Laser cooling of solids

We present an overview of solid-state optical refrigeration also known as laser cooling in solids by fluorescence upconversion. The idea of cooling a solid-state optical material by simply shining a laser beam onto it may sound counter intuitive but is rapidly becoming a promising technology for future cryocoolers. We chart the evolution of this science in rare-earth doped solids and semiconductors.
**Abstract** We present an overview of solid-state optical refrigeration also known as laser cooling in solids by fluorescence upconversion. The idea of cooling a solid-state optical material by simply shining a laser beam onto it may sound counter intuitive but is rapidly becoming a promising technology for future cryocoolers. We chart the evolution of this science in rare-earth doped solids and semiconductors.

Measured cooling efficiency as a function of the pump laser wavelength for a 1.2% Tm$^{3+}$ doped BaY$_2$F$_8$ crystal at room temperature. The solid line is calculated using the absorption spectrum [30].

© 2009 by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

---

**Laser cooling of solids**

*Mansoor Sheik-Bahae$^{1,*}$ and Richard I. Epstein$^{1,2}$*

$^1$ Optical Science and Engineering, Department of Physics & Astronomy, University of New Mexico, Albuquerque, NM, USA

$^2$ Los Alamos National Laboratory, Los Alamos, NM, USA

Received: 1 August 2008, Revised: 8 September 2008, Accepted: 10 September 2008

Published online: 27 October 2008

**Key words:** Solid-state laser cooling, optical refrigeration, anti-Stokes fluorescence, luminescence up-conversion, rare-earth doped solids, direct band-gap semiconductors, all-solid-state cryocooler, external quantum efficiency, GaAs, differential luminescence thermometry.

**PACS:** 32.80.Pj, 78.20.–e, 78.55.Cr, 78.66.Fd

---

**1. Basic concepts**

The term “laser cooling” is most often used in association with cooling and trapping of dilute gases of atoms and ions to extremely low temperatures. This area of science has progressed rapidly in the last two decades and has facilitated the observation of Bose-Einstein condensates and many related phenomena [1, 2]. It is surprising that nearly half a century before Doppler cooling of atoms was ever contemplated and more than thirty years before invention of the laser, German physicist Peter Pringsheim (Fig. 1) proposed cooling of solids by fluorescence up-conversion [3]. In the solid phase, atoms do not possess relative translational motion – their thermal energy is largely contained in the vibrational modes of the lattice. The physics of laser cooling of solids (or optical refrigeration) does have essential similarities to atom cooling: light quanta in the red tail of the absorption spectrum are absorbed from a monochromatic source followed by spontaneous emission of more energetic (blue-shifted) photons. In the case of solids, the extra energy is extracted from lattice phonons, the quanta of vibrational energy in which heat is contained. The removal of these phonons is equivalent to cooling the solid. This process has also been termed “anti-Stokes fluorescence” and “luminescence up-conversion” cooling.

Laser cooling of solids can be exploited to achieve an all-solid-state cryocooler [4–6] as conceptually depicted in Fig. 2. The advantages of compactness, no vibrations, no moving parts or fluids, high reliability, and no need for cryogenic fluids have motivated intensive research. Spaceborne infrared sensors are likely to be the first beneficiaries, with other applications requiring compact cryocooling rea-

* Corresponding author: e-mail: msb@unm.edu
In 1929, Peter Pringsheim suggested that solids could cool through anti-Stokes fluorescence in which a substance absorbs a photon and then emits one of greater energy. The energy diagram on the right shows one way this could occur. An atom with two broad levels is embedded in a transparent solid. The light source of frequency $h\nu$ excites atoms near the top of the ground state level to the bottom of the excited state. Radiative decays occurring after thermalization emit photons with average energy $h\nu > h\nu_f$.

Another potential application of laser cooling of solids is to eliminate heat production in high-power lasers. Even though laser emission is always accompanied by heat production, Bowman [8,9] realized that in some laser materials, the pump wavelength can be adjusted so that the spontaneous anti-Stokes fluorescence cooling compensates for the laser heating. Such thermally balanced laser would not suffer thermal defocusing or heat damage.

The process of optical refrigeration can occur only in special high purity materials (see section III) that have appropriately spaced energy levels and emit light with high quantum efficiency. To date, optical refrigeration research has been confined to glasses and crystals doped with rare-earth elements and direct-band semiconductors such as gallium arsenide. Laser cooling of rare-earth doped solids have been successfully demonstrated, while observation of net cooling in semiconductors has remained elusive. Fig. 1 schematically depicts the optical refrigeration processes for a two level system with vibrationally broadened ground and excited state manifolds. Photons from a low entropy light source (i.e. a laser) with energy $h\nu$ excite atoms from the top of the ground state to the bottom of the excited state. The excited atoms reach quasi-equilibrium with the lattice by absorbing phonons. Spontaneous emission (fluorescence) follows with a mean photon energy $h\nu_f$ that is higher than that of the absorbed photon. This process has also been called anti-Stokes fluorescence. There were initial concerns that the second law of thermodynamics might be violated until Landau clarified the issue in 1946 by assigning an entropy to the radiation [10].

In the aforementioned simple model, the interaction rate between electrons and phonons within each manifold is assumed to be far faster than the spontaneous emission rate, which is valid for a broad range of materials and temperatures. The cooling efficiency or fractional cooling being cooled, the load, by a thermal link; see Fig. 2. This link siphons heat from the load while preventing the waste fluorescence from hitting the load and heating it.

Figure 2 (online color at: www.lpr-journal.org) Schematic of an optical refrigeration system. Pump light is efficiently generated by a semiconductor diode laser. The laser light enters the cooler through a pinhole in one mirror and is trapped by the mirrors until it is absorbed. Isotropic fluorescence escapes the cooler element and is absorbed by the vacuum casing. A sensor or other load is connected in the shadow region of the second mirror. Fig. 2 has been reproduced from [6].
energy for each photon absorbed is
\[ \eta_e = \frac{h\nu_f - h\nu}{h\nu} = \frac{\lambda}{\lambda_f} - 1 \]  
(1a)

where \( \lambda = c/\nu \) is the wavelength. The invention of the laser in 1960 prompted several unsuccessful attempts to observe laser cooling of solids [11–13]. In 1995, net cooling was first achieved by workers at Los Alamos National Laboratory [14]. Two technical challenges were addressed and overcome in these experiments. The Los Alamos researchers had to have a system in which i) the vast majority of optical excitations recombine radiatively and ii) there is a minimal amount of parasitic heating due to unwanted impurities. Both of these critical engineering issues are ignored in the idealized situation described by Eq. (1a), but are key to experimental success.

It is also important that spontaneously emitted photons escape the cooling material without being trapped and re-absorbed, which would effectively inhibit spontaneous emission [15,16]. This is a critical issue for high index semiconductors where total internal reflection can cause strong radiation trapping. In the absence of radiation trapping, the fraction of atoms that decay to the ground state by the desired radiative process is known as the quantum efficiency, \( \eta_q = W_{rad}/(W_{rad} + W_{nr}) \) where \( W_{rad} \) and \( W_{nr} \) are radiative and nonradiative decay rates, respectively. Including a fluorescence escape efficiency \( \eta_{ext} \) defines an external quantum efficiency (EQE): \( \eta_{ext} = \eta_e W_{rad}/(\eta_e W_{rad} + W_{nr}) \), which assumes the fluorescence is re-absorbed within the excitation volume (see section IV). This describes the efficiency by which a photo-excited atom decays into an escaped fluorescence photon in free space. In a similar fashion, an absorption efficiency \( \eta_{abs} = \alpha_e/(\alpha_e + \alpha_b) \) is defined to account for the fraction of pump laser photons that are engaged in cooling [6,17]. Here \( \alpha_e \) is the resonant absorption coefficient and \( \alpha_b \) is the unwanted parasitic (background) absorption coefficient. As will be derived in sections II and IV, the combination of all these effects re-defines the cooling efficiency as:

\[ \eta_c = \eta_{ext} \eta_{abs} \frac{\lambda}{\lambda_f} - 1 \]  
(1b)

where the product \( \eta_{ext} \eta_{abs} \) indicates the efficiency of converting an absorbed laser photon to an escaped fluorescence photon. Note that \( \eta_{abs} \) is frequency-dependent and falls off rapidly below photon energy \( h\nu_f - k_B T \) where \( k_B \) is the Boltzmann constant and \( T \) is the lattice temperature. At pump photon energies much more than \( k_B T \) below \( h\nu_f \), \( \eta_{abs} \) is too small to make \( \eta_c > 0 \) and laser cooling is unattainable. The above analysis defines the approximate condition needed for laser cooling [6,17]:

\[ \eta_{ext} \eta_{abs} > 1 - \frac{k_B T}{h\nu_f} \]  
(2)

This relation quantifies the needed efficiencies: cooling a material from room temperature with a nominal energy gap (pump photon) of 1 eV from room temperature demands that \( \eta_{ext} \eta_{abs} > 97\% \). Although suitable lasers were available in the early 1960’s, more than three decades of progress in material growth were needed to satisfy this condition.

2. The 4-level model for optical refrigeration

Consider the 4-level system of Fig. 3 in which the ground state manifold consists of two closely spaced levels of \(|0\rangle \) and \(|1\rangle \) with an energy separation \( \delta E_g \). The excited manifold consists of two states \(|3\rangle \) and \(|2\rangle \) with an energy separation \( \delta E_u \). Laser excitation at \( h\nu \) is tuned to be in resonance with the \(|1\rangle – |2\rangle \) transition as shown by the solid red arrow. The double-line arrows depict the spontaneous emission transitions from the upper level to the ground states with a rate of \( W_{rad} \); this rate is assumed to be the same for all four transitions. The nonradiative decay rates (indicated by the dotted lines) are also assumed to be equal and given by \( W_{nr} \). The population in each manifold reaches a quasi-thermal equilibrium via an electron-phonon interaction rate given by \( w_1 \) and \( w_2 \) for lower and upper states, respectively.

The rate equations governing the density populations \( N_0, N_1, N_2, \) and \( N_3 \) are:

\[ \frac{dN_1}{dt} = -\sigma_{12} \left( N_1 - \frac{g_1}{g_2} N_2 \right) \frac{I}{h\nu} + \frac{R}{2} (N_2 + N_3) - w_1 \left( N_1 - \frac{g_1}{g_0} N_0 e^{-\delta E_s/k_B T} \right), \]  
(3a)

\[ \frac{dN_2}{dt} = \sigma_{12} \left( N_1 - \frac{g_1}{g_2} N_2 \right) \frac{I}{h\nu} - RN_2 + w_2 \left( N_3 - \frac{g_3}{g_2} N_2 e^{-\delta E_u/k_B T} \right), \]  
(3b)

\[ 
\begin{align*}
|0\rangle &\quad \delta E_g \quad |1\rangle \\
|1\rangle &\quad \delta E_u \quad |2\rangle \\
|2\rangle &\quad W_{rad} \quad |3\rangle \\
|3\rangle &\quad W_{nr} \quad |0\rangle 
\end{align*}
\]

*Figure 3* (online color at: www.lpr-journal.org) The four-level energy model for optical refrigeration consisting of two pairs of closely spaced levels: \(|0\rangle \) and \(|1\rangle \) in the ground state and \(|2\rangle \) and \(|3\rangle \) in the excited state manifolds.
We have also included a term that represents parasitic absorption, which will be revisited when discussing semiconductors in section IV. Second, Eq. (6) shows that the mean fluorescence wavelength combination with the reduction of the resonant absorption reduces the cooling efficiency. At temperature $T = T_m$ the cooling stops (i.e. $\eta_c(T_m) = 0$). This minimum achievable temperature ($T_m$) can be lowered by reducing the background absorption (higher purity), increasing the quantum efficiency, and enhancing the resonant absorption (e.g. choosing a material with a narrow ground state manifold). The effect of fluorescence trapping and its consequent re-absorption by both resonant and parasitic processes will further diminish the quantum efficiency. We will discuss this in detail when we analyze laser cooling in semiconductors where total internal reflection leads to substantial trapping.

### 3. Cooling rare-earth-doped solids.

The advantages of rare-earth (RE) doped solids for laser cooling had been foreseen for decades. Kastler (1950 [11]) and Yatsiv (1961 [13]) suggested these materials could be used for optical cooling. The key optical transitions in RE-doped ions involve 4f electrons that are shielded by host and YAG crystals [27], YLF [32, 33]). Laser-induced cooling has since been demonstrated in 1995 with the ytterbium-doped fluorozirconate glass ZBLANP:Yb$^{3+}$ [14]. Laser-induced cooling has since been observed in a range of glasses and crystals doped with Yb$^{3+}$ (ZBLANP [19–22], ZBLAN [23, 24], CNB:Zn [9, 25] BIG [25, 26], KGeO$_3$ [9], KY(WO$_4$)$_2$ [9], YAG [27], Y$_2$SiO$_5$ [27], K$_2$CO$_3$ [25, 28], BaY$_2$F$_8$ [29–31], and YLF [32, 33]).

Fig. 4 shows the cooling and heating of a sample of Yb$^{3+}$ doped ZBLANP for a range of pump wavelengths [5].

\[
\frac{dN_3}{dt} = -RN_3 - w_2 \left( N_3 - \frac{g_3}{g_2} N_2 e^{-\delta E_u/k_B T} \right)
\]

\[
N_0 + N_1 + N_2 + N_3 = N_t
\]

where $R = 2W_{rad} + 2W_{nr}$ is the total upper state decay rate, $\sigma_{12}$ is the absorption cross section associated with the $|1\rangle \leftrightarrow |2\rangle$ transition, $I$ is the incident laser irradiance and the $g_i$ terms represent degeneracy factors for each level. The weighting factor in the electron-phonon interaction terms ($w_1$ and $w_2$) maintains the Boltzmann distribution among each manifold at quasi equilibrium. The net power density deposited in the system is the difference between absorbed and out-radiated contributions:

\[
P_{net} = \sigma_{12} N_1 \left( 1 - \frac{g_1 N_2}{g_2 N_1} \right) I
\]

\[
- W_{rad} \left[ N_2 (E_{21} + E_{20}) + N_3 (E_{31} + E_{30}) \right]
\]

\[
+ \alpha_b I,
\]

(4)

where the first term is the laser excitation $(|1\rangle \leftrightarrow |2\rangle$ transition) and second term includes the spontaneous emission terms from levels $|2\rangle$ and $|3\rangle$ with their respective photon energies. We have also included a term that represents parasitic absorption of the pump laser with an absorption coefficient of $\alpha_b$. It is straightforward to evaluate the steady-state solution to the above rate equations by setting the time derivatives to zero. To emphasize certain features, we ignore saturation and assume unity degeneracy for all levels. The net power density is then obtained as:

\[
P_{net} = \alpha I \left( 1 - \eta_q \frac{h\nu_f}{h\nu} \right) + \alpha_b I,
\]

(5)

where $\eta_q = (1 + W_{nr}/W_{rad})^{-1}$ is the (internal) quantum efficiency and $h\nu_f$ denotes the mean fluorescence energy of the four-level system given by:

\[
h\nu_f = h\nu + \frac{\delta E_g}{2} + \frac{\delta E_u}{1 + (1 + R/w_2) e^{\delta E_u/k_B T}}
\]

The ground state resonant absorption $\sigma$ is given by:

\[
\sigma = \alpha = \sigma_{12} N_t \left( 1 + e^{\delta E_u/k_B T} \right)^{-1}
\]

(7)

Despite its simplicity, the four level model reveals essential features of solid-state optical refrigeration: First, Eq. (7) exhibits diminishing pump absorption due to thermal depletion of the top ground state at low temperatures, $k_B T < \delta E_g$. This implies that the width of the ground-state manifold ($\delta E_g$) must be narrow to achieve cooling at low temperatures with reasonable efficiency. This issue will be revisited when discussing semiconductors in section IV. Second, Eq. (6) shows that the mean fluorescence photon energy is red shifted at low temperatures, which further lowers the cooling efficiency. This shift would be enhanced if the electron-phonon interaction rate ($w_2$) is smaller than the upper state recombination rate ($R$). This means that if $w_2 < R$, decay of the excited state can occur before thermalization with the lattice, which results in no fluorescence upconversion and no cooling [18]. This extreme limit of cold electron recombination is an issue for semiconductors at very low temperatures where the electrons interact with the lattice primarily by slowly relaxing acoustic phonon [15].

Dividing the Eq. (5) by the total absorbed power density $P_{abs} = (\alpha + \alpha_b)I$ gives the cooling efficiency $\eta_c = -P_{net}/P_{abs}$:

\[
\eta_c = \eta_q \frac{h\nu_f}{h\nu} - 1,
\]

(8)

which is similar to Eq. (1.b) not including the luminescence trapping. The most useful feature of the 4-level model is its description of the temperature-dependence of the cooling in a physically transparent manner. As the temperature is lowered, red-shifting of mean fluorescence wavelength combined with the reduction of the resonant absorption reduces the cooling efficiency. At temperature $T = T_m$ the cooling stops (i.e. $\eta_c(T_m) = 0$). This minimum achievable temperature ($T_m$) can be lowered by reducing the background absorption (higher purity), increasing the quantum efficiency, and enhancing the resonant absorption (e.g. choosing a material with a narrow ground state manifold). The effect of fluorescence trapping and its consequent re-absorption by both resonant and parasitic processes will further diminish the quantum efficiency. We will discuss this in detail when we analyze laser cooling in semiconductors where total internal reflection leads to substantial trapping.
Figure 4 (online color at: www.lpr-journal.org) The temperature change (normalized to incident power) in ytterbium-doped ZBLANP glass as a function of pump wavelength. When the pump wavelength is considerably longer than the mean wavelength of the fluorescence $\lambda_F$ (vertical dashed line), the escaping light carries more energy than the absorbed laser light and the glass cools. Heating at wavelengths greater than $\lambda_F$ is due to imperfect quantum efficiency of the fluorescence and non-resonant light absorption [5].

For wavelengths shorter than the mean fluorescence wavelength $\lambda_F$ (vertical dashed line) the sample heats because of the Stokes shift as well as by non-radiative processes. At longer wavelengths, anti-Stokes cooling dominates and cooling as large as 25 K per watt of absorbed laser power is measured. At still longer wavelengths, absorption by impurities or imperfections dominates, and the sample heats.

In 2000, laser cooling in Tm$^{3+}$ doped ZBLANP was reported at $\lambda \sim 1.9$ µm [34]. The significance of this result was two-fold: First, it verified the scaling law of Eq. (1.a,b) by demonstrating nearly a factor of two enhancement in the cooling efficiency compared to Yb-doped systems. The enhancement scales as the ratio of the corresponding cooling transition wavelengths. Second, it was the first demonstration of laser cooling in the presence of excited state absorption. A record cooling power of $\sim 73$ mW was obtained in this material by employing a multipass geometry [35]. More recently, cooling of Er$^{3+}$ doped glass (CNBZn) and crystal (K$\text{Pb}_2\text{Cl}_5$) at $\lambda \sim 0.870$ µm were reported by a Spanish group [36]. It is interesting to note that the cooling transition used in these experiments is between the ground state and the fourth excited state ($^4I_{9/2}$) of Er$^{3+}$, not the first excited state as illustrated in Fig. 1. High-energy transitions have lower cooling efficiency (Eq. 1) but potentially higher quantum efficiency due to their low nonradiative decay rates to the ground state. The presence of higher excited states in Er$^{3+}$ may prove advantageous since the energy upconversion transitions (i.e. at the cooling wavelengths of the main transition) are endothermic with a high quantum efficiency [36, 37]. This is also the case with cooling of Tm$^{3+}$ [34].

The initial proof-of-principle experiments in ZBLANP:Yb$^{3+}$ achieved cooling by an amount 0.3 K below ambient temperature [14]. The LANL group has since cooled ZBLANP:Yb$^{3+}$ to 208 K starting from room temperature [22] as shown in Fig. 5. Although progress is being made, optical refrigerators need to be more efficient and operate at lower temperatures, below about 170 K, to be competitive with other solid-state coolers such as thermoelectric (Peltier) devices. Several studies have shown that ytterbium- or thulium-doped solids should be capable of providing efficient cooling at temperatures well below 100 K [4, 27, 38].

Figure 5 (online color at: www.lpr-journal.org) Panel (a) shows record cooling to 208 K with ZBLANP:Yb$^{3+}$. The temperatures are measured with thermocouples on the sample and chamber; the internal temperature of the glass is inferred [22]. Panel (b) compares the cooling efficiencies of available thermoelectric coolers (TECs) with ZBLANP:Yb$^{3+}$-based optical refrigerators. Devices based on materials with low parasitic heating will outperform TECs below 200. Coolers made from current materials are less efficient than TECs at all temperatures [39]. Fig. 5 (a) has been reproduced from [22].
There are several factors that limit the cooling of rare-earth-doped solids in available materials. The most significant factor is the choice of laser cooling medium. The ideal cooling efficiency (Eq. 1) shows that there is an advantage of pumping with lower energy photons. This increased efficiency was part of the motivation for investigating thulium-doped cooling materials, since their ground- and excited state manifolds are separated by about 0.6 eV compared to 1.2 eV in ytterbium-doped solids. There are obstacles, however, in going to longer wavelengths. First is the more limited choices of pump lasers since there are fewer available near 0.6 eV than near 1.2 eV. While not a fundamental consideration, it needs to be kept in mind for near-term commercialization. A second and more general reason involves the ratio of radiative to non-radiative relaxation decays. The rate of non-radiative, heat-producing, multi-phonon decay decreases exponentially with the separation between the ground- and first excited-state manifolds; this is the well-know energy-gap law. In practical terms, this means that because of the relatively large energy of the excited level in ytterbium-doped materials, non-radiative decays do not significantly decrease the quantum efficiency. For pure thulium-doped material, non-radiative decay can overwhelm anti-Stokes cooling, depending on the properties of the host material. For materials with low maximum phonon energies, such as ZBLANP and other fluoride hosts, the non-radiative decays are relatively slow. Many thulium-doped oxide crystals and glasses have rapid nonradiative decay rates that prevent laser cooling.

Another consideration in the choice of cooling medium is the width of the ground state manifold. According to Boltzmann statistics, lower energy levels in the manifold are more populated than higher ones. As the temperature falls and $k_B T$ becomes small compared to the energy width of the ground-state manifold, the upper levels become depopulated leading to increased transparency at lower frequencies; this effect is illustrated in the four-level system discussed above. The net effect is that at low temperatures the numerator of Eq. (2) becomes small and cooling efficiency goes to zero; see Eq. (1b). The width of the ground state manifold is typically the result of crystal field splitting and depends on both the dopant ion and the host material. By choosing ions and a host that give narrow ground-state manifolds, the material can cool to lower temperatures before the low-frequency transparency condition sets in.

For the material systems studied so far, cooling is not limited by the reasons outlined above. It is most likely hindered by parasitic heating in the bulk of the cooling material or on its surface. As one can see in Fig. 5b, the cooling efficiencies of currently available ZBLANP:Yb$^{3+}$ are far below that for an ideal material with no parasitic heating. One important source of heating in this material is quenching of excited ytterbium ions by impurities such as iron and copper. The radiative decay time of an excited Yb$^{3+}$ ion is about 1 ms. During this time, the excitation encounters an impurity atom, the energy can be transferred to this atom and rapidly converted into heat. A detailed study by Hehlen et al. [39] found that the ideal cooling efficiency can be approached when concentration of impurities such as Cu$^{2+}$ is less than 0.01 ppm and that for Fe$^{2+}$ is below 0.1 ppm; see Fig. 6.

An additional source of parasitic heating is absorption in the mirrors that trap the pump radiation in the cooling element. In the LANL experiments, the cooling glass has a pair of high-reflectivity mirrors deposited on two surfaces, as depicted in Fig. 2. Pump light is reflected multiple times by each mirror, so that that even relatively low absorption of 0.0001 per surface produces significant heating. Depositing higher quality dielectric mirror may obviate this problem. An alternative approach is to avoid dielectric mirrors all together and exploit the total-internal reflection inside the
4. Prospects for laser cooling in semiconductors

Researchers have examined other condensed matter systems beyond RE-doped materials, with an emphasis on semiconductors [17, 43–46]. Semiconductor coolers provide more efficient pump light absorption, the potential of much lower temperatures, and the opportunity for direct integration into electronic and photonic devices. These materials provide their own set of engineering challenges, however, and no net cooling has been observed yet. The essential difference between semiconductors and RE-doped materials is in their cooling cycles. In the latter, the cooling transition occurs in localized donor ions within the host material while the former involves transition between extended valence and conduction bands of a direct gap semiconductor (see Fig. 7a). Indistinguishable charge carriers in Fermi-Dirac distributions may allow semiconductors to get much colder than RE materials. The highest energy levels of the ground state manifold in the RE-doped systems become less populated as the temperature is lowered, due to Boltzmann statistics. The cooling cycle becomes ineffective when the Boltzmann constant times the lattice temperature becomes comparable to the width of the ground state (see previous section describing the 4-level model). This sets a limit of $T \sim 100$ K for most existing RE-doped systems.

No such limitation exists in pure (undoped) semiconductors – temperatures as low as 10 K may be achievable [47].

Semiconductors should achieve higher cooling power density compared to RE-materials. The maximum cooling power density (rate of heat removal) is $\approx N \times k_B T / \tau_r$, where $N$ is the photo-excited electron (-hole) density and $\tau_r$ is the radiative recombination time. In semiconductors the optimal density $N$ is limited due to many-body processes and does not exceed that of moderately doped RE systems. We can gain 5–6 orders of magnitude in cooling power density because the radiative recombination rates in semiconductors are much faster than in RE ions.

Laser cooling of semiconductors has been examined theoretically [15, 44, 45, 47–52] as well as in experimental studies [46, 53–56]. A feasibility study by the authors outlined the conditions for net cooling based on fundamental material properties and light management [15]. Researchers at the University of Arizona [47, 50] studied luminescence upconversion in the presence of partially ionized excitons, which must be understood to attain temperatures approaching 10 K. The role of bandtail states [52], the possible enhancement of laser cooling by including the effects of photon density of states as well as novel luminescence coupling schemes based on surface plasmon polaritons [57, 58] were recently introduced by Khurgin at Johns Hopkins University. Here, we expand on the basic model of [15] and present the theoretical foundation of laser cooling in semiconductor structures with an arbitrary external efficiency. This treatment accounts for the luminescence red-shift due to re-absorption, the effect parasitic absorption of the pump, luminescence power, and band-blocking effects. We then discuss the latest experimental results attempting to make the first observation of laser cooling in a semiconductor material.

We consider an intrinsic (undoped) semiconductor system uniformly irradiated with a laser light at photon energy $h\nu$. Furthermore, we assume that only a fraction $\eta_e$ of the total luminescence can escape the material while the remaining fraction $(1-\eta_e)$ is trapped and re-cycled, thus contributing to carrier generation. For now, we will ignore the parasitic absorption of luminescence but will later consider its implications. For a given temperature, the rate equations

$$\frac{dN}{dt} = -N \frac{\tau_r}{\tau},$$

$$\frac{dT}{dt} = \frac{1}{k_B} \left( \frac{N \tau}{\tau_r} \right) h\nu.$$
for the electron-hole pair density (N) is given by [15]:

\[
\frac{dN}{dt} = \frac{\alpha I}{h\nu} - AN - BN^2 - CN^3 + (1 - \eta_e)BN^2 \tag{9}
\]

Here \(\alpha(\nu,N)\) is the interband absorption coefficient that includes many-body and blocking factors. The recombination process consists of nonradiative (AN), radiative (BN\(^2\)), and Auger (CN\(^3\)) rates. All the above coefficients are temperature dependent. The last term represents the increase in \(N\) from the re-absorption of the luminescence that does not escape, assuming the re-absorption occurs within the laser excitation volume. The density-dependence of \(\alpha\) results from both Coulomb screening and band-blocking (saturation) effects. The latter can be approximated by a blocking factor such that [59, 60]:

\[
\alpha(N, h\nu) = \alpha_0(N, h\nu)\{f_v - f_e\}, \tag{10}
\]

where \(\alpha_0\) denotes the unsaturated absorption coefficient. The strongly density-dependent blocking factor in the brackets [61] contains Fermi-Dirac distribution functions for the valence \((f_v)\) and conduction \((f_e)\) bands.

Under steady-state conditions, Eq. (9) can be rewritten as

\[
0 = \frac{\alpha(\nu, N)}{h\nu}I - AN - \eta_eBN^2 - CN^3. \tag{11}
\]

This indicates that the fluorescence trapping effectively inhibits the spontaneous emission as it appears through \(\eta_eB\) only. This result has also been shown previously by Asbeck [16]. It is important to note that \(\eta_e\) is itself an averaged quantity over the entire luminescence spectrum.

Next \(S(\nu)\) is the geometry-dependent escape probability of photons with energy \(h\nu\) and \(R(\nu)\) is the luminescence spectral density that is related to the absorption coefficient through reciprocity using a “non-equilibrium” van Roosbroeck-Shockley relation (also known as Kubo-Martin-Schwinger (KMS) relation [59, 62]:

\[
R(\nu, N) = \frac{8\pi\nu^2}{c^2}\alpha(\nu, N)\frac{f_e(1 - f_v)}{f_v - f_e}, \tag{13}
\]

where \(c\) is the speed of light and \(n\) is the index of refraction. Note that the radiative recombination coefficient \(B\) is obtained by \(BN^2 = \int R(\nu)d\nu\) which results in a negligible dependence of \(B\) on \(N\) at the carrier densities of interest. The net power density that is deposited in the semiconductor is the difference between the power absorbed from the laser \(P_{\text{abs}}\) and that of the luminescence that escapes \(P_e\):

\[
P_{\text{net}} = P_{\text{abs}} - P_e = [\alpha I + \Delta P] - [\eta_eBN^2h\nu_f], \tag{14}
\]

where the absorbed power density includes the resonant absorption \(\alpha I\) and a term \(\Delta P\) that accounts for the undesirable effects such as free-carrier absorption and other parasitic absorptive processes. The second term is the escaped luminescence power density at a mean luminescence energy \(h\nu_f\) defined as

\[
h\nu_f = \frac{\int S(\nu)R(\nu)h\nu d\nu}{\int S(\nu)R(\nu)d\nu}. \tag{15}
\]

Note that the escaped mean luminescence energy can deviate (i.e. redshift) from its internal value (S=1) depending on the thickness or photon recycling conditions. With the aid of Eq. (9), we rewrite Eq. (14) as:

\[
P_{\text{net}} = \eta_eBN^2(h\nu - h\nu_f) + ANh\nu + CN^3h\nu + \Delta P. \tag{16}
\]

Eq. (16) rigorously describes laser cooling of a semiconductor in a compact and simple form. It accounts for the practical considerations of luminescence trapping by introducing an inhibited radiative recombination \((\eta_eB)\) and a shifted mean photon energy \(h\nu_f\) for the escaped luminescence. For high external efficiency systems where \(S(\nu) = 1\), Eq. (16) approaches that described in the literature with \(\eta_e = 1\) and \(\nu_f = \nu_t\) with \(\nu_t\) denoting the mean fluorescence energy produced internally in the semiconductor [44-46]. Eq. (16) indicates that laser cooling occurs when \(P_{\text{net}} < 0\), requiring a dominant contribution from the radiative recombination with \(h\nu < h\nu_f\). The cooling efficiency \(\eta_e\) is defined as the ratio of cooling power density \(P_c := -P_{\text{net}}\) to the absorbed laser power density \(P_{\text{abs}} = \alpha I + \Delta P\). With the aid of Eq. (11), this efficiency can be expressed as

\[
\eta_e = -\frac{\eta_eBN^2(h\nu - h\nu_f) + ANh\nu + CN^3h\nu + \Delta P}{\eta_eBN^2h\nu + ANh\nu + CN^3h\nu + \Delta P}. \tag{17}
\]

Ignoring the \(\Delta P\) contributions for the moment, \(\eta_e\) can be written more simply as:

\[
\eta_e = \eta_{\text{ext}}\frac{\nu_f}{\nu} - 1, \tag{18}
\]

where \(\eta_{\text{ext}}\) describes the external quantum efficiency (or EQE):

\[
\eta_{\text{ext}} = \frac{\eta_qBN^2}{AN + \eta_eBN^2 + CN^3} \approx \eta_q^{1/\eta_e}, \tag{19}
\]

with \(\eta_q = BN^2/(AN + BN^2 + CN^3)\) denoting the internal quantum efficiency [46, 63] as also defined more generally following Eq. (5). The approximate equality in Eq. (19) is valid only for \(\eta_{\text{ext}}\) near unity (> 0.9). One simple consequence of Eq. (19) is that there is an optimum carrier density \(N_{\text{op}} = (A/C)^{1/2}\) at which \(\eta_{\text{ext}}\) reaches a maximum:

\[
\eta_{\text{max}}^{\text{ext}} = 1 - \frac{2\sqrt{AC}}{\eta_eB} \tag{20}
\]

Including background parasitic absorption \((\Delta P = \alpha I)\), results in more general form of cooling efficiency:

\[
\eta_e = \eta_{\text{abs}}\eta_{\text{ext}}\frac{\nu_f}{\nu} - 1, \tag{21}
\]
where the absorption efficiency $\eta_{\text{abs}}$ is the fraction of the total absorbed photons from the pump laser that is consumed by the resonant absorption in the cooling region:

$$\eta_{\text{abs}} = \frac{\alpha(\nu)}{\alpha(\nu) + \alpha_b}.$$  (22)

and $\alpha_b$ is assumed to be constant in the vicinity of the band-edge region.

If the pump laser suffers parasitic absorption, so will the luminescence since their frequencies are very close. We now examine the parasitic absorption problem and its effect on the cooling efficiency by revisiting Eq. (9). A small fraction $\varepsilon_f$ of the trapped luminescence is absorbed parasitically and the remaining part $(1-\varepsilon_f)$ is recycled through interband absorption thus contributing to the carrier generation. Eq. (9) is rewritten as:

$$\frac{dN}{dt} = \frac{\alpha I}{\hbar \nu} - AN - BN^2 - CN^3 + (1 - \eta_e)(1-\varepsilon_f)BN^2.$$  (23)

Note that $1-\varepsilon_f = \frac{\tilde{\alpha}_f}{\alpha_f + \alpha_b} \approx 1 - \frac{\alpha_b}{\alpha_f}$ where $\tilde{\alpha}_f \approx \alpha(\nu_f)$ is the interband absorption of the luminescence averaged over its spectrum. Following the same analysis leading to Eq. (21), we obtain the modified cooling efficiency:

$$\eta_e = \bar{\eta}_{\text{ext}} \eta_{\text{abs}} \frac{\tilde{\alpha}_f}{\nu} - 1,$$  (24)

with a modified EQE ($\bar{\eta}_{\text{ext}}$) that is reduced from its ideal value in the high purity ($\alpha_b \approx 0$) approximation to:

$$\bar{\eta}_{\text{ext}} = \frac{1}{1 + \eta_{\text{ext}} \varepsilon_f (1 - \eta_e)/\eta_e}.$$  (25)

This expression is useful for setting an upper bound on the existing intrinsic background absorption of GaAs/InGaP heterostructures. This will be discussed in detail below. Parasitic luminescence absorption is not important in the analysis of photo-carrier density and incident laser irradiance so it is ignored for the moment.

The roots of Eq. (16) define the carrier density range within which net cooling can be observed provided that $\eta_e B (h\nu - \hbar \nu f) \geq 2\hbar \nu \sqrt{AC}$. The equality defines the break-even condition: heating and cooling are in exact balance. At high quantum efficiency, radiative recombination dominates (i.e. $\eta_e B/C \gg N \gg A/\eta_e B$) allowing one to obtain the corresponding laser irradiance from Eq. (11) with the assumption of no band-blocking. We can account for parasitic absorption of the pump by taking $\Delta P = \alpha_b I + \sigma_{\text{fca}} N I$ where $\alpha_b$ denotes an effective background parasitic absorption and $\sigma_{\text{fca}}$ is the free-carrier absorption cross section. In this case, net cooling can occur within an irradiance range of $I_1 < I < I_2$ where $I_{1,2} = (h\nu_B/\alpha(\nu)) N_{1,2}^2$ and

$$N_{1,2} = \left( \frac{h\nu_f - h\nu}{\hbar \nu} - \frac{\alpha_b}{\alpha(\nu)} \right) \eta_e B \left( 1 \mp \sqrt{1 - \frac{\nu}{A_0}} \right).$$  (26)

Here $C' = C + \sigma_{\text{fca}} \eta_e B/\alpha(\nu)$, and

$$A_0 = \left( \frac{h\nu_f - h\nu}{\hbar \nu} - \frac{\alpha_b}{\alpha(\nu)} \right)^2 \left( \eta_e B \right)^2 / 4 C'.$$  (27)

is the break-even (maximum allowable) nonradiative decay rate for a given excitation energy $h\nu$. Free carrier absorption appears as an enhancement of the Auger process. The parameters $B$ and $C$ are fundamental properties of a semiconductor and have been calculated and measured extensively for various bulk and quantum-confined structures [59, 60, 63, 64]. The reported values for these coefficients, however, vary considerably. In bulk GaAs, for example, the published values are $2 \times 10^{-16} < B < 7 \times 10^{-16} \text{m}^3/\text{s}$ and $1 \times 10^{-42} < C < 7 \times 10^{-42} \text{m}^6/\text{s}$ [64]. These variations are primarily due to experimental uncertainties. We assume average values of $B = 4 \times 10^{-16} \text{m}^3/\text{s}$ and $C = 4 \times 10^{-42} \text{m}^6/\text{s}$, while ignoring the effects of background and free carrier absorption. These assumptions allow us to gain insight into the feasibility and requirements for achieving net laser cooling. It should be noted that the theoretical values for these parameters, for different models vary almost within the same range as the experimental results. For the simple two-band model used here, $B \approx 5 \times 10^{-16} \text{m}^3/\text{s}$ [65].

Using Eq. (27), we plot in Fig. 8 the break-even radiative lifetime $\tau_{\text{tr}}^0 = A_0^{-1}$ as a function of $\eta_e$ assuming $h\nu_f - h\nu = k_B T$ with $h\nu_f$ corresponding to $\lambda_f \approx 860 \text{ nm}$ at room temperature. The orange area under the curve is the unwanted (heating) zone. Eq. (27) also suggests that increasing the quantum efficiency $\eta_q$ by decreasing the incident photon energy (e.g. at $h\nu_f - h\nu > k_B T$) relaxes this requirement. Interband absorption drops rapidly as the excitation moves further in the Urbach tail and one may no longer ignore background and free carrier absorption. Recently, it was found that free carrier absorption (FCA) at
the band edge wavelengths is much smaller than previously expected [66, 67]. For GaAs, $\sigma_{FCA} \approx 10^{-24} \text{ m}^2$ [66, 67], which requires $\alpha(\nu) \approx 10^4 \text{ m}^{-1}$ to ensure that free carrier losses are negligible (i.e. $C' \approx C$). This requirement is satisfied even at $\lambda = 890 \text{ nm}$ (corresponding to $h\nu_T - h\nu \approx 2k_B T$) where $\alpha(\nu) \approx 10^4 \text{ m}^{-1}$. We conclude that FCA does not pose a limitation on laser cooling.

We can categorize the possible sources and locations of the parasitic background absorption $\alpha_b$ into three regions: (a) active or core material, (b) cladding layers of the heterostructure, and (c) the substrate. It is also implicit that $\alpha_b$ in Eq. (27) is scaled such that for cases (b) and (c), the actual background absorption coefficient $\tilde{\alpha}_b = \alpha_b \times (d/L)$, where d and L are the thicknesses of the loss and active media (if different) respectively.

While situations (b) and (c) can be controlled experimentally by varying the barrier thickness or using high purity substrate respectively, the parasitic absorption from the cooling layer itself presents the most difficult engineering obstacle. This limitation is revisited in the next section where experiments on laser cooling with GaAs will be analyzed.

It is also instructive to show an alternative and compact way of expressing the cooling condition. With laser excitation at $h\nu < h\tilde{\nu}_f$, the cooling condition defined by EQE reduces to:

$$\eta_{ext} > \frac{\nu}{\tilde{\nu}_f} + \frac{\alpha_b}{\alpha(\nu)}.$$  

Including the parasitic absorption of luminescence lets us replace $\eta_{ext}$ by $\tilde{\eta}_{ext}$ using Eq. (28) to give a more general condition

$$\tilde{\eta}_{ext} = \eta_{ext} - \frac{\alpha_b(1 - \eta_{ext})}{\tilde{\nu}_f \eta_{ext}} > \frac{\nu}{\tilde{\nu}_f} + \frac{\alpha_b}{\alpha(\nu)}.$$  

The above inequality emphasizes the critical role of $\alpha_b$ in achieving net laser cooling. The quantity $\tilde{\eta}_{ext}$ can be measured accurately, so Eq. (28) defines the minimum value of EQE for a given background absorption provided $\alpha(\nu)$ is known. The absorption $\alpha(\nu)$ drops sharply for energies considerably below the bandgap, which means this inequality may never be satisfied for any wavelength if $\alpha_b$ is too large. To quantify this argument, we need to know the band-tail absorption accurately. The nature of the band tail states and their dependence on the impurity type and concentration make the reported experimental values very sample-specific. Most theoretical calculations are accurate only for above and near the bandgap wavelengths. It is best to approach the problem experimentally with absorption and luminescence data that allow accurate estimates of the required EQE using Eq. (9). Starting with the measured low-density photoluminescence (PL) spectrum on a high quality sample, we obtain absorption spectra $\alpha(\nu)$ using the KMS relations of Eq. (13). The low density approximation reduces the occupation factor to a simple Boltzmann factor, $\exp(-h\nu/k_B T)$, where we ignore possible band-filling (saturation) effects in the band-tail. Using Eq. (9), the minimum required $\eta_{ext}$ can be estimated as a function of $h\nu$ for various values of $\alpha_b$, as depicted in Fig. 9. Here we assume an extraction efficiency $\eta_e = 0.1$ which is typical of GaAs on a ZnS dome structure [15, 68].

Fig. 9 indicates that the required EQE for cooling becomes more demanding as the temperature is lowered which is essentially a consequence of diminishing phonon population at low temperatures. This result mirrors the situation in the rare-earth doped materials. Semiconductors, however, have the fortunate property that their EQE increases with decreasing temperature. The loss terms (A and C coefficients) decrease while the radiative rate (B coefficient) increases inversely with lattice temperature. Using the accepted scaling for $C(T) \propto \exp(-\beta(300/T-1))$ with $\beta \approx 2.4$ for GaAs [55, 69], taking $B \approx T^{-3/2}$ [59, 70], keeping $h\tilde{\nu}_f/h\nu - 1 \approx k_B T/E_g$, and ignoring parasitic losses and the small temperature dependence of the band-gap energy, we obtain for the break-even nonradiative decay rate

$$\frac{A_0(T)}{A_0(300)} \approx \left(\frac{300}{T}\right)^{3/2} \exp\left(\beta(300 - T)/T\right).$$  

© 2009 by WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim

www.lpr-journal.org
At $T = 150$ K, for example, the break-even lifetime is lowered by ~40 times compared with the room temperature ($T = 300$ K) condition. This is visualized by plotting $\eta_{\text{ext}}$ versus T as is shown in Fig. 10 for two values of $\eta_e$. This range of values of $\eta_e$ corresponds to a GaAs structure bonded to a high refractive index dome of ZnS or ZnSe [15, 68]. The solid red line indicates the break-even condition described by Eq. (2). This condition together with the fact that $A$ (typically dominated by surface recombination) increases with temperature [70–72], makes the low temperature observation of laser cooling more favorable even though the overall efficiency ($\approx k_B T / E_g$) decreases. The reduction in cooling efficiency is effectively due to the reduction of the electron-phonon absorption probability at lower temperatures. In particular, the population of LO phonons with a corresponding reduction of the exciton linewidth $\Gamma$ [59]:

$$\Gamma(T) = \Gamma_0 + \sigma T + \gamma N_{\text{LO}}(T)$$

(31)

where $\Gamma_0$ is due to impurities and inhomogeneous broadening, $\sigma$ accounts for the contribution of acoustic phonons, and $\gamma$ is the coefficient of LO-phonon scattering with $N_{\text{LO}}(T)$ denoting the corresponding Bose-Einstein phonon distribution. For the exciton densities involved, we can ignore possible broadening due to exciton-exciton scattering [73]. As the lattice temperature approaches 10 K, the acoustic phonon contribution begins to dominate. At such low temperatures, however, the exciton-phonon scattering rate ($\approx \Gamma$) becomes comparable to the radiative recombination rate (BN2) and consequently cold exciton recombination occurs before complete thermalization with the lattice. Similar processes, related to premature hot exciton recombination, have also hindered experimental observation of Bose-Einstein condensation in semiconductors. This problem is significantly alleviated by employing quantum confined systems where $\sigma$ is enhanced by nearly 3 orders of magnitude. This relaxes wave-vector conservation along the confinement directions [74]. Enhanced cooling in quantum confined systems may allow operation at temperatures < 10 K.

Another issue of concern is absorption saturation (band-blocking) and many-body interactions. Band-blocking may be a limiting factor for long wavelength excitation where the low density of states gives rise to a stronger bleaching of the interband absorption. It is therefore necessary to have a good understanding of the absorption and emission spectra and its dynamic nonlinearities.

Theoretical models exist that deal with absorption spectra of semiconductor structures under various carrier densities and lattice temperatures. With different levels of complexity, there are theoretical calculations for 2D and 3D systems that deal with such many-body processes under dense e-h excitation [75–78]. Recently, a rigorous microscopic theory for absorption and luminescence in bulk semiconductors that includes the effects of electron-hole (e-h) plasma density as well as excitonic correlations has been introduced under the quasi-thermal equilibrium approximation [47, 79]. The reader is referred to the above sources for the details. Here, we use a simple model to estimate the effect of band-blocking on achieving the carrier densities of Eq. (26) for GaAs. Using the electron-hole density of states corresponding to a simple two-parabolic band model, we calculate the carrier density at which the occupation factor $\{f_f - f_e\}$ vanishes at $h \nu = h \nu_f - k_B T$. The result is depicted in Fig. 11 where this blocking density is contrasted with $N_1$ and $N_2$ of Eq. (26) evaluated using a constant $A$ coefficient, and a temperature dependent $B$ and

**Figure 10** (online color at: www.lpr-journal.org) The required external quantum efficiency (EQE) as a function of temperature for GaAs under typical parameters.

**Figure 11** (online color at: www.lpr-journal.org) The upper carrier density $N_2$, given by Eq. (26) is seen to be unattainable in GaAs due to band-blocking as temperature is lowered below 200 K. The middle (dashed) line represents the calculated density at which the band-tail absorption at $h \nu = h \nu_f - k_B T$ completely saturates (i.e. $\alpha(N)=0$, Eq. (10)). This is a worst case scenario for which the nonradiative recombination rate is assumed to be constant with temperature.
Experimental work on optical refrigeration in semiconductors:

The first thorough experimental effort was reported by the University of Colorado [46]. No net cooling was achieved, despite realization of an impressive external quantum efficiency of 96% in GaAs. These experiments used a high quality GaAs heterostructure optically-contacted to a ZnSe dome structure for enhanced luminescence extraction. A report of local cooling in AlGaAs quantum wells by a European consortium [53] was later attributed to misinterpretation of spectra caused by Coulomb screening of the excitons [80]. Fig. 7b displays anti-Stokes luminescence in a GaAs heterostructure where excitation at λ = 890 nm produces broadband luminescence with a mean wavelength of λ = 860 nm. Each luminescent photon carries away about 40 meV more energy than an absorbed photon, so one might expect cooling. Why then have we not been able to observe laser cooling in this material or any semiconductor? To answer this question we have to revisit the cooling condition of Eq. (2) where strict requirements on EQE and background absorption are yet to be met. As discussed earlier, currently GaAs appears to be most promising due to the mature growth technology and record quantum efficiency. There are, however, challenges that must be overcome to achieve the break-even cooling condition: (a) reduce the surface recombination rate A, (b) reduce the parasitic background absorption αb, and (c) enhance the luminescence extraction efficiency ηe. Issues (a) and (b) involve material preparation and both concern high purity growth using advance epitaxial methods. Condition (c), on the other hand, is a light management and device engineering challenge that deals with luminescence extraction from semiconductors having high index of refraction. Total internal reflection causes the majority of spontaneous emission to get trapped and re-absorbed. A similar problem limits the efficiency of light emitting diodes (LEDs).

Various methods have been devised to remedy this problem for LEDs but not all are applicable to laser cooling. For example, photon recycling in thin textured or in photonic bandgap structures can substantially enhance the luminescence extraction but at the price of redshifting the luminescence as described by Eq. 15 [15]. Index matched dome lenses have been exploited for LED’s and can be used for laser cooling as well [46] provided the dome material does not introduce unacceptable levels of parasitic absorption. This requirement narrows the dome materials for GaAs to nearly index-matched ZnSe and ZnS due to their currently available high purity grade [81]. GaP substrates or domes can provide a higher ηe due to a better index matching to GaAs, but currently available materials produce unacceptable levels of parasitic absorption [81]. Another method for improving ηe makes use of nanometer dimension vacuum gaps (nanogap) as will be briefly discussed [6].

Highly controlled epitaxial growth techniques such as metal organic chemical vapor deposition (MOCVD) can produce very low surface recombination rates (A < 10⁴ sec⁻¹). This involves a double heterostructure of GaAs/GalnP as shown in Fig. 12 where the lattice-matched cladding layers provide surface passivation as well as carrier confinement. To deal with extraction efficiency, geometric coupling schemes such as nearly index-matched dome lenses (also shown in Fig. 12) have been employed to enhance ηe to 15%–20%. The double heterostructures are lithographically patterned into ≈ 1 mm diameter disks, lifted off from their parent GaAs substrates, and then van der-Waals bonded to a ZnS dome. The EQE of each sample is measured using the technique of fractional heating [46,68] at various temperatures and laser pump powers.

The experimental setup is shown in Fig. 13. The pump laser is CW Ti:sapphire laser producing up to 4.5 W tunable in the wavelength range 750–900 nm. This laser is pumped by an 18 W laser at 532 nm (Verdi, Coherent Inc.). In the fractional heating experiment, the laser is tuned around the mean luminescence wavelength while the temperature ΔT of the sample is measured. According to analysis described earlier, the temperature change in the sample is proportional to the net power deposited which can be written as:

\[
\Delta T(\lambda) = \kappa^{-1} P_{\text{abs}}(\lambda) \left(1 - \eta_{\text{ext}} \eta_{\text{abs}} \frac{\lambda}{\lambda_f}\right)
\]
where \( \kappa \) is the total thermal conductance (W/K) of the system positioned in an optical cryostat. During the experiment, the absorbed power is kept constant: the pump laser is tuned and its power adjusted to keep the luminescence (or a fixed spectral portion of it) constant. This is shown in Fig. 13 where a monochrometer is used to monitor the spectral portion of luminescence that does not overlap with the pump wavelength. The temperature change is monitored using a non-contact method called differential luminescence thermometry (DLT) that was developed for this purpose [68]. DLT exploits the temperature dependence of the luminescence resulting from the bandgap shift and broadening. High temperature sensitivity is obtained by monitoring the differential spectrum before and after pump irradiation. To avoid complications caused by high carrier density [80], low-density luminescence spectra induced by a weak diode laser generate the DLT signals. This is performed by modulating the pump laser with a mechanical chopper while synchronously gating a CCD camera on the spectrometer. This assures that the DLT spectra are recorded when the pump laser is blocked and after the high-density luminescence has decayed. DLT signals are obtained by normalizing the signal and reference spectra before subtraction. The resulting differential signal is a peak-valley or valley-peak feature depending on the sign of \( \Delta T \) which is calibrated in situ before an experiment. This method has exhibited a temperature resolution better than 1 mK [68].

Returning to Eq. (32) it is evident that by keeping \( P_{\text{abs}}(\lambda) \) constant, the measured \( \Delta T \) versus \( \lambda \) follows the wavelength dependence of the cooling (or heating) efficiency represented by the term in the bracket. At short wavelengths \( \lambda < \lambda_f \) where \( \alpha(\lambda) \) is large, \( \eta_{\text{abs}} \) can be taken as unity for moderate to high purity samples, and the fractional heating data follows \( \Delta T(\lambda) \propto (1 - \eta_{\text{ext}} \lambda / \lambda_f) \) which is a straight line. Currently available samples do not possess sufficient purity to make this term negative, i.e. net cooling. When the pump wavelength \( \lambda \) is tuned close to \( \lambda_f \), \( \alpha_{\text{abs}} \) tends to degrade, thus preventing net cooling. Extrapolation of the short wavelength data can be used to obtain a “zero-crossing wavelength” \( \lambda_c \) from which \( \eta_{\text{ext}} = \lambda_f / \lambda_c \) can be deduced. Recall that \( \eta_{\text{ext}} > \eta_{\text{ext}} \) and thus we obtain a lower limit on EQE. When parasitic absorption of luminescence can be ignored then \( \eta_{\text{ext}} \approx \eta_{\text{ext}} \). In the following analysis of the experimental data, we refer to this lower limit measured by fractional heating technique as the EQE. Fig. 14a shows the measured EQE for various thicknesses of the GaAs layer at room temperature. The optimum GaAs thickness is found to be about 1 \( \mu m \), determined by balancing excessive luminescence re-absorption for thicker layers with dominant surface recombination for thinner layers. [55, 82]. The measured temperature dependence of EQE is depicted in Fig. 14b, which is in qualitative agreement with the earlier analysis. Enhancement is observed as the temperature is lowered, reaching a record 99% at 100 K.

The fractional heating data leading to 99% EQE is shown in Fig. 15. We note the increase in temperature at longer wavelengths due to parasitic absorption. Knowing the absorption of GaAs, the data is fitted with a constant background \( \alpha_b \approx 10 \text{cm}^{-1} \) assuming it occurs entirely in the 1 \( \mu m \) thick GaAs layer. The experiment measures \( \eta_{\text{ext}} \) which means the unmodified EQE (\( \eta_{\text{ext}} \)) can be even larger.

**Figure 13** (online color at: www.lpr-journal.org) Experimental setup for measuring the external quantum efficiency (EQE or \( \eta_{\text{ext}} \)) at various lattice temperatures in GaAs/GaInP double heterostructures. A tunable cw laser (Ti:sapphire, 4.5 W) excites a constant density of electron-hole pairs as the pump is tuned near the band edge. This is done by keeping the luminescence intensity (within a certain spectral band) constant using the monochrometer. After recombination is complete, the temperature change of the sample is measured using differential luminescence thermometry (DLT) where the spectral shift of the low density luminescence spectrum induced by a weak cw laser diode is monitored using the gated spectrometer. The inset figure shows a typical DLT signal for \( \Delta T = 10 \text{ mK} \) in GaAs at \( T = 100 \text{ K} \).
Experiments show that even though external quantum efficiency is adequate for achieving net laser cooling, the heating of 1 µm thick GaAs sample as a function of excitation wavelength for a fixed (optimum) electron-hole density at starting temperature of 100 K. The extrapolation of short wavelength data determines the zero-crossing wavelength from which a record EQE of ~99% is measured. Excitation at longer wavelengths causes heating due to background parasitic absorption. [6, 68].

The average absorption coefficient ($\alpha_f \approx 4000 \text{ cm}^{-1}$) allows us to estimate an upper limit for $\alpha_b$ inside the double heterostructure. Even if $\eta_{\text{ext}} = 1$, attaining $\bar{\eta}_{\text{ext}} = 99\%$, requires that $\alpha_b < \frac{(1-\bar{\eta}_{\text{ext}})}{(1-\eta_{\text{f}}/\eta_{\text{e}})} \cdot \bar{\alpha}_f \approx 4 \text{ cm}^{-1}$ assuming $\eta_{\text{e}} \approx 0.1$. The value of $\alpha_b = 10 \text{ cm}^{-1}$ that was used to fit the data in Fig. 15 may have contributions from sources outside the double heterostructure such as dome or cold-finger contacts.

Experiments show that even though external quantum efficiency is adequate for achieving net laser cooling, the purity of the samples is not yet sufficient. Efforts are underway to address this material problem. Researchers at the National Renewable Energy Laboratory are using highly controlled MOCVD growth and scientists at the University of New Mexico are exploring aluminum-free GaAs heterostructure growth using phosphorous molecular beam epitaxy (MBE).

Methods to enhance $\eta_{\text{ext}}$ by improving luminescence extraction efficiency are being explored as well. A novel method based on the frustrated total internal reflection across a vacuum “nano-gap” is being developed at the University of New Mexico [17, 83, 84]. In this scheme, the luminescence photons tunnel through the gap into the absorber region. The vacuum nano-gap maintains a thermal barrier between the heterostructure and the absorber/heat sink. The cooling heterostructure and luminescence absorber are thus optically contacted but thermally insulated. Calculations show that a gap spacing of < 25 nm has higher extraction efficiency than the dome structure (GaAs on ZnS or ZnSe) and preliminary fabrication of such structures show promising results. Using a multi step photolithographic process, a Si-based nano-gap with 50 nm spacing supported by posts have been monolithically fabricated as depicted in Fig. 16 [83]. Recently, we have fabricated GaAs nano-gap structures that will be integrated with a high quality GaAs heterostructure to investigate their performance in cooling experiments.

5. Future outlook

Optical refrigeration has advanced from basic principles to a promising technology. Cooling of rare-earth based materials is approaching cryogenic operation. In semiconductors, much progress has been made in achieving high external quantum efficiency. With advanced heterostructure growth and novel device fabrication currently underway, cooling will soon be attainable. In the coming years, optical refrigeration will be useful in applications such as satellite instrumentation and small sensors, where compactness,
ruggedness, and the lack of vibrations are important. Optical refrigeration is well suited for space-borne applications since it has no moving parts and can be designed for long operational lifetimes. Additionally, the cooling element is not electrically powered so it will not interfere with the electronics being cooled. One can envision optical refrigerators being directly integrated with infrared sensors for thermal imaging of the earth and of astrophysical objects. In terrestrial applications, small size and makeability optical refrigerators attractive for use with high-temperature superconductor sensors and electronics. Optical refrigeration could enable compact SQUID magnetometers for geophysical and biomedical sensing and may be a critical component in the production commercial electronics incorporating fast and efficient superconducting components.

6. Appendix:
Photon waste recycling and the carnot limit

Many applications will be possible if the basic efficiency of optical refrigerators $\sim k_B T / \hbar \nu$ can be improved. This efficiency limit assumes all fluorescence photons are absorbed by a heat sink and thus wasted. This is depicted diagrammatically in Fig. 17a. One expects improved overall efficiency if fluorescence photons are recycled with photovoltaic (PV) elements to convert them into electricity [85]. This recovered energy can be used to drive laser diodes (LD) at the pump photon energy $\hbar \nu$. This process is diagrammed in Fig. 17b.

Deriving the new cooling efficiency is then straightforward: For the same cooling power ($P_c$) as in the open loop system, the required laser power in the photon-waste recycled system is lowered by the amount generated by recycling: $\eta_{PV} \eta_L P_f$ where $P_f$ is the luminescence power and $\eta_{PV}$ and $\eta_L$ are power efficiencies of the photovoltaic and the laser diode, respectively. The enhanced cooling efficiency $\bar{\eta}_c$ is then obtained as

$$\bar{\eta}_c = \frac{\eta_c}{1 - \eta_{PV} \eta_L (1 - \eta_c)}, \quad (33)$$

where the “old” or open loop cooling efficiency $\eta_c = 1 - P_f / P_L$, as before. It is useful to investigate what limits this new efficiency. Assuming $\eta_c$ is given by its quantum limit of $h \nu_f / h \nu - 1$, we take $h \nu = h \nu_f - m k_B T_c$, where m if of the order of unity and $T_c$ is the temperature of the solid-state coolant. This leads to $\eta_c = m k_B T_c / h \nu_f \approx m k_B T_c / E_g$, where $E_g$ is the energy gap of the transition. The obvious choice for the photovoltaic and laser diode system will be a semiconductor with energy gap very close to $E_g$. Assuming both are at a reservoir temperature of $T_R$, it is not unreasonable to take their conversion efficiency $\eta_{PV} = 1 - p k_B T_R / E_g$ and $\eta_L = 1 - q k_B T_R / E_g$, respectively. Current devices reach efficiencies that correspond to $p \approx q \approx 30–50$, so one expects they will be limited by

---

**Figure 16** (online color at: www.lpr-journal.org) (a) A vacuum “nano-gap” structure where the heterostructure is situated (e.g. supported by posts) at a sub-wavelength distance from an absorber. Luminescence photons will escape into the absorber via frustrated total internal reflection (photon tunneling) while the gap provides a thermal barrier. (b) The SEM micrographs of a preliminary nanogap structure (with 50 nm spacing) fabricated using a multi-step photolithographic technique [83].

**Figure 17** (online color at: www.lpr-journal.org) Diagrammatical representation of (a) an open loop optical refrigerator (OR) and (b) a closed loop system in which the luminescence (photon waste) recycling is performed using photovoltaic (PV) followed by a laser diode (LD) both at temperature $T_R$. 
We note that thermodynamic analysis require that \( p \approx q \approx 1 \). In such a limit one writes Eq. (33) as:

\[
\bar{\eta}_c = \frac{T_C}{T_C + \left( \frac{p+q}{m} \right) T_R}
\]

(34)

We note that thermodynamic analysis require that \( p + q \geq m \approx 1 \). With \( p + q = m \approx 1 \), one obtains the Carnot limit of \( T_C / (T_C + T_R) \). Expected technological advances in photovoltaic and laser diode devices will enable efficiencies approaching that corresponding to \( p \approx q \approx 1 \), hence the ultimate efficiency of the optical refrigerator will be given by its Carnot limit.

Acknowledgements This work has been supported by Air Force Office of Scientific Research (MURI program), NASA, and US Department of Energy.

References


[23] A. Rayner, M. E. J. Friese, A. G. Truscott, N. R. Heckenberg, and H. Rubinsztein-Dunlop, Laser cooling of a...
[67] H. Yi, J. Diaz, B. Lane, and M. Razeghi, Optical losses of Al-free lasers for \lambda = 0.808 and 0.98 \textmu m, Appl. Phys. Lett. **69**, 2983–2985 (1996).