials, but without taking into account high-field velocity saturation which leads to an overestimation of current and differences in carriers and field distributions. The hole mobility is modeled as:

$$\mu_p(N_T, T, F) = \frac{\mu_p(N_T, T)}{1 + \mu_p(N_T, T)/V_{\text{psat}}} \quad (12)$$

where V_{nsat} , and V_{psat} are the electron and hole saturation velocities, respectively. As in most previous models, the values of V_{nsat} and V_{psat} are taken to be the same, though this could easily be modified, should different values be well established in the future. The GaAs–AlGaAs material constants used in this paper are taken from well accepted data in literature and are summarized in Table I.

The recombination model used is of the form,

$$R = R_{\text{SRH}} + R_{\text{aug}} + G_{\text{av}} + R_{\text{sp}} + R_{\text{st}} \quad (13)$$

which includes five important recombination mechanisms: spontaneous emission recombination (R_{sp}) and stimulated emission recombination (R_{st}), which couples the electrical and optical equations, in addition to the conventional bulk recombination (R_{SRH}), auger recombination (R_{aug}), and impact ionization generation (G_{av}).

The recombination rate due to stimulated emission is given by:

$$R_{\text{st}} = \frac{c_0}{n_{\text{eff}}} g(x, y) \Sigma_p I_p(x, y) S_p \quad (14)$$

TABLE I
PARAMETERS FOR GaAs–Al_xGa_{1-x}As USED FOR THE SIMULATION

Parameter	Value
Bandgap at Γ point (eV)	$E_{g\Gamma} = 1.4218 + 1.247 * x \quad x < 0.45$ $= 1.4218 + 1.247 * x$ $+ 1.147 * (x - 0.45)^2 \quad x > 0.45$
Bandgap at L point (eV)	$E_{gL} = 1.708 + 0.642 * x$
Bandgap at X point (eV)	$E_{gX} = 1.90 + 0.125 * x + 0.143 * x^2$
Electron affinity (eV)	$\chi = 4.06 - 0.64 * (E_{g\Gamma} - 1.4218)$
Relative dielectric constant.	$\epsilon = 13.1 - 3.0 * x$
Electron effective mass for Γ valley	$m_n^\Gamma = 0.067 + 0.083 * x$
Electron effective mass for L valley	$m_n^L = 0.55 + 0.12 * x$
Electron effective mass for X valley	$m_n^X = 0.85 + 0.07 * x$
Hole effective mass	$m_p = 0.48 + 0.31 * x$

where $g(x, y)$ is the local peak gain, n_{eff} is the effective index, S_p is the photon density for the p th optical mode, and

$$I_p = |E_p(x, y)|^2 \cdot 1 / \iint |E_p(x, y)|^2 dx dy \quad (15)$$

which is the normalized dimensionless optical intensity distribution (to be scaled by a unit area) for the p th optical mode obtained from the solution of the wave equation.

The spontaneous emission recombination rate is modeled as:

$$R_{\text{sp}} = B(n^\Gamma p - n_0^\Gamma p_0) \quad (16)$$

where n^Γ , which is the electron density in the Γ valley, and n_0^Γ , which represents the electron density in equilibrium, are

$$n^\Gamma = \frac{N_c^\Gamma}{N_c^\Gamma + N_c^L + N_c^X} n = \gamma n, \quad n_0^\Gamma = \gamma n_0$$

B is the radiative constant, which depends on carrier density and temperature. We use the following approximation in this paper for GaAs–AlGaAs [3],

$$B = 1.52 \times 10^{-10} - 1.2 \times 10^{-29} \min(n^\Gamma, p) (\text{cm}^3/\text{s}). \quad (17)$$

The remaining recombination processes are modeled conventionally and can be found in [11].

B. Optical Equations

1) *Wave equation:* The optical field $E(x, y)$ in the device satisfies the following Helmholtz wave equation:

$$\nabla^2 E(x, y) + (k_0^2 \bar{n}^2 - \beta^2) E(x, y) = 0 \quad (18)$$

where $E(x, y)$ is the electrical field vector of a wave propagating in the z direction. The orientation of the axes has the x direction parallel and the y direction perpendicular to the heterojunction. In the second term, k_0 is the free space wavenumber, $(2\pi/\lambda_0)$, β is the mode propagation

constant, and \bar{n} is the complex refractive index given by:

$$\bar{n}^2(x, y) = \epsilon(x, y) = \bar{n}_0^2 + (-\alpha_R + j) \frac{\bar{n}_0 g(x, y)}{k_0} - j \frac{\bar{n}_0}{k_0} (c_n n + c_p p) \quad (19)$$

where $j = \sqrt{-1}$, \bar{n}_0 is the bulk refractive index, α_R is the linewidth broadening factor, and $g(x, y)$ is the local peak gain which is calculated on the basis of quantum well theory, which considers the k -selection rules, a reduced density of states, and intraband relaxation [12]. In this model, a linear gain approximation is used,

$$g(x, y) = a(N - n_e); \quad N = \min(n^\Gamma, p) \quad (20)$$

where the parameters a and n_e are the fitted local gain coefficient and the transparency carrier concentration, respectively. For GaAs, we adopt $a = 3.32 \times 10^{-16} \text{ cm}^2$, $n_e = 1.889 \times 10^{18} \text{ cm}^{-3}$. As well, c_n, c_p are the free carrier absorption coefficients, and $c_n = 3 \times 10^{-18} \text{ cm}^2$, $c_p = 7.0 \times 10^{-18} \text{ cm}^2$ for GaAs. The effect of the electrical equations on optical waves in the model are reflected in $\bar{n}^2 - \bar{n}_0^2$. It is noted that a more complicated but more accurate gain model than (20), such as a polynomial function of N , could be as easily adopted should it be necessary for some specific purpose.

Equation (18) is the full vector wave equation. For our purpose of modeling waveguide lasers, the problem can be simplified to solving for just the quasi-TE mode, i.e., $E = (E_x, 0, E_z)$. The quasi-TE mode is an eigensolution of the scalar Helmholtz equation for a horizontal component E_x which is continuous across horizontal interfaces, but discontinuous across vertical interfaces [13]. In order to avoid the complications of interface conditions, most authors [1]–[6] only solve the scalar Helmholtz equation for the E_z component. This component is parallel to all internal dielectric interfaces and is therefore continuous, with a continuous first derivatives throughout the solution domain. In this paper, we have also taken the scalar Helmholtz equation approach which is a good approximation for quasi-TE mode [13]. The incorporation of the interface conditions for the E_x component of the quasi-TE mode into the finite element discretization is numerically cumbersome and is to be completed at a later stage.

The boundary condition for the outer boundary of the solution domain of the wave equation is

$$E(x, y)|_{\partial\Omega} = 0. \quad (21)$$

The regions beside the laser ridge is typically filled with dielectric and is chosen in this work to be filled with SiO₂ (see Fig. 1). The outer boundary $\partial\Omega$ includes the metal contacts and the dielectric or semiconductor surfaces. When symmetric devices are modeled, only half of the device needs to be simulated. The boundary condition along the plane of symmetry ($\partial\Omega_s$), is $\partial E(x, y) / \partial \bar{n} |_{\partial\Omega_s} = 0$.

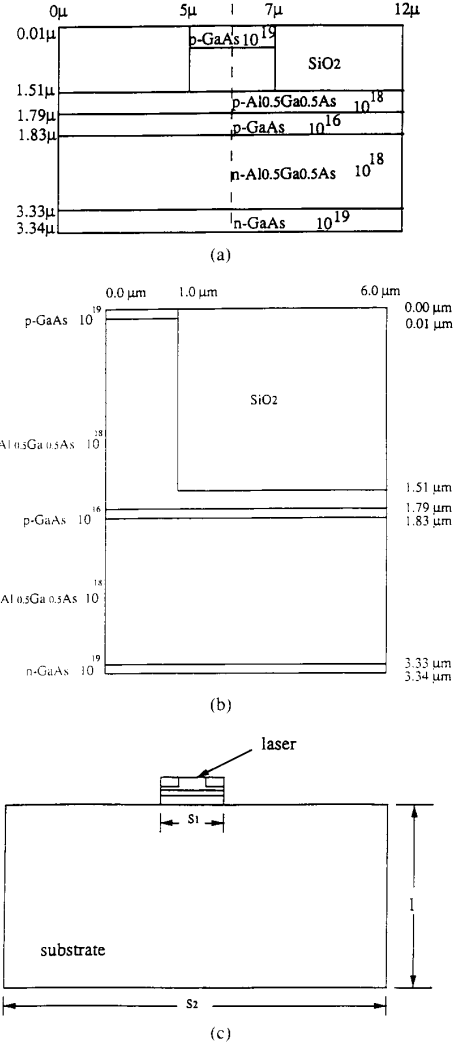


Fig. 1. (a) Structure of a GaAs-AlGaAs ridge laser diode. (b) The simulated device structure. (c) The device structure including the substrate.

sion, is,

$$G_p = \alpha_T \quad (22)$$

where G_p is the p th mode gain, and α_T is the total loss. The modal gain is calculated by averaging over the local gain with the optical intensity,

$$G_p = \frac{\iint g(x, y) |E_p(x, y)|^2 dx dy}{\iint |E_p(x, y)|^2 dx dy} \quad (23)$$

The total loss includes the bulk loss α_a , the free carrier absorption loss α_{fc} and the mirror loss α_M

$$\alpha_T = \alpha_a + \alpha_{fc} + \alpha_M \quad (24)$$

lated from:

$$\alpha_a = \iint \alpha'_a |E_p|^2 dx dy / \iint |E_p|^2 dx dy. \quad (25)$$

$$\alpha_{fc} = \iint \alpha'_{fc} |E_p|^2 dx dy / \iint |E_p|^2 dx dy. \quad (26)$$

The mirror loss, α_M is

$$\alpha_M = \frac{1}{2L_c} \ln \frac{1}{R_f R_r} \quad (27)$$

where L_c is cavity length, R_f and R_r are the front and rear facet reflectivities, respectively. The photon rate equation couples the optical equation and the electrical equations through the photon density. When the device is under low bias, the gain G_p is less than the loss α_T giving a photon density of zero. When the bias is increased to some threshold value, where G_p becomes somewhat larger than α_T , the photon density S_p increases. The increase of the photon density leads to increasing stimulated emission recombination which causes the decrease of the carrier concentration that maintains the equality $G_p = \alpha_T$, through a negative feedback procedure. This determines the steady-state photon density and the solution for the device.

To complete the laser model, it is necessary to calculate the output power of mode p , using the following expression [3]:

$$\begin{aligned} P &= hv\eta_D \int \frac{S_p}{\tau_{ph}} dV \\ &= \frac{hc_0^2 S_p}{2\lambda_0 n_{\text{eff}}} \frac{\ln(1/R_f R_r)}{1 + \sqrt{R_f/R_r} (1 - R_r)/(1 - R_f)} \\ &\quad \cdot \iint |\vec{E}|^2 dx dy \end{aligned} \quad (28)$$

C. Thermal Condition Equation

The thermal conduction equation is given by:

$$\nabla \cdot S + \rho_m c_p \frac{\partial T}{\partial t} = F \cdot J + E_g R_{nr} \quad (29)$$

where ρ_m is the mass density, c_p is the specific heat capacity under constant pressure, F is the static electric field, J is current density, R_{nr} is the nonradiative recombination rate which includes only $R_{\text{SHR}} + R_{\text{aug}}$, and S is the heat flux vector given by:

$$S = -\kappa \nabla T \quad (30)$$

where κ is the thermal conductivity of materials and is temperature dependent. For GaAs, the thermal conductivity, $\kappa = 0.46(T_0/T)^{4/3}$ W/cm-k [12]; for Si, $\kappa = 1.5(T_0/T)^{4/3}$ W/cm-k, ($T_0 = 300$ K) thus Si has a thermal conductivity that is about a factor of 3 of that of GaAs. The RHS of (29) represents the heat sources. The sources consist of Joule heat and recombination heat.

The three types of thermal conduction boundary conditions are the following:

i) The isothermal boundary condition

$$T|_{\partial\Omega} = T_0 \quad (31)$$

where T_0 is the ambient temperature. This condition fixes the temperature at the device boundary while heat is transferred. It corresponds to an ideal heat sink contact with an infinite heat conductivity. Simulations obtained under this condition represent the upper limit in device performance.

ii) The adiabatic boundary condition,

$$\left. \frac{\partial T}{\partial \vec{n}} \right|_{\partial\Omega} = 0 \quad (32)$$

where $\partial T / \partial \vec{n}|_{\partial\Omega}$ is the normal derivative of T at the boundary surface. This boundary is ideally insulated with a zero thermal conductivity. The device performance obtained from a simulation under this condition represents the worst heat dissipation case.

iii) The thermal exchange boundary condition,

$$-\kappa \left. \frac{\partial T}{\partial \vec{n}} \right|_{\partial\Omega} = \alpha(T - T_0). \quad (33)$$

Under this condition, the thermal flux through the boundary surface is proportional to the difference between temperature at the boundary surface and the ambient temperature. The constant α is the heat exchange coefficient. This mixed boundary condition depends on the temperature and its derivative, representing a more realistic situation in actual devices. Generally speaking, the temperatures at various points on the boundary surfaces are not the same and may change after solving the thermal conduction equation. Moreover, if the boundary happens to be an ohmic contact boundary, its electrical boundary conditions change with the temperature on boundary points as well. This is due to the fact that on the ohmic contact boundary the potential is

$$\begin{aligned} \Psi &= \phi_n + \frac{V_p - V_n}{2} + V_T \ln \left[\frac{N_s}{2} e^{-(V_p - V_n)/2V_T} \right. \\ &\quad \left. + \sqrt{1 + \left(\frac{N_s}{2} e^{-(V_p - V_n)/2V_T} \right)^2} \right] \end{aligned} \quad (34)$$

and the Slotboom variables $\Phi_n = \exp(-\phi_n/V_T)$, $\Phi_p = \exp(\phi_p/V_T)$ depend on the lattice temperature. This makes the convergence of a decoupled solution very difficult.

III. COMPUTATION TECHNIQUE

In this section, the schematic flowchart for the simulator is presented and the solution procedure of the electrical equations, the thermal conduction equation, and the optical equation is discussed. All the equations in the

model are discretized in a uniform way using the finite element method. The six equations are solved in a decoupled manner to save on memory requirements. Each system of discretized nonlinear algebraic equation (except the optical equation) is solved using a damped Newton iteration method [14].

A. Flowchart

The flowchart of FELES-1 simulator is shown in Fig. 2. After the data for device structure (including dimensions, composition, and doping concentration) are read, the mesh is generated and refined automatically [16]. The simulator first solves Poisson's equation at zero bias and obtains electron and hole distributions. These distributions will influence the complex refractive index in (19). Then the wave equation is solved in order to obtain the mode distribution. Initially, the photon density is set to zero, then the bias voltage is increased and the gain-loss balance is checked for photon generation. At each bias step, the electrical equations (first continuity equations followed by Poisson's equation) and thermal conduction equation are solved in two nested loops:

- To begin, the electrical equations are solved with a lattice temperature distribution which is uniform. In later iterations, the lattice temperature is obtained from the solution of the thermal conduction equation. The simulator then solves the continuity equations and Poisson's equation, in turn, each using the Newton iteration. These two iterations are repeated until convergence is achieved. We call this the adiabatic approximate solution.
- After convergence is achieved for the electrical equations, the thermal conduction equation is solved, with heat sources which arise from the electric field and current, and recombination distributions which were obtained in a).

After this, the temperature distribution will have changed and therefore the electrical equations must be resolved since most of the parameters in the electrical equations are temperature dependent. The Fermi-Dirac statistical distributions of charge carriers are strongly dependent on the lattice temperature. The iterations between a) and b) are repeated until self-consistent convergence results. We call this the nonisothermal solution.

Next, the optical gain and loss in the device can be calculated, and their equality tested. If $G_p \neq \alpha_T$, the photon density is adjusted accordingly using a linear extrapolation with $\Delta S/\Delta G$. This factor is a negative quantity, and is obtained from the two most recent values of S_p and G_p . Initially, it is assumed that $\Delta S/\Delta G = c$, and a small value is chosen for c , e.g., $c = 0.01$. It should be noted that the adjustment of photon density does not start until the first occurrence of $G_p > \text{loss}$.

At this point, the nonisothermal solution of the device has to be recalculated and the procedure repeated until equality is reached. Then, the wave equation is solved

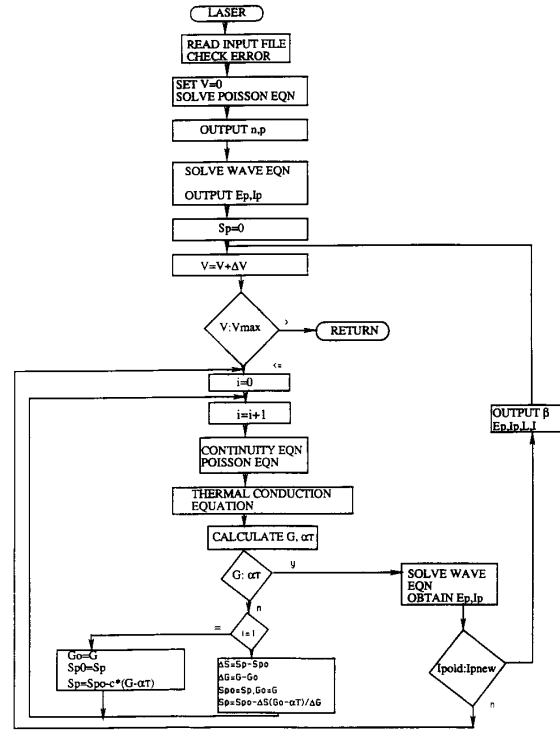


Fig. 2. The flowchart of laser simulation.

again because of changes in the complex refractive index. The last two optical intensity distributions are compared. If they are not identical, the solution is recalculated until self-consistent convergence is reached. From this procedure the solution for the device at one bias point is obtained.

The procedure would be very difficult to carry out without a very accurate discretization and robust nonlinear iteration strategy. For example, [4] has shown that using Gummel's method to solve only the Poisson, continuity and wave equations, without the thermal conduction equation, may fail to converge at high injection if a standard five-point finite difference scheme is used. This condition occurs near the onset of lasing. To overcome this problem, the author in [4] used a type of continuum approximation that modifies the doping and composition in certain layers such that current flow in the device becomes nearly one-dimensional. After reaching the lasing voltage, the doping and composition are changed back to their original values.

For this paper, the procedure is more difficult because the carrier concentrations depend dramatically on the lattice temperature (8)–(9) and the electrical boundary conditions change whenever the thermal conduction equation is solved.

We used our FEM-QSG discretization method [7], [16] and ILUV-CGS nonlinear iteration algorithm [17] to solve the nonstable problem with the coupling thermal conduction equation.

B. Analysis of the Electrical and the Thermal Conduction Equations

The semiconductor device equations are convection dominated flow equations, and each of the equations (1), (4), (5), and (29) can be expressed in the following form:

$$\nabla \cdot (a(x, u, \nabla u) \nabla u) = f(x, u, \nabla u, t) \quad x \in \Omega \quad (35.1)$$

$$u|_{\Omega_D} = d \quad (35.2)$$

$$\left. \frac{\partial u}{\partial n} \right|_{\Omega_N} = q \quad (35.3)$$

where Ω is the domain of the device and Ω_D and Ω_N are forced and natural boundaries of the device, respectively.

It is well known that the convection-diffusion equation presents serious numerical difficulties when “ a ” depends strongly on “ u ,” (especially exponentially). These difficulties stem from the combination of the essentially elliptic and parabolic nature of the two terms $a \nabla^2 u$ and $\nabla a \cdot \nabla u$. If we use the central difference scheme to discretize the differential operators, the solution would oscillate when the mesh size exceeds a critical value. The classical finite element is equivalent to the central difference and thus encounters the same difficulties.

In the finite difference (FD) approximation these difficulties are overcome by the upwind difference method. For finite element analysis, various upwind schemes have been proposed to avoid the spurious oscillation in solving the convective dominated flow equations. In the case of semiconductor device equations, the Scharfetter–Gummel scheme [18] is the most widely used method in the discretization of the continuity equations in a finite difference scheme, referred to as the box integral method (BIM-SG). Essentially, the optimal upwind formulation for the one-dimensional case is equivalent to the S - G scheme. However, the flow direction is not always along the grid line, and it is difficult to extend the formulation to two or three dimensions in a straightforward manner [19]. In order to extend the S - G scheme to multidimensions in a finite element method, we have the S - G scheme embedded in the quadrature of a finite element assembly of the equations (1), (4), (5), (29). The method, referred to as FEM-QSG, not only avoids the problem of flux direction in the discretization of the flow terms, but also extends the S - G scheme to the discretization of the source terms in a consistent way. The detail of the FEM-QSG is given in Appendix I.

This method is more accurate than the conventional finite difference one, thus, a coarser mesh could be used to achieve the same accuracy. Moreover, the convergence for the Newton iteration of each single equation is sped up. Specifically, for Poisson’s equation and the thermal conduction equation, which have very strong non-linear dependencies, only a few iterations are sufficient for convergence. This method has a second power convergence in contrast with the linear convergence for the conventional finite element method [17], [16].

The discretized nonlinear algebraic equations are solved

using the nonlinear ILUV-CGS method [17] which is presently one of the fastest methods. The system of equations for the whole device is solved in a decoupled manner in two main loops: a) adiabatic approximation solution, and b) non-isothermal solution (see Fig. 2)

C. Analysis of Wave Equation

Standard laser devices, such as ridge waveguide lasers, have structures which are essentially 2-D. For accurate modeling, the wave equation (18) can not be considered separable in the x and y directions. So schemes that only use the effective index method yield results that are not accurate enough. In this paper, the finite element method is used to provide a 2-D analysis for the optical equation. This method of solution is also consistent with that used for the electronic device equations.

The normalized wave equation is

$$\nabla^2 E(x, y) + (\bar{n}^2 - \beta^2)E(x, y) = 0 \quad (36)$$

and its finite element discretized form is

$$\begin{aligned} & - \int \nabla E \cdot \nabla \phi_i \, d\Omega + \int \phi_i \bar{n}^2 \, d\Omega - \beta^2 \int \phi_i E \, d\Omega \\ & = - \int (\phi_i \nabla E) \cdot d\vec{S} \end{aligned} \quad (37)$$

where ϕ_i is the shape function. According to the boundary conditions of the wave equation, the integral in the RHS can be evaluated as:

$$\int \phi_i \nabla E \cdot d\vec{S} \begin{cases} = 0 & i \in \Omega^0 \\ \neq 0 & i \in \Omega_D \\ = \int \phi_i \frac{\partial E}{\partial n} d\vec{S} & i \in \Omega_N \end{cases} \quad (38)$$

where Ω^0 is the inner part of the domain Ω , Ω_D , and Ω_N are the forced and natural boundary of the domain. Generally speaking, the outline of a device is a forced boundary whose boundary value is zero for eigenvalue problem, or nonzero for scattering problem. The exception is for symmetric surfaces which has a natural boundary for even modes. We are only concerned with equations for inner points and natural boundary points, thus, the integral is zero in our case. If we use the shape function as interpolation function, then have

$$\sum_j K_{ij} E_j = \lambda \sum_j B_{ij} E_j \quad (39)$$

or in matrix form,

$$KE = \lambda BE \quad (40)$$

where $\lambda = \beta^2$, and the elements of the matrix K are

$$\begin{aligned} K_{ij} &= \sum_e [K_{ij}^e] = \sum_e - \int_e \nabla \phi_i \nabla \phi_j \, d\Omega + \int_e \bar{n}^2 \phi_i \phi_j \, d\Omega \\ &= \sum_e \left[- \frac{B_i^e B_j^e + C_i^e C_j^e}{4\Delta_e} \right. \\ & \quad \left. + 2\Delta_e \sum_{k \in e} (\bar{n}^2)_k (\phi_i)_k (\phi_j)_k W_k \right] \end{aligned} \quad (41)$$

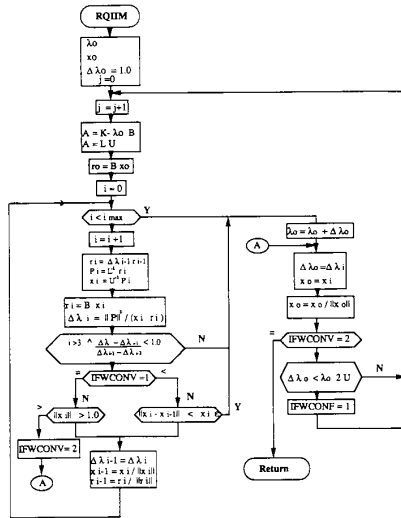


Fig. 3. The flowchart of RQIIM.

and the elements of the matrix B are

$$\begin{aligned} B_{ij} &= \sum_e [B_{ij}^e] = \sum_e \int_e \phi_i \phi_j d\Omega \\ &= \sum_e 2\Delta_e \sum_{k \in e} (\phi_i)_k (\phi_j)_k W_k \end{aligned} \quad (42)$$

where linear triangular elements are used with a seven-point Gaussian quadrature. The integral \int_e is the integral over an element; the sum \sum_e is the sum of all elements and $\sum_{k \in e}$ is the sum over the seven points basis on an element and $\nabla \phi_i = (B_i^e / 2\Delta_e, C_i^e / 2\Delta_e)$ is the gradient vector.

The complex refractive index is a function of the carrier concentrations, thus we have integrals of the form: $\int d\Omega n \dots$, $\int d\Omega p \dots$, once again. The method used to calculate these integrals has important implications for the discretization accuracy because of the exponential dependence of carrier concentrations on the potential and temperature. The generalized S - G formulas (see Appendix I) can be used to calculate these integrals by obtaining an accurate discretization of the wave equation.

The matrix equation (40) is a generalized complex eigenvalue problem for the matrix pair (K, B) , of which K is a complex symmetric matrix for the scalar wave equation, or a complex nonsymmetric matrix for the quasi-TE wave equation, and B is a real symmetric positive-definite matrix. The eigenvalue λ and the eigenvector E are complex. Fast solutions for the eigenvalue equation are important, because solving the device equations coupled with the thermal conduction equation is very computationally intensive.

It is well known that the complex spectra of (40) can be obtained by applying the QR and QZ algorithm from the EISPACK Library [20]. However, the QR and QZ schemes require storage for the "full" matrices. Memory restrictions limit the largest matrix dimension which could

be handled (for example, 800). A matrix of this size is too small to obtain finite element meshes of the resolutions required for accurate solutions of a multilayer and multi-region device such as a semiconductor laser. The suitable methods for solving large-scale complex eigenvalue problem are the Lanczos algorithm [21] and Rayleigh quotient-inverse iteration method [22]. The Lanczos algorithm can be used to obtain the full spectrum of eigenvalues but it is very time consuming. The latter method is much faster, but can only obtain a single eigenvalue at a time and requires an initial approximation. Since we are only interested in a few dominate eigenvalues in a laser simulation, the inverse iteration method was preferred.

A very efficient Rayleigh quotient-inverse iteration method (RQIIM) was developed for solving the generalized eigenvalue problem of finite element discretized equation for the wave equation. Our RQIIM algorithm (see Fig. 3) is given in Appendix II.

IV. SIMULATION RESULTS

The GaAs-AlGaAs ridge laser diode whose structure is shown in Fig. 1(a), was simulated on our DEC3100 workstation. The device has a ridge width of $2.0 \mu\text{m}$, active layer thickness of 400 \AA , cladding layers of $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$ with a doping 10^{18} cm^{-3} , cavity length of $300.0 \mu\text{m}$, and a front and rear facet refractivity of 0.3. Due to the symmetry of the device, only half of the device needs to be simulated to save computation time. Also for the sake of saving computation time, the relatively large substrate has been truncated and replaced with an ohmic contact [Fig. 1(b)]. We found the truncation has almost no influence on the solution of the wave equation in this example. Including more of the substrate slightly increases the total voltage drop at the same current level and the overall thermal resistance, but it has almost no influence to characteristics of optical power versus current and the distribution of field and current. The regions beside the ridge are filled with SiO_2 which leads to the boundary condition for the wave equation, $\vec{E} = 0$ along the whole outer boundary of the device, and $\partial \vec{E} / \partial \vec{n} = 0$ is assumed along plane of symmetry of the device. The resulting traverse fundamental mode is plotted in Fig. 4. On a DEC3100 workstation, the cpu-times needed for each bias point were approximately 1 h (low bias), 3 h (lasing threshold), and 5 h (above lasing threshold) for this example.

Several studies were conducted for this device. First the simulated L - I curve of the device without considering thermal conduction effects was compared (shown in Fig. 5) with the measured L - I curve of an actual device with a contact temperature fixed at 30°C . Even though fitting parameters were not used, the agreement between the simulation result and experiment was rather good. But the calculated threshold current I_{th} is less and the external quantum efficiency η_{ext} is greater than those measured. The difference can be attributed to the fact that the simulation assumes isothermal condition across the entire device which is, as shown below, an even better (but fictitious) condition than the ideal heat sink boundary condition.

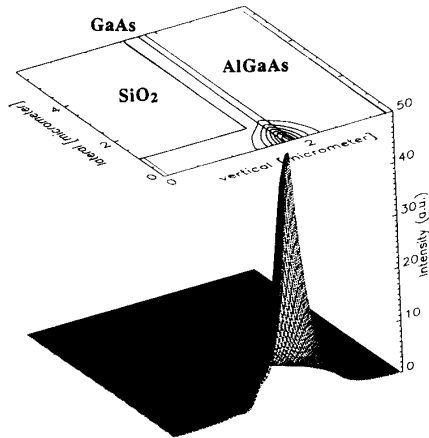


Fig. 4. The optical mode distribution of the laser diode.

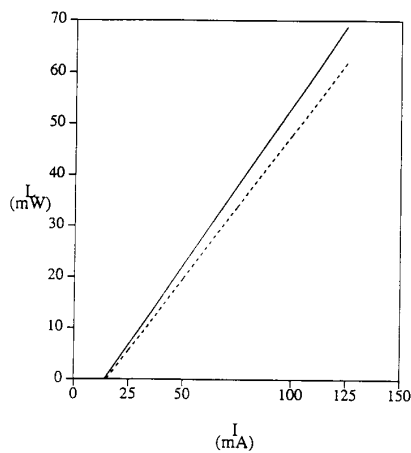


Fig. 5. The comparison of the simulated $L-I$ curve (solid) with measured $L-I$ curve (dashed).

Second, thermal effects were examined with a comparison between the isothermal model (without considering thermal conduction and heat generation) and the nonisothermal model. For the latter, we first assume isothermal boundary conditions along the contact boundaries at the top and bottom of the simulated structure. Under this assumption, heat is readily transferred to the metal contact due to the "infinite" heat conductivity assumed in the isothermal boundary condition. It is observed from Fig. 6 that the current in the nonisothermal case is larger than the one in the isothermal case at the same bias. Furthermore, the nonisothermal model has a larger threshold current and smaller quantum efficiency, both in support of the explanation of Fig. 5. However, the differences are relatively small due to the small temperature increase (the temperature increase in the device is less than 1 K at lasing bias). This is understandable since the isothermal boundary condition implies an infinitely fast exchange rate at the contacts (ideal heat sink), thus the dissipation of the internally generated heat is limited only by the finite ther-

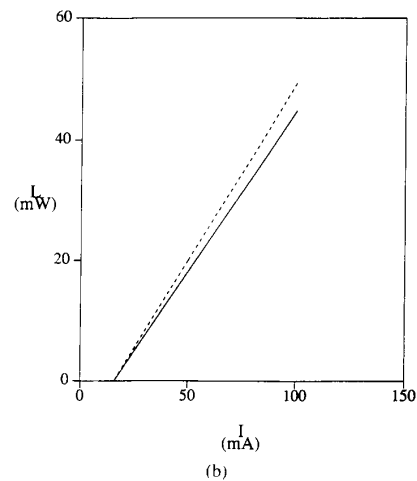
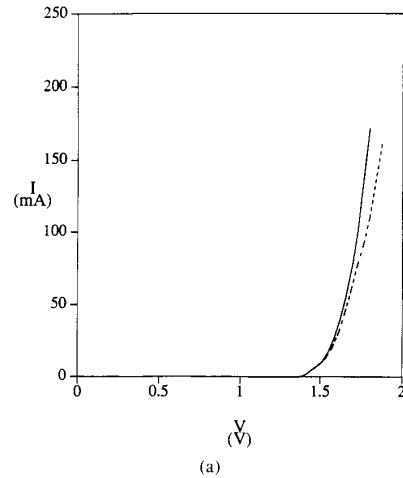


Fig. 6. The comparison of simulated results using isothermal model and nonisothermal model with isothermal boundary condition: (a) IV curve (solid—nonisothermal model, dashed—isothermal model); (b) LI curve (solid—nonisothermal model, dashed—isothermal model).

mal conductivity of the semiconductor layers which are quite thin in the modeled structure. The plot in Fig. 7 shows the distribution of the temperature in the structure while lasing at a bias of $V = 1.6$ V as an example. The temperature profiles at higher biases are similar to Fig. 7, but with increased peak temperatures. The active layer under the ridge is shown to have higher temperatures. This is consistent with the fact that the injection current in the laser diode is concentrated in this region. However, less than obvious, is that the point having maximum temperature is above the active layer, where a poor thermal conduction condition co-exists with the high current density.

Actual devices do not have the idealized thermal exchange conditions. In order to model the situation more realistically, the boundary conditions for the thermal conduction equation are modified as follows:

A mixed boundary condition [(33)], was used. Again, to save computation time, the substrate is simplified to a pure heat resistor, by ignoring the heat generated by the

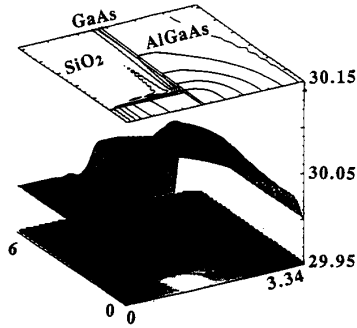


Fig. 7. The distribution of temperature in the laser diode using isothermperature boundary condition at contract boundary.

current passing through the substrate which is reasonable since the substrate is highly conductive electrically. We could have done this differently by simulating the substrate, which is a few hundred times larger than the active region, together as a part of the device, but we could not afford the exceedingly long cpu time (\approx a month) on a workstation. Now we estimate the heat exchange coefficient, [refer to Fig. 1(c)]. According to the continuity of thermal flow, the quantity of heat released through the substrate is

$$Q_{th} = J_T S_1 = \alpha(T - T_0)S_1 = \kappa \frac{T - T_0}{l} S_2 = \frac{T - T_0}{\left(\frac{l}{\kappa S_2}\right)} \quad (43)$$

where J_T is heat flux, l is thickness of the substrate, and S_1 and S_2 are the width of active layer and the substrate, respectively. For GaAs substrates, $\kappa = 0.46$ W/cmK, giving $\alpha = 352.82$ W/cm² K. At the top contact, the heat exchange coefficient which depends on the device package, was arbitrarily assumed to be compatible with that of the substrate, i.e., $\alpha = 300$ W/cm² K. The ambient temperature T_0 is set as 30.0°C.

Next, we consider the use of silicon substrate in response to the interest in monolithic integration of (Al,Ga)As photonic devices on silicon substrates which has been pursued in order to incorporate optical capability with advanced silicon high-speed electronic device technology. For example, in [23] it was reported that the fabrication of strain-free (Al,Ga)As heterostructure lasers on silicon substrate was successful. Because silicon has higher thermal conductivity than that of GaAs, one expect AlGaAs-GaAs heterostructure lasers on silicon to have better performance. The effects of thermal conductivity of the substrate on the performance of a laser diode are examined by simulating two device having the same structure but different substrates. One laser has a GaAs substrate and the other has a Si one which has a thermal conductivity about three times that of GaAs. The plot in

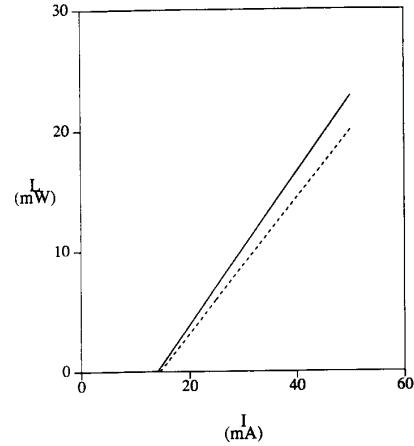


Fig. 8. The $L-I$ curves: solid for Si substrate; dashed for GaAs substrate using thermal exchange boundary condition at contract boundary.

Fig. 8 shows $L-I$ curves of the two devices. It can be seen that the laser on Si has a slightly smaller threshold current while a higher quantum efficiency. From this point of view, the push for GaAs-on-Silicon is well warranted. In general, it shows that improving the thermal conductivity of the substrate improves laser diode performance. Fig. 9 shows the maximum temperature increase ΔT_m in the laser diodes as a function of the bias voltage. It is clear that at below the threshold bias voltage, the temperature increase is small which explains why the improvement of threshold current for a laser on Si is small. However, at voltages higher than the threshold voltage, the temperature increases in the devices on GaAs are substantial. The maximum temperature increase ΔT_m for Si substrate seems to increase almost linearly as the bias increases, while ΔT_m for GaAs substrate on the other hand seems to increase exponentially. This explains the improvement of quantum efficient for a laser on Si. Fig. 10(a) and (b) show 2-D distributions of temperature for the GaAs and Si substrate cases at a bias of $V = 1.6$ V. From Fig. 10, the temperature in the center of the device is higher than that at the edge of the device along the active layer direction (y direction) which is consistent with the distribution of the current density and electron concentration [see Fig. 12(b)]. Because, in the center of the device, there are higher current and carrier densities, thus resulting in higher Joule heating and recombination heating. However, the temperature variation along the ridge direction (x direction) is smaller. To see more clearly the temperature distribution in the device, we plot the two cross section figures [Fig. 11(a) and (b)] along the ridge and along active layer for the laser on the GaAs substrate. From Fig. 11(a), it is observed that the temperature along the ridge is not uniform, and the temperature on top of the device is slightly higher than the one on the bottom of the device. There are significant gradients of temperature near two ohmic contacts caused by the thermal exchange boundary condition. The distribution of temperature along active

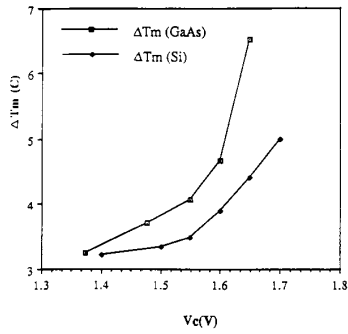


Fig. 9. The maximum temperature increase in laser diode versus the bias voltage.

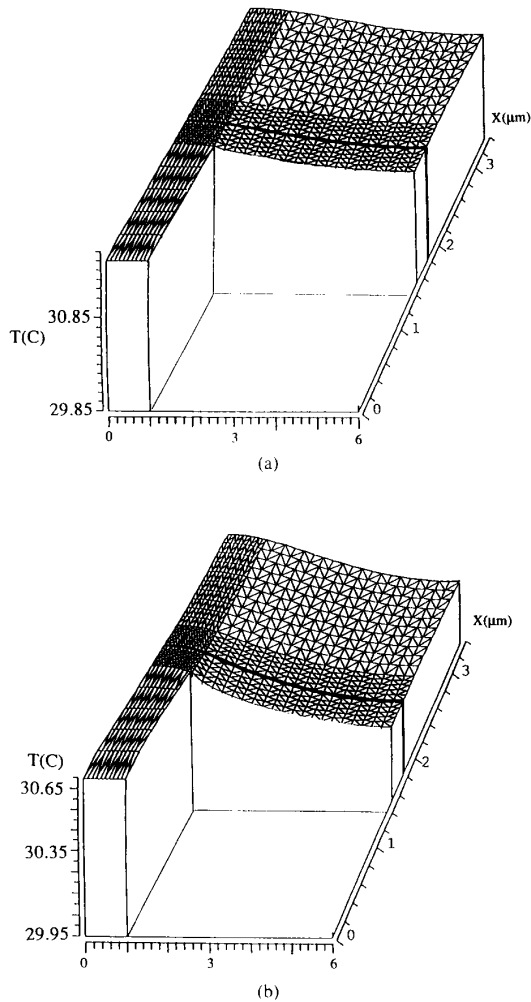


Fig. 10. 2-D distribution of temperature in the laser diode (a) GaAs substrate; (b) Si substrate at lasing bias ($V = 1.6$ V).

layer [Fig. 11(b)] is similar to the distribution of carrier concentration [see Fig. 12(b)]. This shows that the non-radiative recombination heat is the main heat source in

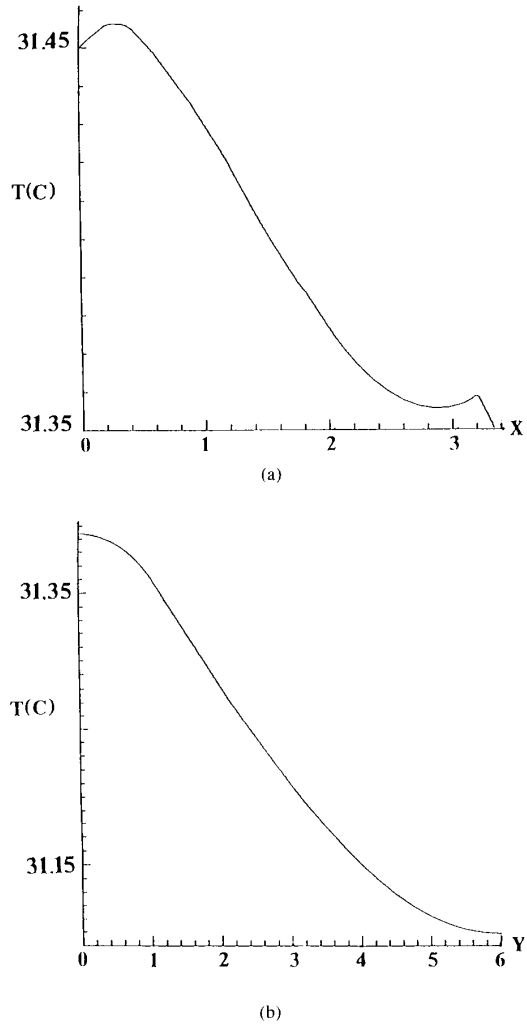


Fig. 11. The distributions of temperature along (a) $y = 0.55 \mu\text{m}$ (within the ridge); (b) $x = 1.815 \mu\text{m}$ (in active layer) for the device on GaAs substrate.

this region. Fig. 12(a) and (b) compare the distributions of the temperature and the electron concentration for lasers with GaAs and Si substrate along the active layer at a bias of $V_c = 1.6$ V, respectively. Clearly the GaAs substrate case has a higher temperature and higher electron concentration at the same bias, which cause higher leakage current and nonradiative recombination. Moreover, a higher carrier concentration and temperature lead to an increase of optical loss. Therefore, the performance of a laser diode on GaAs is inferior to that of a laser on Si. Although the laser diodes on Si substrate are better in terms of thermal properties, this aspect is minor with respect to the stress effects. However, with the progress of the processing technology to overcome the stress problems, this advantage of the thermal effects of Si substrate may become more pronounced in the future.

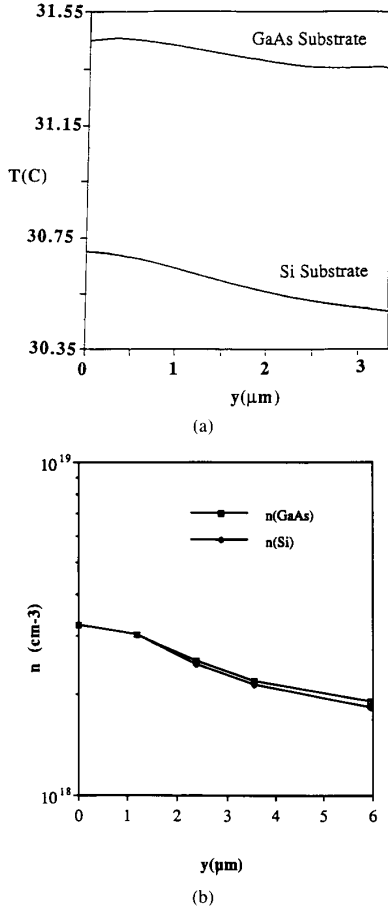


Fig. 12. The comparison of the distribution of (a) the temperature and (b) the electron concentration along the active layer ($x = 1.811 \mu\text{m}$) at lasing bias ($V = 1.6 \text{ V}$).

V. CONCLUSION AND DISCUSSION

We have presented a nonisothermal model of laser devices which solves the conventional electronic device equations (Poisson's equation, electron and hole continuity equations), thermal conduction equation, wave equation, and rate equation consistently. An accurate finite element discretization method (FEM-QSG) is used to discretize the device equations along with the thermal conduction equation as well as the wave equation. An efficient nonlinear iteration method (ILUV-CGS) is used to solve the electronic device equations and a modified Rayleigh quotient-inverse iteration method, RQIIM, is used to calculate the eigenvalue and eigenvector of the wave equation. The laser model has been implemented as a general-purpose finite element light emitter simulator FELES-1 and can operate in a workstation environment. As a demonstration of the model's capabilities, a GaAs-AlGaAs ridge laser diode was simulated by FELES-1. The results show: 1) a good agreement with the measured results without fitted parameters; 2) the differences between the non-isothermal model and the isothermal model are

considerable; 3) the thermal conductivity of the substrate plays an important role, supporting the use of Si substrate from the thermal conduction point of view. It is noted that besides being able to run on standard workstations as opposed to supercomputers, FELES can also handle novel laser devices with multiterminals and arbitrary geometries, owing to the incorporation a finite-element scheme and efficient numerical algorithm.

APPENDIX I FEM-QSG DISCRETIZATION METHOD

To discretize (35), we write the weak form as

$$\int_{\Omega} a \nabla u \cdot \nabla \phi_i ds + \int_{\Omega} f \phi_i ds = \int_{\Omega_N} a \nabla u d\Gamma \quad (\text{P1.1})$$

where ϕ_i is the shape function. This results in the formulation:

$$\sum_{j=1}^N (K_{ij} + G_{ij}) \Delta u_j = \left(\sum_{j=1}^N K_{ij} + f_i \right) \quad (\text{P1.2})$$

where the stiffness matrices K_{ij} and G_{ij} and load vector f_i are defined as

$$K_{ij} = \iint a \nabla \phi_i \cdot \nabla \phi_j ds \quad (\text{P1.3})$$

$$G_{ij} = \iint \left[\frac{\partial a}{\partial u} \nabla u \cdot \nabla \phi_i \phi_j + \frac{\partial f}{\partial u} \phi_i \phi_j \right] ds \quad (\text{P1.4})$$

$$f_i = \iint f \phi_i ds. \quad (\text{P1.5})$$

If we adopt the seven-point Gaussian quadrature rule to calculate the integrals (P1.3-5), we have the integrals of the following form:

$$\begin{aligned} & \iint_e G(n, p, e^{(\psi+V_n)}, e^{(-\psi+V_p)}) ds \\ & = 2 \Delta_e \sum_k G(n_k, p_k, (e^{(\psi+V_n)})_k, (e^{(-\psi+V_p)})_k) w_k \end{aligned} \quad (\text{P1.6})$$

where e is a triangular element, k is one of the seven points: either of the three vertices (i, j, k) of a triangular element, the middle points (l, m, n) of the three edges or the center (p) of the triangle. In FEM-QSG, we use a set of generalized S-G formulas to calculate the values of "a" at points l, m, n , and p . Similar to the derivation of the S-G formulation, we assume the derivatives $d\phi_n/dx$, $d\phi_p/dx$, $d\psi/dx$, and dT/dx are constant along the line $\bar{12}$ between points 1 and 2. From the current density equation (6), we obtain the formula for the product $n\mu$ of electron concentration and electron mobility at any point on the line $\bar{12}$ between the points 1 and 2:

$$n\mu = (n\mu)_1(1 - G_0) + (n\mu)_2 G_0 \quad (\text{P1.7})$$

where

$$G_0 = \frac{\exp [(\alpha + 1) \ln (T_1/T)] - 1}{\exp [(\alpha + 1) \ln (T_1/T_2)] - 1}. \quad (\text{P1.8})$$

Expressions (P1.7) and (P1.8) are used to calculate the values of $(n\mu)$ at points $l, m, n,$ and p in a triangle in the discretization of the continuity equations and the thermal conduction equation.

For Poisson's equation, we use the following formula to calculate the integral $\int n \dots ds$:

$$n = n_i(1 - g) + n_j g \quad (\text{P1.9})$$

$$g = \frac{1 - \exp [((\Delta\psi - \alpha_n \Delta T_n)/\bar{T}_n)(x - x_i)/h]}{1 - \exp ((\Delta\psi - \alpha_n \Delta T_n)/\bar{T}_n)} \quad (\text{P1.10})$$

where $\Delta\psi = \psi_j - \psi_i$, $\Delta T_n = T_{nj} - T_{ni}$, $\bar{T}_n = (T_{nj} + T_{ni})/2$. A similar result is obtained for p (holes). If the devices are under isothermal condition, the formulas reduce to (22)–(29) in [7].

APPENDIX II

RQIIM ALGORITHM FOR SOLVING THE GENERALIZED EIGENVALUE PROBLEM OF FINITE ELEMENT DISCRETIZED EQUATION FOR THE WAVE EQUATION

To solve the generalized eigenvalue problem,

$$Kx = \lambda Bx \quad (\text{P3.1})$$

where K and B are complex matrices, B is symmetric and positive definite, and K may be arbitrary, an initial guess eigenvalue λ_0 for (P3.1), can be improved, to obtain the eigenvalue λ where,

$$\lambda = \lambda_0 + \Delta\lambda_0 \quad |\Delta\lambda_0| \ll |\lambda_0| \quad (\text{P3.2})$$

from solving the eigenvalue equation for the λ_0 correction,

$$(K - \lambda_0 B)x = \Delta\lambda_0 Bx. \quad (\text{P3.3})$$

Using the iteration system,

$$(K - \lambda_0 B)x_i = \Delta\lambda_{i-1} Bx_{i-1} \quad (\text{P3.4})$$

the pair $(x_i, \Delta\lambda_i)$ can be calculated. The new vector x_i is computed as a solution of equation (P3.4). A random vector is used for x_0 , and the new eigenvalue correction $\Delta\lambda_i$ is calculated by means of the Rayleigh quotient:

$$\Delta\lambda_i = \frac{x_i^T (K - \lambda_0 B)x_i}{x_i^T Bx_i} \quad (\text{P3.5})$$

here the transpose (the symbol T) is taken because the matrices K and B are complex symmetric.

If the matrix K is a nonsymmetric complex matrix, a second sequence is defined by the left-hand eigenvectors, yielding the iteration:

$$y_0 = x_0$$

$$(K - \lambda_0 B)x_i = \Delta\lambda_{i-1} Bx_{i-1} \quad (\text{P3.6a})$$

$$(K - \lambda_0 B)^T y_i = \Delta\lambda_{i-1} B y_{i-1}. \quad (\text{P3.6b})$$

In the nonsymmetric case an improved eigenvalue is de-

termined from both the right-hand and left-hand eigenvectors x_i, y_i by using the generalized Rayleigh quotient,

$$\Delta\lambda_i = \frac{y_i^T (K - \lambda_0 B)x_i}{y_i^T Bx_i}. \quad (\text{P3.7})$$

The system of linear equations is solved by LU factorization, i.e.,

$$A = K - \lambda_0 B = LU. \quad (\text{P3.8})$$

This LU decomposition preserves the sparse structure of A and can be performed efficiently. With the definitions

$$r_{i-1} = \Delta\lambda_{i-1} Bx_{i-1} = \Delta\lambda_{i-1} \bar{x}_{i-1} \quad (\text{P3.9a})$$

$$s_{i-1} = \Delta\lambda_{i-1} B y_{i-1} = \Delta\lambda_{i-1} \bar{y}_{i-1} \quad (\text{P3.9b})$$

the following equations are solved for x_i and y_i

$$Lp_i = r_{i-1}, Ux_i = p_i \quad (\text{P3.10a})$$

$$U^T q_i = s_{i-1}, L^T y_i = q_i. \quad (\text{P3.10b})$$

Utilizing the decomposition of A , it holds that

$$y_i^T (K - \lambda_0 B)x_i = y_i^T L U x_i = q_i^T p_i \quad (\text{P3.11})$$

and the Rayleigh quotient assumes the form

$$\Delta\lambda_i = \frac{q_i^T p_i}{y_i^T Bx_i} = \frac{q_i^T p_i}{y_i^T \bar{x}_i} \quad (\text{P3.12})$$

where $\bar{x}_i = Bx_i$, defined in (P3.9a), is needed to compute the new vector $r_i = \Delta\lambda_i \bar{x}_i$. If A is a complex symmetric matrix, the above equation may be reduced to

$$\Delta\lambda_i = \frac{|p_i|^2}{x_i^T \bar{x}_i}. \quad (\text{P3.13})$$

The iteration is terminated if the error of successive eigenvectors is smaller than a defined tolerance $\epsilon = 4*U$, where U is the machine epsilon. Termination is also forced if the eigenvalue is not monotonically approximated, in order to avoid pathologic iteration paths, (i.e., if $|\Delta\lambda_i - \Delta\lambda_{i-1}| > |\Delta\lambda_{i-1} - \Delta\lambda_{i-2}|$ for $i > 3$). If convergence is achieved at $i = n$ step, the new eigenvalue is given by $\lambda = \lambda_0 + \Delta\lambda_n$, and the eigenvector by $x = x_n$. It is important to realize that an eigenvalue is obtained faster by choosing a new shift rather than by performing more inner iterations. This is especially true when the initial guess value is far from real eigenvalue. Thus an outer iteration is added to modify the shift λ_0 . Termination of the outer iteration is achieved if $\Delta\lambda_0 < \lambda_0 * 2U$. The flowchart of our RQIIM algorithm is described in Fig. 3. Usually four to five inner iterations and two to three outer iterations are needed for convergence. On a DEC3100 workstation it takes about 20 s cpu-time for the above examples.

We always use the last eigenvalue and eigenvector as an initial guess value at the start of the computation for each new bias point. For zero bias, an estimate could be used for the initial value or the Lanczos algorithm [21] could be used to get an approximate eigenvalue, which is fast if the initial eigenvector is not calculated and a low accuracy is used.

REFERENCES

- [1] D. P. Wilt and A. Yariv, "A self-consistent static model of the double-heterostructure laser," *IEEE J. Quantum Electron.*, vol. QE-17, pp. 1941-1949, 1981.
- [2] T. Kumer, "Self-consistent analysis of gain-guided triple stripe laser," *Solid-State Electron.*, vol. 30, pp. 21-31, 1987.
- [3] T. Ohtosh, K. Yamaguchi, C. Nagaoko, Y. Murayama, and N. Chino, "A two-dimensional devices simulator of semiconductor laser," *Solid-State Electron.*, vol. 30, pp. 627-638, 1987.
- [4] K. B. Kahen, "Two-dimensional simulation of laser diodes in the steady state," *IEEE J. Quantum Electron.*, vol. 24, pp. 641-651, 1988.
- [5] M. Ueno, S. Asada, and S. Kume Shiro, "Two-dimensional numerical analysis of lasing characteristics for self-aligned structure semiconductor lasers," *IEEE J. Quantum Electron.*, vol. 26, pp. 972-981, 1990.
- [6] G. H. Song, K. Hess, T. Kerkhoven, and U. Ravaioli, "Two-dimensional simulator for semiconductor lasers," in *Tech. Dig.*, IEEE Int. Electron Devices Meeting, 1989, pp. 143-146.
- [7] G. L. Tan, X. L. Yuan, Q. M. Zhang, W. Ku, and A. Shey, "Two dimensional semiconductor device analysis based on a new finite element discretization employing the S-G scheme," *IEEE Trans. on CAD/ICAS*, vol. 8, pp. 468-478, 1989.
- [8] R. Stratton, "Diffusion of hot and cold electrons in semiconductor barriers," *Phys. Rev.*, vol. 126, no. 6, pp. 65-87, 1982.
- [9] S. P. Gaur and D. H. Navon, "Two dimensional carrier flow in a transistor structure under nonisothermal conditions," *IEEE Trans. Electron Device*, vol. ED-23, pp. 50-57, 1976.
- [10] M. S. Lundstrom and R. J. Schuelke, "Numerical analysis of heterostructure semiconductor devices," *IEEE Trans. Electron Device*, vol. ED-30, pp. 1151-1159, 1983.
- [11] Z. Yu and R. W. Dutton, "SEDAN—a generalized electronic material device analysis program," Stanford University, 1985; M. R. Pinto, C. S. Rafferty, and R. W. Dutton, "PISCES-II-Poisson and Continuity equation solver," *Tech. Rep.*, Stanford Electron. Lab., Stanford University, CA, 1984.
- [12] M. Asada, A. Kameyama, and Y. Sumatsu, "Gain and intervalence band absorption in quantum-well lasers," *IEEE Quantum Electron.*, vol. QE-20, pp. 745-753, 1984.
- [13] M. S. Stern, "Semivectorial polarized finite difference method for optical wave guides with arbitrary index profiles," *IEE Proc.*, vol. 135, pp. 56-63, 1988.
- [14] R. E. Bank and D. J. Rose, "Global approximate Newton method," *Numer. Math.*, vol. 37, pp. 279-295, 1981.
- [15] J. S. Blakemore, "Semiconducting and other major properties of gallium arsenide," *J. Appl. Phys.*, vol. 53, pp. R123-R81, 1982.
- [16] G. L. Tan, Q. M. Zhang, and J. M. Xu, "Computation of field and charge transport in compound semiconductor devices—some new features and methods," *IEEE Trans. Magnet.*, vol. 27, pp. 4158-4161, 1991.
- [17] Z. Y. Zhao, Q. M. Zhang, G. L. Tan, and J. M. Xu, "A new preconditioner for CGS iteration in solving large sparse nonsymmetric linear equation in semiconductor device simulation," *IEEE Trans. CAD-ICS*, vol. 11, pp. 1432-1440, 1991.
- [18] D. L. Scharfetter and H. K. Gummel, "Large-signal analysis of a silicon Read diode oscillation," *IEEE Trans. Electron Devices*, vol. ED-16, pp. 64-77, 1969.
- [19] S. J. Polak *et al.*, "Semiconductor device modeling from the numerical point of view," *Int. J. Num. Meth. Eng.*, vol. 24, pp. 763-838, 1987.
- [20] B. T. Smith *et al.*, *Matrix Eigensystem Routines—EISPACK Guide*, Lecture Notes in Comput. Sci. 6. New York: Springer-Verlag, 1976. B. S. Garbow *et al.*, *Matrix Eigensystem Routines—EISPACK Guide Extension*, Lecture Notes in Comput. Sci. 51. New York: Springer-Verlag, 1977.
- [21] J. K. Cullum and R. A. Willoughby, "Lanczos algorithm for large symmetric eigenvalue computations," vol. 1 Theory; vol. 2, Programs, *Progress in Scientific Computing Series*, Birkhauser/Boston/Basel/Stuttgart, 1985.
- [22] W. Kermer, "Large-scale complex eigenvalue problem," *J. Computation Phys.*, vol. 85, pp. 1-85, 1989.
- [23] G. F. Burn and C. G. Fonstad, "Monolithic fabrication of strain-free (Al,Ga)As heterostructure lasers on silicon substrates," *IEEE Photon. Technol. Lett.*, vol. 4, pp. 18-20, 1992.



Gen-Lin Tan graduated from the Department of Engineering Physics, Qinghua University, China, in 1961.

Since then he has been engaged in teaching and research with the Department of Electrical Engineering, Beijing Polytechnic University, Beijing, China, where he was a professor. His research concerns microcomputer application, integrated circuit analysis and optimization, semiconductor device simulation, and layout optimization of IC design. He is the author or co-author of several articles and three books. During 1987-1988 he was a visiting scholar with the Department of Electrical Engineering and Computer Science, University of California, San Diego, La Jolla. Since 1989, he has been a senior research associate with the Department of Electrical Engineering, University of Toronto, Canada. His current research interests include simulation and optimization of high speed semiconductor devices, optoelectronic devices, and integrated circuits.



Naveen Bewtra was born in New Delhi, India. He received the B.A.Sc. and M.A.Sc. degrees from the University of Windsor in 1988 and 1992, respectively. The topic for his M.A.Sc. thesis was optically programmable neural networks.

From 1988 to 1990, he was with Northern Telecom Canada as a design engineer. He is currently a Ph.D. student at the University of Toronto in the Department of Electrical Engineering. He is working in the areas of thermal interaction in OEIC's and optically programmable

neural networks.



Keith Lee was born in Hong Kong on July 28, 1965. He received the B.A.Sc. degree in engineering physics in 1988 and the M.Sc. degree in 1990 in physics, both from the University of Toronto.

Currently, he is working towards the Ph.D. degree in the area of semiconductor lasers with the Photonics Group, Department of Electrical Engineering, University of Toronto.

Mr. Lee is a member of the C.A.P. and P.F.N.W.

J. M. Xu (M'87-SM'91) received the Ph.D. degree in electrical engineering from the University of Minnesota, Minneapolis, in 1987.

He joined the faculty of the Department of Electrical Engineering at the University of Toronto, Toronto, Canada, in 1987. He is now the J. M. Ham Chair Professor in Optoelectronics, jointly sponsored by the Natural Science and Engineering Research Council of Canada and Bell-Northern Research Ltd. He has authored and co-authored over 65 refereed papers in physics and engineering journals, and over 20 refereed conference papers. He holds 7 patents on electronic and photonic devices. His current research interests include semiconductor physics, nanostructures, and compound semiconductor device design, modeling, and measurements. Currently, he leads a group of 14 researchers and graduate students of the Optoelectronics Laboratory at the University of Toronto and conducts research primarily in the areas of heterostructure transistors, quantum well devices, lasers, waveguiding devices, and other optoelectronic devices, as well as large scale computer simulations. The Optoelectronics Laboratory is sponsored by Bell-Northern Research Limited through the BNR-NSERC Chair program. He is a principal investigator of the Ontario Laser and Lightwave Research Center. He is also an Investigator of the Ontario Center for Materials Research and a Key Associate of the Information Technology Research Center. He has served on a number of conference program committees, and has been a referee for seven physics and engineering journals.

Dr. Xu is an associate editor for IEEE TRANSACTIONS ON ELECTRON DEVICES.