Toward Silicon-Based Lasers for Terahertz Sources

Stephen A. Lynch, Senior Member, IEEE, Douglas J. Paul, Senior Member, IEEE, Paul Townsend, Student Member, IEEE, Guy Matmon, Student Member, IEEE, Zhang Suet, Student Member, IEEE, Robert W. Kelsall, Zoran Ikonic, Paul Harrison, Senior Member, IEEE, Jing Zhang, David J. Norris, Anthony G. Cullis, Carl R. Pidgeon, Pawel Murzyn, Ben Murdin, Mike Bain, Harry S. Gamble, Member, IEEE, Ming Zhao, and Wei-Xin Ni

Abstract—Producing an electrically pumped silicon-based laser at terahertz frequencies is gaining increased attention these days. This paper reviews the recent advances in the search for a silicon-based terahertz laser. Topics covered include resonant tunneling in p-type Si/SiGe, terahertz intersubband electroluminescence from quantum cascade structures, intersubband lifetime measurements in Si/SiGe quantum wells, enhanced optical guiding using buried silicided layers, and the potential for exploiting common impurity dopants in silicon such as boron and phosphorus to realize a terahertz laser.

Index Terms—Boron, far infrared, germanium, impurity, lifetime, phosphorus, pump-probe, quantum cascade laser, resonant tunneling diode (RTD), silicide, silicon, terahertz, waveguide.

I. INTRODUCTION

SEMICONDUCTOR lasers are gaining increasing importance in industrial applications and there have been sustained efforts over the past few years to develop a silicon-based laser source. Such a silicon-based device would represent a considerable advancement in optoelectronic technology, since it would open the way for inexpensive monolithic integrated optical and electronic components.

While III–V semiconductor lasers covering the ultraviolet to the far infrared wavelength ranges are now commonplace, efficient silicon-based lasers have consistently eluded researchers.

This is largely due to the indirect bandgap in silicon, which prevents fast electron–hole recombination. Quantum cascade lasers, however, rely on intersubband transitions, effectively negating effects resulting from the detrimental indirect bandgap [1]. Intersubband transitions in Si/SiGe heterostructures have small energy gaps, resulting in emission in the mid- and far-infrared (terahertz) part of the electromagnetic spectrum [2]–[4]. Recently, there has also been considerable interest in using impurity dopants in silicon such as boron and phosphorus, since it was demonstrated that such materials can show electroluminescent emission in the terahertz region of the electromagnetic spectrum [5]–[9].

With the emergence of several potential applications, sources that emit terahertz radiation have also recently become a “hot” research topic. These include medical imaging [10] (including dental and skin cancer), biological weapons detection [11], explosives detection [12], gas sensing, pollution monitoring, and molecular spectroscopy [13].

The paper is organized as follows. Section II deals with some of the issues particularly pertinent to growth in a strained material system such as Si/SiGe. This is one of the main constraining factors in the design of Si/SiGe active regions. Section III covers resonant tunneling in p-type Si/SiGe. Resonant tunneling through the barriers separating subsequent quantum wells in a quantum cascade active region controls the transport of carriers across the heterostructure. The amount of carriers that can travel through the active region has a direct effect on the gain. Section IV covers electroluminescence measurements at terahertz frequencies from Si/SiGe cascade structures. The amount of gain in any laser depends on the lifetimes of the upper and lower laser levels. In Section V, lifetime measurements performed on Si/SiGe quantum wells are discussed. Another important aspect regarding gain is how much of the optical mode overlaps with the active region. Conventional semiconductor lasers have exploited the refractive index step between epitaxial layers with different atomic mole fractions to achieve a high mode overlap in the vertical (growth) direction. In contrast, at terahertz frequencies, it is necessary to exploit surface-plasmon-enhanced guiding in order to achieve the same effect. Section VI explores the possibility of using buried silicided layers to obtain a high mode overlap in the vertical direction. Electroluminescence measurements in Si/SiGe quantum cascade structures so far have exhibited wide (>10 meV) intersubband emission peaks. Narrow peaks in the emission spectrum of some samples have been attributed to dopant impurities such as boron and phosphorus. Section VII discusses whether these impurities transitions might themselves
be exploited to realize an electroluminescent terahertz silicon-based laser.

II. MATERIAL GROWTH ISSUES

While a III–V terahertz quantum cascade laser has already been demonstrated [1], a silicon-based quantum cascade laser is yet to be realized. This has a lot to do with inherent difficulties of the less mature strained Si/SiGe growth system, the chemistry of dopant materials, and the band structure of the system [14]. All Si/SiGe quantum cascade emitter designs produced to date have been p-type using holes as the unipolar carriers. This is due to the large effective mass for electrons tunneling in the growth direction of Si (1 0 0) being 0.94 $m_e$, where $m_e$ is the free electron mass [14]. This would result in the requirement of subnanometer tunnel barriers to allow significant tunneling through potential barriers. Also, most n-type dopants surface segregate, making the doping of the thin n-type layers difficult without doping layers grown after the As or P has been released in the growth chamber.

Typically the critical thickness at which dislocations occur is several orders of magnitude smaller than the necessary thickness of a quantum cascade active region and strain symmetrization is required [14]. The thickness and mole fractions of each successive SiGe epilayer are carefully adjusted so that each epilayer under compressive strain has neighboring epilayers under tensile strain, so that the average strain across the entire structure is zero. Strain does not have to be balanced over each individual pair of compressive and tensile layers but must be balanced over a distance less than the critical thickness. Therefore, typically the strain is balanced over a cascade period. To date we have demonstrated the growth of over 4 $\mu$m of quantum cascade active region that has been perfectly strain balanced. A good example of this is shown in Fig. 1. This figure shows a transmission electron micrograph (TEM) of few of both the top and bottom epilayers of a 600-period, 4-$\mu$m-thick, quantum well stack grown by gas source MBE.

![TEM showing both the top and bottom few epilayers of a 600-period, 4-$\mu$m-thick, quantum well stack grown by gas source MBE.](image)

![Current versus voltage for three different RTD devices with different areas. Current peak corresponds to the tunneling between the contacts and the HH2 states in the quantum well.](image)

III. RESONANT TUNNELING IN P-TYPE SiGe

The principle of the quantum cascade relies heavily on the basic quantum mechanical principle of tunneling, i.e., a particle may tunnel through a finite potential energy barrier even though it may not classically have enough energy to otherwise surmount the barrier. Resonant tunneling occurs when two energy states on either side of a potential energy barrier have the same energy. Resonant tunneling in a quantum cascade laser controls the transport of carriers across the heterostructure, and therefore has a direct effect on the gain. Significant understanding of the tunneling processes can be obtained from two tunnel barriers with a single quantum well sandwiched in between. Such a device is called a resonant tunneling diode (RTD).

There are many reports of Si/SiGe RTDs in the literature especially p-type grown pseudomorphically on bulk silicon substrates [15], but relatively few reports on high-quality strain-relaxed buffers [16], [17]. In particular, there are very few examples of strain symmetrized p-type RTDs with energy level separations at terahertz frequencies. A number of wafers were grown to gain an understanding of tunneling of holes in Si/SiGe structures with the same composition and barriers as designs for terahertz cascades, and the results from one wafer are presented here. The epitaxial layer structure consisted of a 4-nm i-Si$_{0.6}$Ge$_{0.4}$ quantum well sandwiched between 4-nm i-Si barriers. Either side of the Si barriers had 11 nm of graded i-Si$_x$Ge$_{1-x}$ leading to an Si$_{0.8}$Ge$_{0.2}$ ohmic contact layer doped at a concentration greater than $10^{19}$ cm$^{-3}$. The entire structure just described was grown on top of a Si$_{0.8}$Ge$_{0.2}$ virtual substrate. Mesas were etched using SiCl$_4$ reactive ion etching and Al (1% Si) evaporated to form shallow ohmic contacts. A rapid thermal anneal below 400 °C was performed to prevent spiking of the contacts. A thick SiO$_2$ layer was used to isolate the top and bottom contacts before a via hole was used to contact the top ohmic contact of the device. Fig. 2 shows typical current–voltage curves for the three mesa sizes processed from a single wafer. In each case, a large nonlinearity is observed at around 0.5 V indicating resonant tunneling between the ohmic contacts and the heavy-hole 2 (HH2) state in the quantum well.
details of these calculations are published elsewhere [18]. Fig. 3 (TE) shows one well-defined spectral feature centered on 8 meV, attributed to the lowest energy HH to light-hole (LH) interwell transition, HH1-LH1. It is very difficult to completely rule out thermal origins for this feature (black body). However, since HH-LH or LH-HH transitions allow TE-polarized intersubband radiative transitions, while for HH-HH, LH-LH, or electron-to-electron radiative intersubband transitions only TM polarization can be observed at small k-parallel, it should be possible to observe a discernible difference between the two spectra. This is indeed the case. In the TM spectrum shown in Fig. 4, the 8-meV feature is strongly suppressed while an additional feature between 20 and 30 meV is clearly observed. This feature is not present in the TE spectrum and is identified as the degenerate LH1-LH1 and HH1-HH1 interwell transition. Between 30 and 40 meV, three sharp features are observed labeled 1, 2, and 4. These correspond to the well known p$_{3/2}$ series of Si:B [19], [20]. The considerable differences between Figs. 3 and 4 along with their good agreement with theoretical predictions provide strong evidence that these features do indeed result from intersubband transitions. Further evidence can be found in [4], where the frequency shift of these peaks with applied bias is discussed. The best theoretical fits to the experimental emission spectra require population inversion to be present in the calculations. In contrast, there is poor agreement between experimental and theoretical spectra if thermal equilibrium populations are used. The theoretical modeling therefore suggests that population inversion is obtained in these samples although there is no experimental proof to confirm this prediction.

V. INTERSUBBAND LIFETIME MEASUREMENTS

In order to produce gain in a potential laser material, it is necessary to achieve population inversion. This typically happens when the upper and lower states in the radiative transition have different lifetimes, specifically when the upper laser level has a long lifetime, and the lower laser level has a short lifetime. Using FELIX, the free electron laser facility at Utrecht in the Netherlands, a number of pump-probe experiments were
performed to determine the upper state lifetime of SiGe quantum wells that were engineered to have an intrawell LH1-to-HH1 energy level spacing at terahertz frequencies. Full details of the band structure and experimental details can be found elsewhere [21], [22]. Fig. 5 shows the differential probe transmission against optical delay for three temperatures between 4.2 and 300 K. From this measurement, it is observed that lifetimes as long as 25 ps persist for temperatures up to 300 K. More importantly, this measurement shows that the lifetime remains relatively constant over this large temperature range due to the lack of polar optical phonon scattering in Group IV materials [18]. The intersubband lifetimes decrease as the energy is increased toward the Ge–Ge optical phonon energy as this mechanism becomes the dominant nonradiative process. This is in stark contrast to what occurs in the III–V material system. Here, the lifetime rapidly collapses to subpicosecond values above 40 K due to polar optical phonon scattering. The situation is summarized in Fig. 7, which shows the intrawell LH1-to-HH1 intersubband lifetime plotted against energy. The lifetimes can be observed to decrease with increasing energy, and at 30 meV are limited by the measurement resolution of the pump-probe technique. The Ge–Ge and Si–Si phonon energies are highlighted for clarification.

VI. BURIED SILICIDE WAVEGUIDES

Realizing an active region that exhibits large material gain is of course like winning only half the battle. If the electric field of the propagating electromagnetic mode cannot interact with the material, then no laser can be produced. The magnitude of this interaction is described by the modal overlap. In conventional visible and near infrared semiconductor lasers, this is achieved by engineering a large refractive index step around the active region in order to tightly confine the electromagnetic mode. Unfortunately, as the wavelength gets longer it becomes progressively harder to engineer a significant refractive index step. One of the key enabling technologies in the development of the III–V terahertz quantum cascade laser was the introduction of surface plasmon waveguides [1]. Surface plasmon waves form at the interface between metallic and dielectric media. In the case of the III–V cascade laser, sandwiching the active region between the top Au-doped semiconductor ohmic contact and a bottom highly doped semiconductor layer forms a double plasmon waveguide. For the bottom doped-semiconductor layer to become a reflector, the doping of the layer must be made sufficiently high to result in a negative dielectric constant. As higher doping will result in a better reflector but also will increase the
waveguide losses due to free carrier absorption, the doping density is selected to allow a compromise between modal overlap and waveguide loss.

In silicon, however, the mobility is much lower than in the corresponding III–V material system. This means that the conductivity is also much lower. Unfortunately, a high conductivity is required for the imaginary part of the dielectric constant to become larger than the real part. This is necessary for a negative dielectric constant and a good reflector. For this reason, a heavily doped Si or SiGe layer will not act as a good plasmon reflector unlike \( n^+ \) GaAs in the III–V laser. Assuming a 4-\( \mu \)m-thick SiGe active region sandwiched between a top Al-p\(^{++}\)-SiGe ohmic contact and a bottom contact layer with a realistic doping density of \( 3 \times 10^{20} \) \( \text{cm}^{-3} \), it is possible to calculate the mode overlap as a function of bottom contact layer thickness (Fig. 8) [23]. This is a reasonable approximation of the quantum cascade structures that have been grown, fabricated, and characterized to date with active regions of up to 4-\( \mu \)m thickness on top of \( \sim 3.5-\mu \)m virtual substrates. Fig. 8 shows that this configuration provides a waveguide with very poor modal overlap, with a best case scenario of between 10% and 20% overlap at 4.8 THz.

The situation can be dramatically improved if advantage is taken of a technology that is only available in the Si (and SiGe) material system, i.e., replacing the bottom doped-contact layer with a buried silicide layer. The silicide layer has the advantage of having a much higher electrical conductivity than a semiconductor, and typically within one order of magnitude of good metals. A full description of the fabrication of a buried silicide layer can be found elsewhere [24]. The four most mature silicide technologies are WSi\(_2\), CoSi\(_2\), TiSi\(_2\), and NiSi. So far, we have concentrated on WSi\(_2\) because the low-resistivity phase is stable to temperatures over 800 °C allowing Si and SiGe hetero-layer growth after a silicide has been formed. Fig. 9 shows the mode overlap for a 4-\( \mu \)m-thick SiGe active region, sandwiched between a top Al-p\(^{++}\)-SiGe ohmic contact and a bottom WSi\(_2\) contact layer, as a function of silicide thickness [23]. The dramatic improvement in the situation is quite noticeable. Using a buried silicide of thickness 0.3 \( \mu \)m or larger, the overlap is now approximately 95% at 4.8 THz.

Fig. 10 shows the calculated waveguide losses as a function of active region thickness for WSi\(_2\). Again it is assumed that the active region is sandwiched between a top Al-p\(^{++}\)-SiGe ohmic contact and a bottom WSi\(_2\) contact layer. Once the active region thickness becomes larger than 4 \( \mu \)m, the losses become quite small and certainly comparable to the waveguide losses and modal overlaps quoted in terahertz GaAs quantum cascade lasers [1].

Fig. 11 shows a TEM of a wafer we have grown with a buried WSi\(_2\) layer wafer [24] to produce high modal overlap. It consists of 4.005 \( \mu \)m of SiGe quantum cascade active region, preceded by a 0.990-\( \mu \)m-thick Si\(_{0.2}\)Ge\(_{0.8}\) constant composition layer, a 2.830-\( \mu \)m graded Si\(_x\)Ge\(_{1-x}\) buffer, and a 1.280-\( \mu \)m-thick Si layer, on top of a 440-nm-thick WSi\(_2\) layer.
VII. IMPURITY-BASED TERAHERTZ ELECTROLUMINESCENCE

In Section IV, it was noted that between 30 and 40 meV three sharp features are observed in both the TE-polarized (Fig. 3) and the TM-polarized (Fig. 4) electroluminescence spectra. These transitions were attributed to the well known $p_{3/2}$ series of Si:Bi [19], [20]. The conventional wisdom has been that these transitions are excited by impact ionization, though recent work (to be published) has shown that thermal effects may play a more important role. Under certain bias conditions and at the lowest temperatures in some structures, the intersubband transition can be suppressed and the terahertz electroluminescence arises purely from the impurity transitions [5], [9].

The possibly of harnessing these transitions to make an impurity-based laser has generated considerable excitement over the past few years. The first of these impurity-based lasers was demonstrated in 2002 [25]. A pulsed CO$_2$ pump laser was used to excite electrons high into the conduction band. These electrons rapidly relax into the upper $2p_0$ lasing state via intervalley acoustic and optical phonon scattering. The long lifetime of the $2p_0$ state causes a relaxation bottleneck, and leads to a population inversion. Stimulated emission on the $2p_0 \rightarrow 1s(T_2)$ transition at 22.4 meV (5.41 THz) then occurs [26]. A similar mechanism was exploited in bismuth-doped silicon [27] (Si:Bi). Here, lasing transitions occur between the upper $2p_{\pm}$ level and the closely spaced $1s(E)$ and $1s(T_2)$ lower levels. This results in two emission lines at 23.76 meV (5.74 THz) and 25.49 meV (6.16 THz). Lasing emission was also observed from similar $2p_{\pm} \rightarrow 1s(E)$ and $2p_{\pm} \rightarrow 1s(T_2)$ transitions in arsenic-doped silicon (Si:As) [28]. This time the emission lines were at 24.94 meV (6.03 THz) and 26.33 meV (6.36 THz). Antimony-doped silicon (Si:Sb) was also shown to exhibit lasing [29]. One laser emission line at 24.07 meV (5.15 THz) from the $2p_0 \rightarrow 1s(T_2)$ transition is observed.

Therefore, an obvious question is: Can such impurity transitions be used to produce an electrically pumped laser? Fig. 12 shows two terahertz electroluminescence spectra for a sample containing boron and another containing phosphorus. All electroluminescence measurements so far suggested that the appearance of these transitions in the electroluminescence spectra was highly dependent on the temperature of holes in the device. This hole temperature in turn was dependent on both the heat-sink temperature (i.e., the temperature of the cold finger of the cryostat) and on the amount of current flowing through the particular device. In the case of a biased device, if sufficient current is flowing, then it can contribute to an additional rise in temperature through joule heating.

In order to properly assess the effect of temperature on the impurity transitions, it is necessary to consider a passive device with no current flowing. In the simplest case, this can be achieved by considering the effect of temperature on the absorption properties of a moderately doped bulk silicon wafer ($10^{15} - 10^{16}$ cm$^{-3}$). Fig. 13 shows the absorption from a boron-doped silicon wafer with temperature at terahertz frequencies (2.4–14.4 THz). At very low heat sink temperatures, three very sharp absorption features can be observed between 30 and 40 meV. These correspond to the lines 1, 2, and 4 of the well known $p_{3/2}$ series of Si(B) [19], [20]. As the heat sink temperature is raised between 4.2 and 90 K, the transitions become progressively weaker, until they are finally extinguished around 100 K [9]. This corresponds closely with a reduction in the total absorption up to 100 K. In contrast, the behavior of the phosphorus-doped sample in Fig. 14 is quite different. Unlike the boron-doped sample, the maximum absorption of the phosphorus-doped sample is not at 4.2 K, but at 50 K. Close examination of Fig. 14 reveals that the absorption spectrum is dominated by three features in the range 30–45 meV at 4.2 K. These correspond to the $2p_0$, $2p_{\pm}$, and $3p_{\pm}$ Lyman series transitions to the ground state $1s(A_1)$ [20]. As the heat sink temperature is increased, a further set of thermally activated absorption lines appear between 20 and 30 meV. By considering the binding energies in [19], it is possible to show that these transitions are between the $2p_0$, $2p_{\pm}$ energy levels.
ideal passive device with no current flowing. In a real device with current flowing and associated joule heating, the actual operating range is likely to be much narrower and it is unlikely such a device would have a maximum operating temperature above 40 K as in Fig. 12. This is quite consistent with the value quoted in [27]. The major problem with such impurity devices is the tradeoff between doping, current, gain, and free carrier losses. Initial estimates of the gain for electrically pumped impurity structures suggest that it is going to be difficult to overcome the waveguide losses with simple device structures.

VIII. CONCLUSION

In this paper, we have discussed several aspects pertinent to the realization of an electrically pumped silicon-based terahertz laser. Growth in a strained material system such as Si/SiGe has been discussed. It has been shown that it is possible to grow a low dislocation density, thick multiperiod Si/SiGe superlattice, with good layer thickness uniformity. Quantum mechanical tunneling between valance band states in p-type Si/SiGe RTD structures has been demonstrated, which is a prerequisite for the operation of a p-type Si/SiGe quantum cascade laser. Lifetime measurements in quantum wells with an energy spacing at terahertz frequencies have also been performed. These measurements show that the lifetime remains relatively constant over a large temperature range and provides strong evidence that a successful SiGe quantum cascade laser has the potential for operation over a wider temperature range than its III–V counterpart. The question of high vertical confinement in silicon has been addressed and it has been shown that buried silicide technology provides a possible route forward. Modeling suggests that modal overlap and waveguide losses comparable to GaAs terahertz quantum cascade lasers can be produced. Finally, the possibility of exploiting impurities such as boron and phosphorus to realize an electrically pumped terahertz silicon-based laser has been explored. Absorption measurements show that it is unlikely such an impurity-based device would have a maximum operating temperature above 40 K. The impressive results from optically pumped Si-impurity lasers demonstrate that lasers can be produced with these impurities although the additional free carrier losses from higher doping levels required to reduce resistive joule heating and increase gain make the realization of an electrically pumped silicon-impurity laser far more difficult than the optically pumped version.

In summary, all the necessary individual components to realize a terahertz quantum cascade laser are in place. The challenge, which remains, is to design and grow a quantum cascade active region that exhibits sufficient gain to provide lasing action.

REFERENCES


Zhang Suet (S’02), photograph and biography not available at the time of publication.

Anthony G. Cullis, photograph and biography not available at the time of publication.

Robert W. Kelsall, photograph and biography not available at the time of publication.

Carl R. Pidgeon, photograph and biography not available at the time of publication.

Zoran Ikonic was born in Belgrade, Yugoslavia, in 1956. He received the B.Sc., M.Sc., and Ph.D. degrees in electrical engineering from the University of Belgrade, Belgrade, Yugoslavia, in 1980, 1984, and 1987, respectively. During 1981–1999, he was with the Faculty of Electrical Engineering, University of Belgrade, where he was a Full Professor from 1998. Since 1999, he has been with the School of Electronic and Electrical Engineering, University of Leeds, Leeds, U.K. His current research interests include electronic structure, optical and transport properties of semiconductor nanostructures, and devices based upon them.

Pawel Murzyn, photograph and biography not available at the time of publication.

Ben Murdin, photograph and biography not available at the time of publication.

Paul Harrison (M’99–SM’99) received the B.Sc. degree from the University of Hull, Hull, U.K., in 1988, and the Ph.D. degree from the University of Newcastle-upon-Tyne, Newcastle-upon-Tyne, U.K., in 1991. He was a Postdoctoral Research Assistant at the University of Hull in 1995. He obtained a Fellowship at the University of Leeds, Leeds, U.K., and joined the Institute of Microwave and Photonics. He is currently Chair in quantum electronics and Head of the School of Electronic and Electrical Engineering, University of Leeds. He is the author of Quantum Wells, Wires and Dots (Wiley, 2005). He has been working on ways to adapt his theoretical and computational experience in semiconductor heterostructures to terahertz sources and detectors.

Mike Bain, photograph and biography not available at the time of publication.

Harry S. Gamble (M’96), photograph and biography not available at the time of publication.

Jing Zhang, photograph and biography not available at the time of publication.

Ming Zhao, photograph and biography not available at the time of publication.

David J. Norris, photograph and biography not available at the time of publication.

Wei-Xin Ni, photograph and biography not available at the time of publication.