Theory of Gain Spectra for Quantum Cascade Lasers and Temperature Dependence of their Characteristics at Low and Moderate Carrier Concentrations

Vera B. Gorfinkel, Serge Luryi, *Fellow, IEEE,* and Boris Gelmont, *Senior Member, IEEE*

***Abstract*—We have developed a theory describing the operation of lasers based on intersubband transitions in a quantum well. The theory combines a first-principles description of the intersubband lineshape and the optical gain with kinetic models for carrier heating. The theory is consistent with the experimental data available and suggests new ways of improving the laser design for room-temperature operation with high output power. At low carrier concentrations, it is possible to achieve positive values of the gain at room temperature even in the absence of an overall population inversion between quantum-well subbands. For higher (but still moderate) concentrations, the theory predicts a peculiar dependence of the output wavelength on the pump current, including a regime where the lasing wavelength switches "digitally" between two stable values.**

I. Introduction

IMPLEMENTATION of the quantum cascade laser (QCL) [1] is a great milestone in the history of semiconductor devices. The intersubband laser has many potential applications in the mid-infrared range ($\lambda \gtrsim 4$ μm). Since the first demonstration, the Bell Laboratories group has continued a rapid development of the QCL, improving such characteristics as the threshold and the maximum temperature of operation [2]–[4]. Nevertheless, the QCL operation is not yet fully understood. The temperature behavior of the threshold, the abrupt disappearance of lasing above a certain critical temperature $T_{\rm cr}$, the nature of $T_{\rm cr}$, the low differential efficiency, and the disappearance of lasing with increasing current—this is only a partial list of questions which lack clear understanding. Theory lags behind the dramatic experimental progress, despite a quarter-century-long quest for an intersubband laser [5].

One reason for the lack of a quantitative theory of the QCL is inadequacy of the usual approach which treats intersubband lasers by analogy with a two-level atomic system, assuming parallel parabolic subbands. This analogy fails not so much because of the nonparabolicity effects—which have been taken into account in one form or another by many workers—but because of the *very existence of the transverse degrees of freedom*, corresponding to in-plane motion of carriers. These degrees of freedom store a considerable amount of energy dissipated in the carrier motion through a QCL period, fundamentally changing the lineshape of intersubband resonance and the spectral characteristics of gain. Theoretical expression for these characteristics has been obtained in our earlier work [6]. Based on these results, we presently develop a kinetic model of the QCL operation—including the all-important hot carrier effects.

The intersubband gain spectral function $g(\Omega)$ has been described [6] as a functional on the distribution functions, $f_1(\varepsilon_1)$ and $f_2(\varepsilon_2)$, corresponding to the occupation probabilities of kinetic energy states in the subbands 1 and 2, respectively. Generally, these distributions are nonthermal and must be derived from a transport model. As carriers cascade down the QCL heterostructure, they gain energy. Depending on the laser design, there may be different scenarios of how this energy is distributed among various electrons and dissipated into the lattice. In this work, we consider two simplified situations, referred to as the "high-concentration" and the "low-concentration" models. The former corresponds to the situation when the electron–electron (ee) interaction is sufficiently fast to equilibrate the input power among all free carriers in a QCL period. The low-concentration model describes the opposite limit, when the ee interaction is negligible and hence different groups of electrons have different average energies and different energy distributions.

The intersubband gain has a strong dependence on the shape of the carrier energy distributions in the two subbands because these distributions control the rate of scattering which breaks the phase of intersubband resonance. For high-carrier kinetic energies, the dominant phase breaking mechanism is "transverse," i.e., controlled by intrasubband scattering [7]. The intersubband processes are typically slower because they involve a large momentum transfer.

Among the possible transverse phase-breaking mechanisms, we shall focus on the interaction with polar optic phonons [8]–[10]. This implies the dependence of $g(\Omega)$ on the lattice temperature T, which governs the population $N_{\rm ph}$ of optic phonons $\hbar\omega_{\rm ph}$. Most importantly, the gain becomes strongly influenced by the nonequilibrium distributions $f_{1,2}(\varepsilon)$ in the two subbands, because the dominant scattering process, *emission* of optic phonons, has a threshold nature, its rate $\gamma(\varepsilon)$ being proportional to the step function $\theta(\varepsilon - \hbar\omega_{\rm ph})$.

Assumption of the dominant role of optical phonons in transverse phase breaking involves the neglect of two key

Manuscript received March 26, 1996; revised July 30, 1996.

V. B. Gorfinkel and S. Luryi are with the Department of Electrical Engineering, State University of New York at Stony Brook, Stony Brook, NY 11794-2350 USA.

B. Gelmont is with the Department of Electrical Engineering, University of Virginia, Charlottesville, VA 22903-2442 USA.

Publisher Item Identifier S 0018-9197(96)08040-2.

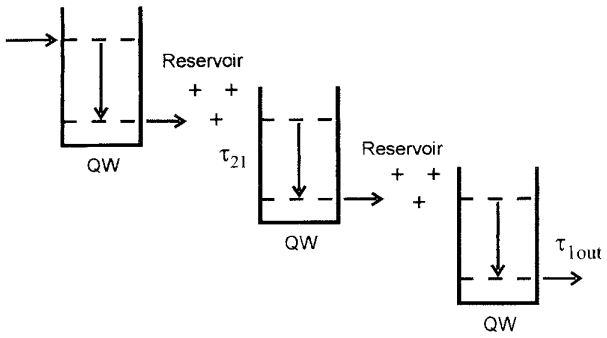


Fig. 1. Schematic diagram of the quantum cascade laser structure and the kinetics of carrier transport. To ensure identical conditions, each cascade period must be neutral. The mobile charge in the undoped quantum well is compensated by the positive donor charge in the reservoir. The rates of (nonradiative) transitions between the subbands and the escape rate from the lower subband into the reservoir are characterized by time constants τ_{21} and $\tau_{1\text{out}}$, respectively.

mechanisms, the impurity scattering and the ee interaction. We have considered these effects and will discuss them in a separate publication [11]. Qualitatively, we found that distant impurities had only a minor influence, owing to a characteristic cancellation effect for transverse phase relaxation by elastic scattering [7]. On the other hand, we found that the ee scattering is very important and becomes dominant at high carrier concentration.[1] Our present approach is valid when the total sheet carrier concentration n_D per period is not too high $n_D \lesssim 10^{11}$ cm^{-2}. The theory presented in this paper can be considered nearly rigorous in the low-concentration regime and semiquantitative in the moderate-concentration regime.

II. Rate Equations

Kinetics of the carrier transport in the QCL and the device structure are illustrated in Fig. 1. Under the operating conditions, each period must have net zero charge: by Gauss' law, this is necessary in order that subsequent periods could replicate each others' electrostatic state. To avoid the space charge accumulation associated with current flow, one therefore needs a reservoir of positive fixed charge, compensating the negative mobile charge in each period. Introduction of such a reservoir, implemented as a doped superlattice region, is the key design innovation that led to the successful implementation of a unipolar laser [1]. Subsequent QCL designs [3] entrusted the reservoir with an additional mission to suppress the unwelcome tunneling from the upper level into the continuum. For this purpose, the superlattice is implemented as an electronic Bragg filter with a stop band in the range of energies of upper level states. This elegant approach places stringent demands on band structure modeling and layer control by molecular beam epitaxy. We shall assume that carriers enter the quantum well (QW) only via the upper and leave only via the lower subband, neglecting a leakage from the upper subband directly into the reservoir.

[1] Increasing the sheet carrier concentration per QCL period above $n_D \lesssim 10^{11}$ cm^{-2} degrades the performance of an intersubband laser, as has been noted empirically [12]. However, this effect should not be attributed to impurity scattering. It arises mostly due to the phase breaking by ee collisions.

As the carrier density in the QW rises with the current, the total sheet concentration in both subbands $n_1 + n_2$ may not exceed the doping level n_D (per unit area), provided in a reservoir outside the QW. The subband concentrations n_i are found from the following rate equations:

$$\frac{\partial n_2}{\partial t} = J - \frac{n_2}{\tau_{21}} - \bar{g}S \tag{1a}$$

$$\frac{\partial n_1}{\partial t} = \frac{n_2}{\tau_{21}} + \bar{g}S - \frac{n_1}{\tau_{1\text{out}}} \tag{1b}$$

where J is the current flux and S is the photon density per unit area in the lasing mode (frequency Ω_L). The gain $\bar{g} = \bar{c}g(\Omega_L)$ is the intersubband gain (in units of sec^{-1}) at the lasing frequency, where $\bar{c} = c/\sqrt{\kappa_\infty}$ is the speed of light in the lasing mode. The time constants τ_{21} and $\tau_{1\text{out}}$ describe, respectively, the rate of nonradiative intersubband transitions and carrier removal from the bottom subband. The ratio of these constants

$$\xi_0 \equiv \frac{\tau_{1\text{out}}}{\tau_{21}} \tag{2}$$

is an important design parameter of an intersubband laser, as it determines the carrier density ratio in a steady state below threshold, $n_1 = \xi_0 n_2$. Above the threshold, the ratio n_1/n_2 may deviate from ξ_0 due to the influence of lightwave S. However, the relation $n_1 = J\tau_{1\text{out}}$ remains rigorous. In order to calculate the characteristics above threshold, we need to solve (1), including terms with S. The steady-state values of n_1 and n_2 are given by $n_1 = J\tau_{1\text{out}}$ and

$$\xi \equiv \frac{n_1}{n_2} = \frac{\xi_0}{1 - \bar{g}S/J} = \frac{\xi_0}{1 - \bar{c}\alpha_{\text{Loss}}S/\Gamma J}. \tag{3}$$

The subband concentrations n_1 and n_2 continue to vary above the threshold, even though the gain itself is pinned at the threshold value $\bar{g} = \bar{c}\alpha_{\text{Loss}}/\Gamma$. We remark that in conventional bipolar lasers, dependence of the gain on the carrier temperature T_e does also give rise to a variation in the carrier concentration above the threshold [13]. A new feature of the unipolar laser is the dependence of gain $\bar{g}$ on the n_1/n_2 ratio, the latter being affected by the optical wave. This dependence has no analogy in the bipolar case, where *local* neutrality condition ensures *equal* concentration of injected electrons and holes. In the QCL, the total QW charge $n_1 + n_2$ is compensated only by the donor charge in the reservoir (see Fig. 1).

III. Spectral Gain Function and the Scattering Model

The dispersion relations $\varepsilon_{1,2}(k)$ in both subbands are different and nonparabolic. An excellent approximation to all the effects of interest here is obtained by regarding the subbands themselves as parabolic, but characterized by different effective masses m_1 and m_2. In this approximation, the density of states in each subband becomes constant and the functions f_i are normalized in a simple way as

$$\frac{m_i}{\pi\hbar^2}\int_0^\infty f_i(\varepsilon_i)d\varepsilon_i = n_i, \quad i = 1, 2. \tag{4}$$

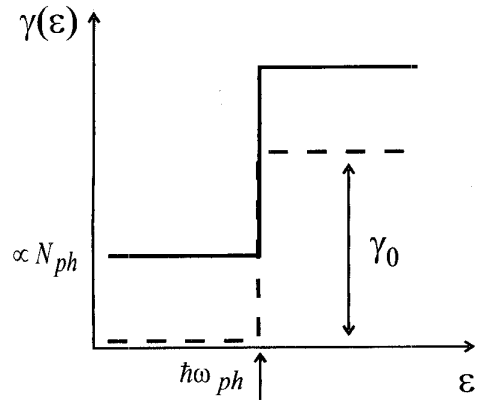


Fig. 2. Optical phonon scattering rate in two-dimensional subbands. Electron energy ε is referenced to the subband bottom. The dashed and the solid lines describe the cases of low and high lattice temperatures, respectively.

Assuming that the characteristic electron scattering rates are much lower than optical frequency Ω, the gain spectra can be expressed in the form [6]

$$g(\Omega) = \frac{4e^2|z_{12}|^2 m_2 \Omega}{\hbar^3 a c\sqrt{\kappa_\infty}} \int_0^\infty \frac{d\varepsilon \gamma(\varepsilon)[f_2(\varepsilon) - f_1(\varepsilon_1)]}{[\Omega - \Omega_\varepsilon]^2 + [\gamma(\varepsilon)]^2} \quad (5)$$

where z_{12} is the transition matrix element, $\sqrt{\kappa_\infty}$ the refractive index, a the QW width, Ω_ε the optical transition frequency for the in-plane electron momentum $\hbar k = \sqrt{2m_2\varepsilon}$, namely, $\hbar\Omega_\varepsilon \equiv \hbar\Omega_0 + \varepsilon_2 - \varepsilon_1$, where $\varepsilon_2 \equiv \varepsilon$ and $\varepsilon_1 = \hbar^2 k^2/2m_1$ are kinetic energies in the upper and lower subbands, respectively, characterized by the effective masses m_2 and m_1 and the distribution functions f_2 and f_1. The function $\gamma(\varepsilon)$ describes the transverse phase relaxation rate due to intrasubband scattering.

Validity of (5) is not restricted to any particular scattering mechanism, responsible for the transverse phase relaxation. In this work, it is assumed that $\gamma(\varepsilon)$ is dominated by the interaction with polar optical phonons. Optical phonon scattering in two-dimensional subbands has a sharp step at the phonon emission threshold (see. Fig. 2)

$$\gamma(\varepsilon) = \gamma_0 \times \begin{cases} N_{\rm ph} \\ (N_{\rm ph}+1)\theta(\varepsilon - \hbar\omega_{\rm ph}) \end{cases} \quad (6)$$

where the top line corresponds to absorption and the bottom line to emission of optical phonons, $N_{\rm ph}$ is the phonon Planck function, and $\theta(\varepsilon)$ is a step function. The threshold nature of $\gamma(\varepsilon)$ has important consequences for the lineshape. However, the ultimate sharpness of the step, peculiar to the QW case, is not essential for our calculations below. Results obtained with the three–dimensional (3-D) scattering rate function are quite similar (in the 3-D case, where the density of states vanishes at low energies, the step is softer). For a sufficiently narrow QW of any shape, the constant $\gamma_0 \approx 10^{13}$ s^{-1} is given by [8], [9]

$$\gamma_0 = \frac{\pi}{2}\frac{e^2}{\hbar}\left(\frac{1}{\kappa_0} - \frac{1}{\kappa_\infty}\right) q_{\rm ph} \quad (7)$$

where $q_{\rm ph} = \sqrt{2m_e\omega_{\rm ph}/\hbar}$.

The above expression (5) for the intersubband optical gain takes into account the transverse degrees of freedom of QW electrons and in essence replaces the conventional two-level model of intersubband transitions by a two-band model, which includes from the first principles such effects as energy-dependent scattering and the subband nonparabolicity. Inclusion of these effects leads to a qualitative change in the lineshape of the intersubband resonance, so that the maximum spectral intensity is no longer related to the linewidth, as it would be in a parabolic model (see the Appendix).

IV. Nonequilibrium Carrier Distributions and Intersubband Gain Spectra

Deviations from equilibrium arise from the power $\mathcal{P}_{\rm in} \approx J \cdot \hbar\Omega$ per unit area, dissipated in each cascade period. Depending on the QCL design, there may be different scenarios of how this power is distributed among various electrons and dissipated into the lattice. In this work, we consider two simplified situations, referred to as the high-concentration (A) and the low-concentration (B) models. Model A corresponds to the situation when $n_D \gtrsim 10^{11}$ cm^{-2} and the ee interaction is sufficiently fast that $\mathcal{P}_{\rm in}$ is shared among all free carriers in the period. In this limit, it is permissible to characterize the distribution functions of carriers in each subband by a quasi-equilibrium, characterized by an effective temperature T_e and, moreover, assume the same T_e for both subbands (and also for the ensemble of carriers remaining in the reservoir). Model B describes the opposite limit, $n_D \ll 10^{11}$ cm^2, when the ee interaction is negligible and hence different groups of electrons have different energy distributions. The concept of an effective temperature is no longer valid in this situation. At low concentrations, most of the input power is dissipated via the lower subband.

A. High-Concentration Model

It is well known [14] that when the bulk carrier concentration is well above 10^{17} cm^{-3}, the ee scattering process dominates over optical phonon emission. In this regime, it is perfectly justified to use the electron temperature approximation for $f_i(\varepsilon_i)$ and assume, in the energy balance equation, that all n_D electrons in the cascade period acquire the same T_e. However, in this range, one should also consistently take into account the phase-breaking effect of ee collisions, as well as the dependence of the electron cooling rate on the carrier density and effective temperature. We shall discuss these effects quantitatively in a separate publication. Here we confine ourselves to the consideration of optical phonon scattering as the dominant phase-breaking mechanism, which limits the validity of model A to a range near but not much above $n_D \approx 10^{11}$ cm^{-2}, which is similar to the experimental situation in reported intersubband lasers [1]–[4].

At these moderate levels of injection, we can use Boltzmann's statistics, since even for $T_e = 100$ K, we have $n_2/N_C \ll 1$, where $N_C = mkT_e/\pi\hbar^2$ is the effective density of states in a 2-D subband. The distribution functions f_1 and f_2, normalized as in (4), are then given by

$$f_i(\varepsilon_i) = \frac{\pi\hbar^2 n_i}{m_i k T_e} e^{-(\varepsilon_i)/(kT_e)}. \quad (8)$$

In the Boltzmann approximation, the energy balance equation is of the form $\mathcal{P}_{\rm in} = n_D k(T_e - T)/\tau_\varepsilon$, where τ_ε is the energy relaxation time. Whence we see that the carrier temperature is a linear function of the current

$$k(T_e - T) = \frac{J\tau_\varepsilon \hbar\Omega}{n_D}. \tag{9}$$

In the Boltzmann approximation, our expression (5) for the gain can be further simplified to

$$g(\Omega) = \frac{4e^2|z_{12}|^2 n_2 \Omega}{\hbar a c\sqrt{\kappa_\infty} k T_e} \times \int_0^\infty \frac{d\varepsilon \gamma(\varepsilon) e^{-\varepsilon/kT_e}}{[\Omega - \Omega_\varepsilon]^2 + [\gamma(\varepsilon)]^2}\left(1 - \frac{f_1}{f_2}\right) \tag{10}$$

where $\varepsilon_2 \equiv \varepsilon$ and $\varepsilon_1 \equiv \varepsilon m_2/m_1$, so that

$$\frac{f_1}{f_2} = \frac{n_1}{n_2}\frac{m_2}{m_1} e^{(\varepsilon/kT_e)(1-m_2/m_1)}. \tag{11}$$

Equations (10) and (11) suggest that the gain can be positive even in the absence of inversion between the two subbands, i.e., for $n_1 \gtrsim n_2$. This effect—in model A—occurs owing to the nonparabolicity ($m_2 > m_1$): at a sufficiently high wavevector k, the occupation probability of state $\varepsilon_2(k)$ in the upper subband is higher than that of state $\varepsilon_1(k)$, even though the lower subband has higher overall population. This can be also seen in a different way: in the high-concentration regime, where both subbands are characterized by the same effective temperature, it is possible to introduce the subband quasi-Fermi levels E_{F_1} and E_{F_2} for the distributions f_1 and f_2, respectively. One can then show that the range of $g(\Omega) > 0$ corresponds to the relation $\hbar\Omega < E_{F_2} - E_{F_1}$, familiar from the theory of conventional semiconductor lasers [15].

Gain spectra calculated in the high-concentration limit, $n_D \gtrsim 10^{11}$ cm^{-2}, are shown in Fig. 3. We have assumed a Maxwellian distribution in both subbands, characterized by an effective temperature T_e, which for concreteness we take to coincide with the lattice temperature, $T_e = T$. Note the existence of a range of $g > 0$ in Fig. 3(a), even though we assumed no overall inversion between the two subbands, $\xi = 1$. In this regime, the peak gain is rather low and the peak shifts to longer wavelengths at higher T. These effects have been observed experimentally [4]. In the high-concentration regime, the range of positive gain for $n_1 \geq n_2$ arises entirely due to nonparabolicity. Dependence of the gain spectra on the subband inversion parameter $\xi = n_1/n_2$ is shown in Fig. 3(b) for a constant temperature. As expected, the gain spectra improve, as the ratio ξ decreases.

B. Low-Concentration Model

In contrast to the case described in Section A, this model is rigorous in the limit $n_D \ll 10^{11}$ cm^{-2} and requires no additional assumptions. In the low carrier density limit, all of the power $\mathcal{P}_{\rm in}$ is dumped into the lower subband and the distribution functions of carriers in the upper subband and in the reservoir both remain near equilibrium. Optical phonon bottleneck effects are not important in this limit either.

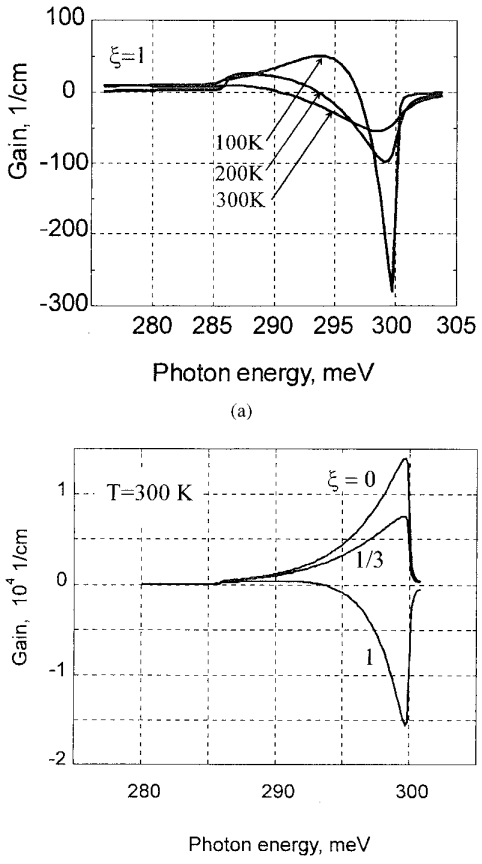


Fig. 3. Calculated intersubband gain spectra in the high-concentration limit, $n_D \approx 10^{11}$ cm^{-2}, and the following material parameters: $E_G = 1$ eV and $a = 76$ Å, resulting in $E_1 = 138$ meV, $E_2 = 438$ meV, ($\hbar\Omega_0 = 0.3$ eV), and $m_1 = 1.28\ m_e$, $m_2 = 1.88\ m_e$, where $m_e = 0.04\ m_0$ is the Kane effective mass at the conduction band bottom [6]. We assume the transition matrix element $z_{12} = 15$ Å, $n_2 = 5 \cdot 10^9$ cm^{-2} and $\tau_{21} = 1$ ps, i.e., the current density $J = 800$ A/cm^2. (a) Gain spectra at different temperatures assuming no overall inversion between the two subbands: $n_1 = n_2$. (b) Room temperature gain spectra for different values of the subband inversion parameter $\xi = n_1/n_2$.

For $n_D \ll 10^{11}$ cm^{-2}, the rate of ee collisions is low and the shape of the distribution functions is far from Maxwellian. The dominant scattering process is due to optical phonons causing electronic transitions within the same subband. We assume that electrons enter the upper subband "cold"[2] and their distribution $f_2(\varepsilon)$ corresponds to a quasi-equilibrium at the lattice temperature T. Immediately after a nonradiative intersubband transition (rate $\tau_{21}^{-1} \approx 10^{12}$ s^{-1}), the lower subband electrons are in a state of high kinetic energy $\varepsilon_1 \gg \hbar\omega_{\rm ph}$. Subsequently, they cascade down emitting optical phonons $\hbar\omega_{\rm ph}$ at the rate $\gamma_0 = \gamma(\varepsilon > \hbar\omega_{\rm ph}) \approx 10^{13}$ s^{-1}. This process is illustrated in Fig. 4. Neglecting thermal re-excitation of hot carriers up the optical phonon ladder, the distribution function $f_1(\varepsilon_1)$ is established by solving a simple kinetic equation

$$\frac{df_1(\varepsilon)}{dt} = \gamma_0(f_1(\varepsilon + \hbar\omega_{\rm ph}) - f_1(\varepsilon)) - \tau_{\rm 1out}^{-1} f_1(\varepsilon). \tag{12}$$

[2]This implies a sufficiently fast energy relaxation in transport through the reservoir.

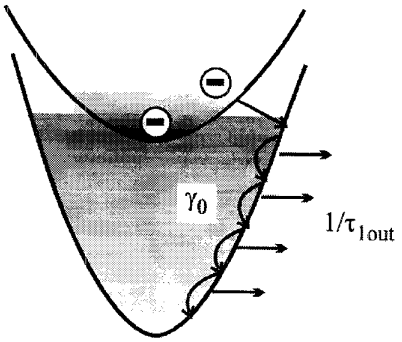


Fig. 4. Formation of the electron energy distribution in the bottom subband at low carrier concentrations.

The escape rate $\tau_{1\text{out}}^{-1} \approx 1$ ps is assumed independent of the kinetic energy. The resultant steady-state distribution is given by a quasi-discrete ladder with the occupation probabilities *decreasing* toward the subband bottom

$$f_1(\varepsilon_1) = f_{\max} \cdot (1+\mu)^{\tilde{\varepsilon}/\hbar\omega_{\text{ph}}} \approx f_{\max} e^{\mu\tilde{\varepsilon}/\hbar\omega_{\text{ph}}} \qquad (13)$$

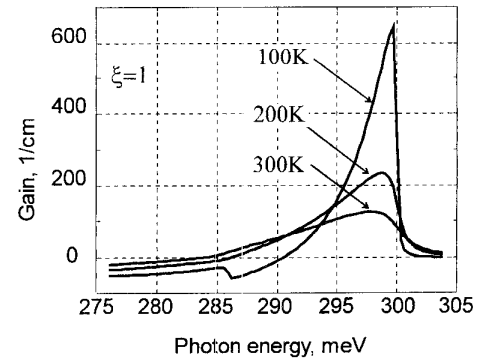
where $\tilde{\varepsilon} < 0$ is referenced to the entry level in the lower subband, $\tilde{\varepsilon} \equiv \varepsilon_1 + \hbar\omega_{\text{ph}} - \hbar\Omega_0$, $f_{\max} \equiv f(\hbar\Omega_0 - \hbar\omega_{\text{ph}})$, and $\mu \equiv 1/(\gamma_0\tau_{1\text{out}})$. The approximate expression in (13) results when $\mu \ll 1$ and describes a negative T_e ensemble. The distribution (13) has no hot-electron tail overlapping the upper subband and hence there is no significant backflow into the upper subband, even though the average energy $\langle\varepsilon_1\rangle \gg \hbar\omega_{\text{ph}}$ may be very large. The detailed shape of the distribution (13) and its quasi-discrete nature are of little importance, as all that matters is the fact of negligible occupation of states at the bottom of the lower subband and the absence of hot-electron effects in the upper subband.

The calculated gain spectra are shown in Fig. 5(a) for several temperatures. The peak gain is substantial even at $T = 300$ K, although we again assumed no overall population inversion between the subbands, $n_1/n_2 \equiv \xi = 1$. In the absence of lasing, the ratio ξ is determined by nonradiative kinetics, (2). In the low-concentration regime, the peak wavelength does not depend on the temperature. To our knowledge, this regime has not yet been realized experimentally.

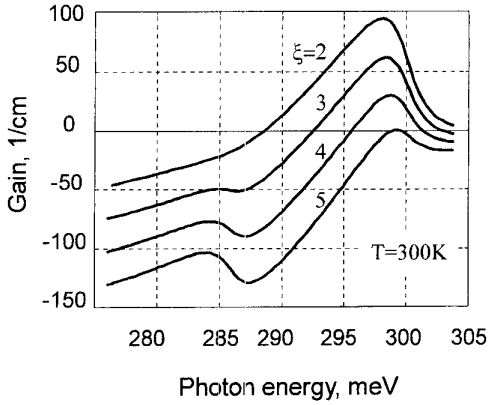
Room-temperature gain spectra calculated for several values of ξ are presented in Fig. 5(b). In contrast to the high-concentration regime, the existence of positive gain in the low-concentration limit does not rely on nonparabolicity and persists to concentrations far from the overall inversion.

V. Calculation of the Laser Characteristics

Our general expression (5) for the intersubband gain allows one to calculate the laser characteristics under both steady-state and dynamic conditions, provided we possess an adequate kinetic model for the variation of the distribution functions f_1 and f_2 under a variable input power $\mathcal{P}_{\text{in}}(t) \approx J(t) \cdot \hbar\Omega$. In as much as the characteristic electron relaxation times are typically faster than the recombination times in a bipolar laser, the QCL can be expected to exhibit ultrafast modulation characteristics. In the present work, however, we shall be



(a)



(b)

Fig. 5. Gain spectra in the low-concentration limit ($n_D \ll 10^{11}$ cm^{-2}). As in Fig. 3, we assume $n_2 = 5 \cdot 10^9$ cm^{-2} and $\tau_{21} = 1$ ps, i.e., the current density $J = 800$ A/cm^2. Other parameters are also as in Fig. 3. (a) Gain spectra at different temperatures assuming no overall inversion between the two subbands: $n_1 = n_2$. (b) Room-temperature gain spectra for different values of the subband inversion parameter $\xi = n_1/n_2$.

concerned only with the steady state or quasi-stationary characteristics, corresponding to the situation when frequencies of the input current variation are much lower than inverse energy relaxation times for electron ensembles in the QW subbands. In this case, we can let the time derivatives in rate equations (1) vanish and solve the resultant algebraic equations for any input current.

As the first step, one must evaluate the spectral gain function $g(\Omega) = g[f_1(\varepsilon_1), f_2(\varepsilon_2), T]$, where the functions f_1 and f_2 are determined from a kinetic model for a given current J. In this way, we find the lasing frequency Ω_L, which maximizes $g(\Omega)$, and the modal gain $g_M \equiv \Gamma g(\Omega_L)$, where Γ is the confinement factor for the optical wave. The threshold is then found from the usual generation condition

$$g_M = \alpha_{\text{Loss}} \qquad (14)$$

where $\alpha_{\text{Loss}} = \alpha_{\text{Mirr}} + \alpha_{\text{Cav}}$ includes both mirror and intracavity losses. If the modal gain g_M were a monotonically rising function of the pump current, we would expect the existence of a lasing threshold at a sufficiently high current. However, in general, we find that $g_M(J)$ is not a monotonic function of the current and, depending on the ambient temperature, the

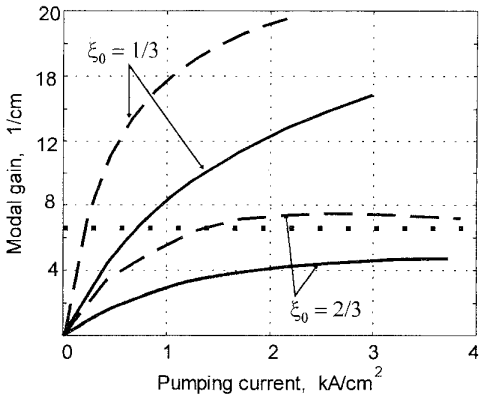


Fig. 6. Graphical solution of (14) describing the steady-state generation condition. Dashed and solid lines correspond, respectively, to $T = 100$ K and $T = 300$ K. Dotted horizontal line is the assumed $\alpha_{\rm Loss} = 7$ cm^{-1} ($\alpha_{\rm Mirr} = 5$ cm^{-1} and $\alpha_{\rm Cav} = 2$ cm^{-1}). Other assumed and calculated parameters are as in Fig. 3; also, we take the energy relaxation time $\tau_\varepsilon = 1$ ps, the cascade period $D = 450$ Å, the number of periods 25, the confinement factor $\Gamma = 0.1$, and the cavity dimensions $L \times W = 2.4$ mm $\times$ 14 μm.

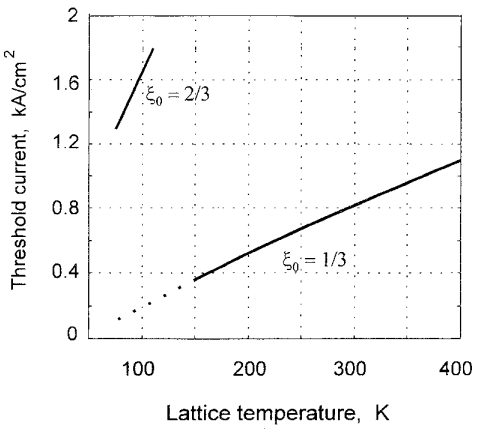


Fig. 7. Temperature dependence of the calculated threshold current density for two quantum cascade lasers, characterized by design parameters ξ_0. Other parameters are as in Fig. 6.

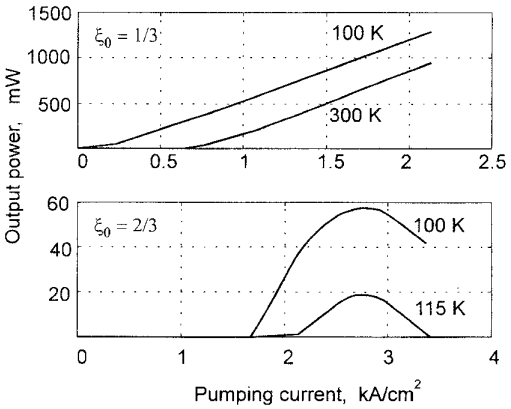


Fig. 8. The calculated light-current characteristics of two cascade lasers with different ξ_0 at different lattice temperatures.

threshold condition may or may not be achievable in a QCL of a given design. Moreover, as we shall see below, there may be a range of currents—with a second threshold—beyond which no lasing is possible.

A. High-Concentration Model

Below threshold, both n_2 and T_e are linear functions of the current [compare with (3) and (9)]. For each value of J in the absence of generation, we can determine the gain spectra by substituting the values of $n_2(J)$, ξ_0, and $T_e(J)$ into our theoretical expression (10) for $g(\Omega)$, which in this model has the structure

$$g(\Omega) = g(n_2, \xi, T_e, T). \tag{15}$$

Graphical solution of (14) is illustrated in Fig. 6 for two different lasers, characterized by the design parameter ξ_0. It turns out that the function $g_M(J)$ in general has a maximum, whose value strongly depends on ξ_0. If this maximum is below $\alpha_{\rm Loss}$ (as is the case for $\xi_0 = 2/3$ at $T = 300$ K), then the generation threshold cannot be achieved. For a given laser (fixed ξ_0), the threshold goes up with increasing temperature. At the same time, the operating range of currents shrinks and the output power decreases. Above a critical temperature $T_{\max}$, the threshold no longer exists. Evidently, $T_{\max}$ is higher for lower ξ_0. Calculated temperature dependence of the threshold current is shown in Fig. 7.

To calculate the light-current characteristics above the threshold, we substitute the values of n_2, n_1/n_2, and T_e obtained from (3) and (9) into (10)–(11) for $g(\Omega)$, evaluated at $\Omega = \Omega_L$ Whence, setting $\bar{g} = \bar{c}\alpha_{\rm Loss}/\Gamma$ in accordance with (14), we find the light-current relationship $S(J)$ in an explicit form (which is somewhat too cumbersome to be displayed). Fig. 8 shows the calculated output power characteristics of a QCL with parameters same as in Figs. 6 and 7.

An interesting feature of the intersubband laser in the high-concentration regime is the fact that the lasing wavelength may vary with the pump current. Fig. 9 shows this variation in the instance of a QCL with $\xi_0 = 2/3$. The peak frequency Ω_L that maximizes the gain $g(\Omega)$ is displayed in Fig. 9(a) together with the output power as functions of pumping current eJ. The corresponding gain spectra are shown in Fig. 9(b) at several representative values of eJ. In general, the value of Ω_L is close to $\Omega_0 \equiv \Omega(\varepsilon = 0)$ at low T_e and, as the current (and hence T_e) increases, Ω_L shifts to $\Omega_{\rm ph} \equiv \Omega(\varepsilon = \hbar\omega_{\rm ph})$. At high T_e, the spectral peak is pinned very tightly near the frequency $\Omega_{\rm ph} \approx \Omega_0 - (m_2/m_1 - 1)\omega_{\rm ph}$, because those transitions that correspond to $\varepsilon > \hbar\omega_{\rm ph}$ are sharply broadened by optical phonon emission, so that the gain spectra are depressed above $\Omega_{\rm ph}$. In the present example, virtually all of the variation of Ω_L occurs below the threshold. This effect may be advantageous for applications requiring a very stable lasing frequency.

On the other hand, the QCL can be designed in such a way that part or most of the variation in Ω_L will occur above threshold. An example of this situation is provided by the $\xi_0 = 1/3$ laser, as illustrated in Fig. 10. Note the sharp transition from Ω_0 and $\Omega_{\rm ph}$ in a narrow interval of currents.

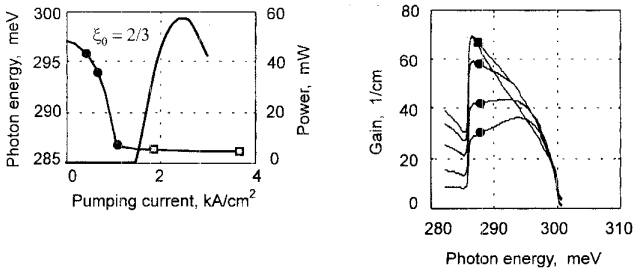


Fig. 9. Variation of the peak frequency of the intersubband gain as a function of the current density for a laser with $\xi_0 = 2/3$ in the high-concentration model. (a) The marked curve corresponds to $\hbar\Omega_L(J)$ and the other curve to the output optical power. (b) Gain spectra at selected values of the current density. Subthreshold spectra are marked by circles and above-threshold spectra by squares. Maximum value of the gain above threshold is pinned by the generation condition (14).

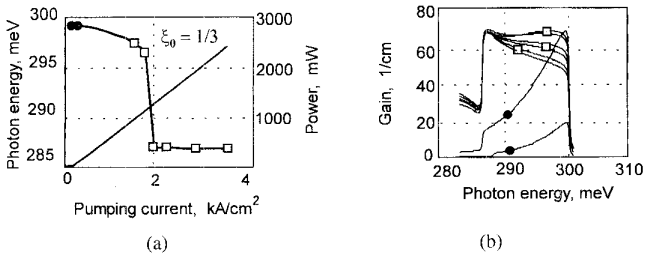


Fig. 10. Same as in Fig. 9 but for a laser with $\xi_0 = 1/3$. The gain above threshold is pinned at the same maximum value, which is either at the peak near Ω_0 or at $\Omega_{\rm ph}$, depending on the current. Most of the variation of the output optical frequency in this laser occurs when the highest value of the gain spectral function jumps from one peak to the other. Carrier concentration in the upper subband varies in the 10^{10} cm^{-2} range.

This means that the QCL has a unique capability of controlling the output wavelength "digitally" between two stable values. This is another unique property of the unipolar laser, which is likely to find interesting applications.

B. Low-Concentration Model

Light-current characteristics at low concentrations are obtained in a similar fashion. In our model, the upper subband distribution function f_2 is thermal, while the lower subband distribution function f_1 corresponds to the negative T_e ensemble (13). We stress that our results are very tolerant to the crudeness of the approximation chosen for f_1—all that really matters is the relatively low probability density near the subband bottom where most of the transitions occur. On the other hand, the assumption of thermal distribution in the upper subband does make a significant difference. High performance of the QCL in the low-concentration limit relies on the possibility for hot carriers to lose most of their excess energy while transiting the reservoir. To ensure this possibility is one of the important design considerations in this limit.

Fig. 11 shows the light-power characteristics calculated in the low-concentration regime for a QCL with $n_D < 10^{10}$ cm^{-2}. Note the low values of n_2 indicated by the dashed lines. Inasmuch as in the low-concentration regime the n_1 electrons are smeared over the subband with very low occupation

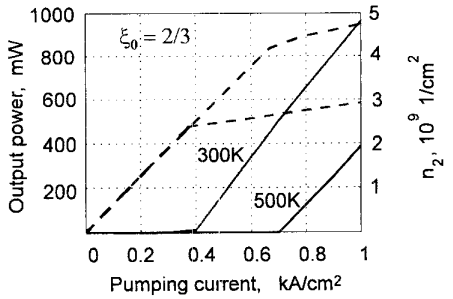


Fig. 11. Calculated light-current characteristics of a quantum cascade laser with low built-in carrier concentration [model B]. Dashed lines show the variation of carrier concentration in the upper subband. The design parameter $\xi = 2/3$, temperature $T = 300$ K (curves 1) and $T = 500$ K (curves 2).

probability at the bottom, low values of ξ are not required for high-temperature operation.

We believe that focusing on the low-concentration regime represents the best strategy for the implementation of low-threshold and high-output-power intersubband lasers. In this regime, electron heating are beneficial: they do not affect the distribution of electrons in the upper subband and at the same time "improve" the distribution in the lower subband by shifting most of the occupation probability to harmless states of high kinetic energy.

Admittedly, our model of the low-concentration regime corresponds to an ideal case, which neglects leakage from the upper subband directly into the reservoir. As has been shown by the Bell Labs group [3], this problem can be solved by designing a suitable Bragg reflector in the reservoir. Perhaps one should also not underestimate the difficulty of maintaining the identical conditions in different cascades at low n_D, in view of doping fluctuations and possible violations of cascade neutrality. Nevertheless, overcoming these difficulties appears to be well within the capability of modern molecular beam epitaxy—of which one of the most impressive triumphs was the implementation of the QCL itself.

VI. Conclusion

We have developed a theory of the quantum cascade laser that includes a rigorous description of the intersubband lineshape and optical gain as well as the very important carrier heating effects. The theory has been applied to two opposite regimes of carrier heating, as controlled by the built-in doping level. Both regimes allow for a high-temperature operation. In the high-concentration regime, the laser wavelength can be made a "digital" function of the current with two sharply defined and switchable wavelength states. The low-concentration regime offers possibility of higher performance, because the carrier heating effects in this case actually improve the effective population inversion for states participating in the lasing transition.

Given the tremendous applications potential of the QCL, we believe the device deserves a major development effort. The improvement strategies that have born most fruit so far have dealt with the kinetics of subband population, effectively the reduction of parameter ξ_0, and with the laser

engineering, such as improved heat sink, cavity design, etc. Present work identifies three new strategies. First, one should strive to reduce the effects of nonparabolicity by employing new material systems and designs to engineer subbands with approximately equal kinetic masses in the QW plane. Next, one can attempt to reduce the carrier heating effects by introducing new carrier cooling channels, e.g., via intervalley phonon emission in the doped reservoir region. Third, one can render hot-carrier effects harmless by implementing the low-concentration regime of operation. Our work shows that the QCL still has an enormous reserve for improvement.

APPENDIX
HOT-ELECTRON EFFECTS AND TRANSVERSE RELAXATION IN A TWO-LEVEL MODEL

If γ were independent of energy, $\gamma(\varepsilon) = \gamma_0$, and the nonparabolicity could be neglected, then for the peak frequency $\Omega = \Omega_0$ one could reduce (5) to the usual expression for a two-level system, $g = g_n'(n_2 - n_1)/a$, where (compare, e.g., with [16])

$$g_n' \equiv \frac{4\pi e^2 |z_{12}|^2 \Omega}{\hbar c \sqrt{\kappa_\infty} \gamma_0}. \tag{A1}$$

However, for the heterosystem used in the experimentally realized intersubband lasers [1]–[4], neither of these approximations is adequate. Thus, in the entire range of temperatures considered in this work, the peak value of the gain is suppressed compared to the parabolic case by at least an order of magnitude. This means that temperature dependence of the gain cannot even be estimated from (A1).

It may be instructive, however, to discuss qualitatively the threshold behavior in a model where (A1) is valid, assuming a QCL implemented in a parabolic material (say, a wide-gap semiconductor heterostructure). Consider the high-concentration limit. Using (A1) together with (14) and (9), we find that the threshold condition is described by the following equation for T_e:

$$\frac{\hbar \gamma_0(T_e)}{k(T - T_e)} = \mathcal{R} \tag{A2}$$

where $\mathcal{R}$ is a dimensionless constant

$$\mathcal{R} = \frac{4\pi e^2 |z_{12}|^2 \Gamma n_D}{\hbar c \sqrt{\kappa_\infty} a \alpha_{\text{Loss}}} \frac{\tau_{21} - \tau_{1\text{out}}}{\tau_\varepsilon}. \tag{A3}$$

Consider a graphical solution to (A2), Fig. 12, assuming that the intersubband linewidth γ_0 is an increasing and *concave* function of the effective carrier temperature.[3] The threshold carrier temperature is determined by the first intersection of a line $(T_e - T) \cdot \mathcal{R}$ with the concave upward curve $\hbar\gamma_0(T_e)$. Depending on the value of $\mathcal{R}$ and the ambient temperature, (A2) may have no solutions at all or—if the value of $\mathcal{R}$ is sufficiently large—two solutions. For a given $\mathcal{R}$, the highest lattice temperature $T = T_{\max}$ at which threshold is possible, is determined by the condition that the line is tangent to the curve. Finally, there is a minimum value of $\mathcal{R}$ such that $T_{\max} > 0$. If $\mathcal{R}$ is less than this minimum value, then there can be no lasing at any temperature.

[3] Intersubband photoluminescence experiments indicate that $\gamma_0(T)$ is an increasing concave function of the *lattice* temperature [17].

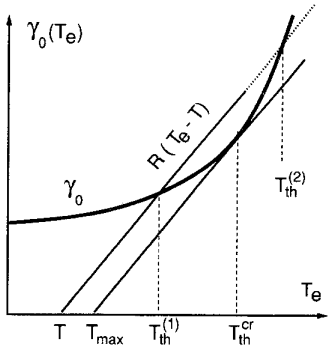


Fig. 12. Graphical solution of the threshold equation (A2) for a parabolic model of the quantum subbands. For a concave upward curve $\hbar\gamma_0(T_e)$ and $T < T_{\max}$, the lasing range is defined by $T_{\text{th}}^{(1)} < T_e(J) < T_{\text{th}}^{(2)}$. In view of (9), the two thresholds $[T_{\text{th}}^{(1)}$ and $T_{\text{th}}^{(2)}]$ in the effective temperature T_e determine both the lower and the upper thresholds for the pump current.

From this example, we conclude that both the transverse relaxation of the intersubband resonance and the hot-carrier effects are indispensable for the correct physical understanding of the temperature behavior of quantum cascade lasers.

REFERENCES

[1] J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, and A. Y. Cho, "Quantum cascade laser," *Science*, vol. 264, pp. 553–556, 1994.
[2] J. Faist, F. Capasso, D. L. Sivco, A. L. Hutchinson, C. Sirtori, S. N. G. Chu, and A. Y. Cho, "Quantum cascade laser: Temperature dependence of the performance characteristics and high T_0 operation," *Appl. Phys. Lett.*, vol. 65, pp. 2901–2903, 1994.
[3] J. Faist, F. Capasso, C. Sirtori, D. L. Sivco, A. L. Hutchinson, and A. Y. Cho, "Vertical transition quantum cascade laser with Bragg confined excited state," *Appl. Phys. Lett.*, vol. 66, pp. 538–540, 1995.
[4] ——, "Continuous wave operation of a vertical transition quantum cascade laser above $T = 80$ K," *Appl. Phys. Lett.*, vol. 67, pp. 3057–3059, 1995.
[5] R. F. Kazarinov and R. A. Suris, "Possibility of the amplification of electromagnetic waves in a semiconductor with a superlattice," *Fiz. Tekh. Poluprovodn*, vol. 5, pp. 797–800, 1971.
[6] B. Gelmont, V. B. Gorfinkel, and S. Luryi, "Theory of the spectral lineshape and gain in quantum wells with intersubband transitions," *Appl. Phys. Lett.*, vol. 68, pp. 2171–2173, 1996.
[7] R. F. Kazarinov and R. A. Suris, "Electric and electromagnetic properties of semiconductors with a superlattice," *Fiz. Tekh. Poluprovodn*, vol. 6, pp. 148–162, 1972.
[8] P. J. Price, "Polar-optical-mode scattering for an ideal quantum well heterostructure," *Phys. Rev.*, vol. B30, pp. 2234–2235, 1984.
[9] B. Gelmont, M. Shur, and M. Stroscio, "Polar optical-phonon scattering in 3-dimensional and 2-dimensional electron gases," *J. Appl. Phys.*, vol. 77, pp. 657–660, 1995.
[10] S.-C. Lee, I. Galbraith, and C. R. Pidgeon, "Influence of electron temperature and carrier concentration on electron-LO-phonon intersubband scattering in wide GaAs/$Al_xGa_{1-x}As$ quantum wells," *Phys. Rev. B*, vol. 52, pp. 1874–1881, 1995.
[11] B. Gelmont, V. B. Gorfinkel, and S. Luryi, unpublished.
[12] F. Capasso and J. Faist, private communication.
[13] V. B. Gorfinkel and S. Luryi, "Fundamental limits for linearity of CATV lasers," *IEEE J. Lightwave Technol.*, vol. 13, pp. 252–260, 1995.
[14] A. A. Grinberg, S. Luryi, N. L. Schryer, R. K. Smith, C. Lee, U. Ravaioli, and E. Sangiorgi, "Adiabatic approach to the dynamics of nonequilibrium electron ensembles in semiconductors," *Phys. Rev. B*, vol. 44, pp. 10536–10545, 1991.
[15] M. G. A. Bernard and G. Duraffourg, "Laser conditions in semiconductors," *Phys. Status Solidi*, vol. 1, pp. 699–703, 1961.
[16] A. Yariv, *Optical Electronics*, 4th ed. Philadelphia, PA: Saunders, 1991, ch. 5.
[17] J. Faist, private communication.

Vera B. Gorfinkel, photograph and biography not available at the time of publication.

Serge Luryi (M'81–SM'86–F'89) received the Ph.D. degree in physics from the University of Toronto, Toronto, ON Canada, in 1978.

Between 1980 and 1994, he was a Member of the Technical Staff at AT&T Bell Laboratories, Murray Hill, NJ. During this time, he served as a Group Supervisor in several device research departments dealing with VLSI, quantum phenomena, and optoelectronics. In 1994, he joined the faculty of the State University of New York at Stony Brook, where he is currently a Leading Professor and Chairman of the Department of Electrical Engineering. He has published over 140 papers and filed 28 US patents in the areas of high-speed electronics and photonic devices, material science, and advanced packaging. From 1986–1990, he served on the Editorial Board of IEEE TRANSACTIONS ON ELECTRON DEVICES first as an Associate Editor and then as Editor.

Dr. Luryi is a Fellow of the American Physical Society. In 1990, Bell Laboratories recognized him with the Distinguished Member of Technical Staff award.

Boris Gelmont (SM'92) received the M.S.E.E. degree (with honors) from the Leningrad Electrical Engineering Institute, Leningrad, U.S.S.R., in 1960 and the Ph.D. and Dr.Phys.Math.Sc. degrees from the A. F. Ioffe Institute of Physics and Technology, Leningrad, U.S.S.R., in 1965 and 1975, respectively.

In 1960, he joined the Ioffe Institute, where he remained until 1991. From 1984–1990, he was also a Professor at the Polytechnic University, Leningrad. In 1990, he joined the University of Virginia, Charlottesville, where he is currently a Senior Scientist with the Department of Electrical Engineering. He has published more than 200 papers in the areas of material science and electronic devices, lasers, and other photonics devices.

Dr. Gelmont is a member of the Commission D of the International Union of Radio Science. He received the State Prize (the highest Soviet award) for the theory of narrowband and zero-band gap semiconductors.