Invited Comment

Why the laser linewidth is so narrow: a modern perspective

Alexander Cerjan and A Douglas Stone

Department of Applied Physics, Yale University, New Haven, CT 06520, USA

E-mail: douglas.stone@yale.edu

CrossMark

Abstract

We review and interpret a modern approach to laser theory, steady-state *ab initio* laser theory (SALT), which treats lasing and amplification in a unified manner as a non-unitary scattering problem described by a nonlinear scattering matrix. Within the semiclassical version of the theory the laser line has zero width as the lasing mode corresponds to the existence of an eigenvector of the S-matrix with diverging eigenvalue due to the occurrence of a pole of the scattering matrix on the real axis. In this approach the system is infinite from the outset and no distinction is made between cavity modes and modes of the Universe; lasing modes exist both in the cavity and in the external region as solutions satisfying Sommerfeld radiation boundary conditions. We discuss how such solutions can be obtained by a limiting procedure in a finite box with damping according to the limiting absorption principle. When the electromagnetic and matter fields are treated as operators, quantum fluctuations enter the relevant correlation functions and a finite linewidth is obtained, via a generalization of SALT to include noise (N-SALT). N-SALT leads to an analytic formula for the linewidth that is more general than all previous corrected versions of the Schawlow–Townes formula, and can be evaluated simply from knowledge of the semiclassical SALT modes. We derive a simpler version of this formula which emphasizes that the noise is dominated by the fluctuations in the polarization of the gain medium and is controlled by the rate of spontaneous emission.

Keywords: laser theory, laser linewidth, microcavities, semiconductor lasers

(Some figures may appear in colour only in the online journal)

1. Introduction

The laser has been a critical enabler of the modern discipline of quantum optics via its highly monochromatic and intense emission properties. The physical understanding of steady-state laser emission has presented challenges since the early days after its invention, as first semiclassical [1, 2] and then fully quantum laser theories were developed [3–6]. Although these theories were able to capture most of the relevant physics, they tended to incorporate some phenomenological or approximate steps due to the complexity of dealing with the full nonlinear and space-dependent semiclassical or quantum laser equations. Over forty years ago, several pioneers of early laser theory, Lang, Scully and Lamb, addressed some of the weak points in the theory in a classic paper entitled, 'Why is the laser line so narrow? A theory of single quasi-mode laser operation' [7]. This work was aimed at a fundamental challenge in laser theory, dealing with the openness of the laser system, which invalidated standard Hermitian modal decompositions appropriate for closed systems. Prior to this work essentially all theories were framed in terms of 'quasi-modes' of the type studied by Fox and Li [8], and were based on phenomenological modal equations with cavity damping inserted to represent the outcoupling of the modes. The goal of Lang *et al* was to go beyond these quasi-mode equations for the laser field amplitudes to a more fundamental description which included both the cavity and exterior regions explicitly (i.e. a model of 'the Universe'), and

then work back from the modes of the Universe to the quasi-modes of the cavity region alone. Using this approach the authors (a) 'hope(d) to understand the mechanism leading to the extreme monochromaticity of the laser radiation', and (b) 'to investigate the sense in which the present calculation, which does not include the cavity dissipation as a pheno-mological loss mechanism, still leads to and implies a fluc-tuation–dissipation (FD) theorem' [7].

While this approach to understanding the laser equations was groundbreaking at the time, there has been much progress in laser theory in the subsequent years and it seemed to us a good time to re-examine these questions from a modern perspective. One reason for this is that during the past decade a new theoretical framework has been developed, known as the steady-state *ab initio* laser theory (SALT), which treats the openness of an arbitrary cavity exactly within the semi-classical theory [9–12], and thus has provided new insights into lasing phenomena. The resulting nonlinear SALT equations for steady-state lasing can be solved efficiently using a non-Hermitian basis set method and have been shown to agree to very high accuracy with brute force simulations of the atomic lasing equations [13, 14], including even *N*-level models with multiple independent lasing transitions [15]. The theory can also be generalized to treat injected signals (I-SALT), as will be relevant below [16]. Moreover, recently Pick *et al* have added Langevin noise to the SALT equations (N-SALT) and, using the SALT steady-state lasing modes as a starting point, calculated the quantum-limited laser line-width analytically [17]. Their analytic results have very recently been confirmed, again by direct simulations of the lasing equations, but with appropriate noise terms included [18]. It should be said that many aspects of the results obtained through SALT and N-SALT were anticipated by prior theoretical approaches, a partial list of which includes the work by Spencer and Lamb [19], Fu and Haken [20–22], and Henry [23]; however, SALT and N-SALT appear to be the most general, flexible and inclusive formalism yet developed for steady-state lasing. It treats the fields over the entire 'Universe' from the outset, but in a different manner from Lang *et al*, and does not arrive at modal equations with damping in the conventional sense. We will now examine the questions raised by Lang *et al* from the perspective of SALT and N-SALT.

2. Lasers as scattering systems

2.1. Motivation and definition of the problem

Most early lasers emitted from lasing modes which were concentrated along some optical axis, e.g. the axis joining the centers of two mirrors or the geometric or index profile defining a waveguide or fiber, and so there was a tendency to develop laser theories focused on one-dimensional or axially symmetric equations, and this is overwhelmingly the type of formulation one finds in text books [24–26]. However, in the past two decades a great variety of laser systems have been developed with complex two or three-dimensional cavities,
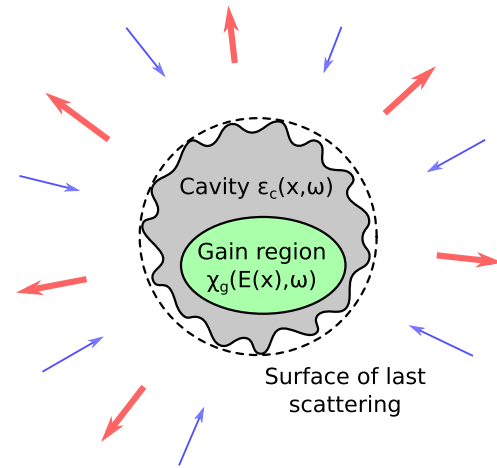

Figure 1. Schematic of lasers as scattering systems: a singly connected region surrounded by a last scattering surface containing an irregularly shaped cavity defined by a linear dielectric function $\varepsilon_c \neq 1$ and a nonlinear gain region. Any incoming signals (blue arrows) are amplified and scattered by the cavity to outgoing states (red arrows); lasing corresponds to outgoing solutions without any input, and is only possible at discrete frequencies for a given cavity.

such as microdisks and microspheres [27–29], micro-toroids [30], deformed (chaotic) disk cavities [31–33], photonic crystal defect-based and surface emitting lasers [34–36] and random lasers [37–43]. To describe such devices it was necessary to formulate laser theory in a manner that could both describe, and be implemented computationally, for cavities with an arbitrary geometry, and to include the full Maxwell wave equation, not its paraxial approximation.

Moreover, some of the new lasers (e.g. random lasers) are extremely open systems, not even cavities in the usual sense; nor do they always emit well-behaved narrow beams of light but rather generate complex emission patterns in two and three-dimensions. Thus it was crucial to find a description which did not over-simplify the spatial properties of the cavities and emission patterns, and which hence treated the cavity boundary and outcoupling exactly. The solution to this challenge provided by SALT is to describe the laser as a scattering system. Any laser consists of two essential ele-ments: a cavity region, given as a linear dielectric function, $\varepsilon_c(\mathbf{x})$, which describes the scattering and confinement of photons, either injected externally or internally generated, and a gain region, containing the medium which will be pumped to provide optical gain (sometimes cavity and gain medium essentially coincide). To include all cases the 'cavity' will refer simply to a singly connected region large enough to enclose all the regions in which $\varepsilon_c(\mathbf{x}) \neq 1$; and for simplicity we will assume the gain medium is fully contained within this region as depicted in figure 1.

The laser so defined can function both as an amplifier of input radiation (when the gain medium is inverted) and as a self-organized oscillator/source of radiation in the absence of input, above the first threshold for lasing. From this point of view the laser is a scattering system which both scatters and amplifies input radiation, and also, at discrete frequencies,

scatters specific internally generated modes of outgoing radiation when sufficiently pumped. However, unlike linear scattering systems, the laser is subject to nonlinear saturation effects which are fundamental to its stability and must be taken into account.

Other nonlinear effects within lasers, such as four-wave mixing, can destabilize their simple harmonic time-response, but these effects are neglected in SALT, and also (typically) in the earlier laser theories which motivated the work by Lang *et al* [7]. Leaving out such effects is equivalent to assuming a stationary inversion density, giving rise to a gain susceptibility which does not mix frequencies. This assumption is always valid for single-mode lasing and amplification, which will be the focus of this work; it is also valid under certain rather general conditions for multimode lasing in microlasers, and SALT has been shown to apply quantitatively to such cases [13–16].

Before proceeding, it should be mentioned that the input/ output theory developed by Collett and Gardiner (and others) also considers the cavity to be a scattering element [44, 45] and is able to treat the full quantum operator properties of the scattered fields. This approach has certainly contributed to our understanding of open quantum systems and the role played by the reservoirs and noise processes. However, while there has been substantial effort expended to describe lasing within this framework [46–49], it remains an open problem how to include the strong nonlinearities present in active lasing systems in this formalism; whereas at the semiclassical level this is done quite accurately by SALT.

2.2. Scattering in the SALT formulation

Initially we focus on the single-mode lasing case, as was done by Lang *et al*, so that the positive frequency component of the electric field will be given by:

$$\mathbf{E}^{(+)}(\mathbf{x}, t) = \mathbf{\Psi}(\mathbf{x}) e^{-\mathrm{i}\omega t}, \qquad (1)$$

where the *real* number, ω, is either the input radiation frequency (for the amplifier) or the unknown frequency of the laser oscillator, which must be self-consistently determined. $\mathbf{\Psi}(\mathbf{x})$ is the unknown spatial lasing (or amplified) mode, not assumed to be equal to any passive cavity mode or quasi-mode and is also self-consistently determined; such lasing/ amplifier modes *describe the solution over all space*, not just inside the cavity.

Assuming stationary inversion and neglecting quantum and thermal fluctuations, we can describe $\mathbf{\Psi}(\mathbf{x})$ for an arbitrary laser by a nonlinear wave equation of the form:

$$\left[\nabla \times \nabla \times - \left(\varepsilon_{\mathrm{c}} + 4\pi\chi_g \right) k^2 \right] \mathbf{\Psi}(\mathbf{x}) = 0, \qquad (2)$$

where $k = \omega/c$, and χ_g is the complex, frequency-dependent saturable gain susceptibility, which can be calculated for an arbitrary N-level atomic gain medium in steady-state [14]. For

the case of a two-level gain medium it takes the form

$$\chi_g(\omega, \mathbf{x}) = \frac{\frac{1}{4\pi} D_0 \gamma_\perp}{\left(\omega - \omega_{\mathrm{a}} + \mathrm{i}\gamma_\perp \right)\left(1 + \Gamma(\omega - \omega_{\mathrm{a}})|\mathbf{\Psi}(\mathbf{x})|^2 \right)}, \qquad (3)$$

where D_0 is the pump parameter (proportional to the inversion), $\gamma_\perp$ is the dephasing rate of the polarization, ω_{a} is the atomic resonance frequency, and $\Gamma(\omega - \omega_{\mathrm{a}})$ is a Lorentzian centered at ω_{a}. D_0 and $\mathbf{\Psi}$ are measured here in appropriate dimensionless units [13], and this form assumes that there is only one mode, either lasing or amplified, active in the medium. (This is the only case we will consider explicitly here, but we will review the more general multimode SALT formalism in appendix A). The nonlinear gain susceptibility represents the polarization of the gain medium and is the source of the amplification and (above threshold) of laser emission. The lack of any Langevin force terms in (2) is a result of the semiclassical approximation, in which the operators for the electromagnetic and polarization fields are replaced by their average values, eliminating the effects of quantum and thermal fluctuations. Such terms are taken into account in N-SALT, which will be discussed briefly in section 7, but were neglected (intentionally) in the work of Lang *et al*.

Assuming the system is pumped below the first lasing threshold and the inputs are small enough to generate negligible saturation of the gain, this equation defines a linear non-unitary scattering (S-)matrix for input fields. Depending upon the geometry there are different natural basis states describing input and output modes, $\mathbf{H}_\alpha^{(-)}(\mathbf{x})$, $\mathbf{H}_\alpha^{(+)}(\mathbf{x})$, from and to infinity, i.e. sufficiently far away from the cavity region

$$\begin{aligned} \mathbf{\Psi}(\mathbf{x}) &= \sum_\alpha \left[c_\alpha \mathbf{H}_\alpha^{(-)}(\mathbf{x}) + d_\alpha \mathbf{H}_\alpha^{(+)}(\mathbf{x}) \right] \\ &= \sum_\alpha \left[c_\alpha \left(\mathbf{H}_\alpha^{(-)}(\mathbf{x}) + \sum_\beta S_{\alpha,\beta} \mathbf{H}_\beta^{(+)}(\mathbf{x}) \right) \right]. \end{aligned} \qquad (4)$$

If the pump is strong enough to invert the medium and create gain at the input frequency, the S-matrix will be amplifying, with eigenvalues greater than unity. When functioning as an amplifier the boundary conditions are familiar: an input wave from infinity interacts with the cavity plus gain medium, scattering according to the appropriate dielectric or metallic boundary conditions in different regions of the cavity and giving rise to an outgoing wave which differs from passive scattering in that it is amplified and more photon flux emerges from the scattering region than enters it. The case of large inputs leading to a saturated gain and the resulting nonlinear S-matrix can be handled by I-SALT and will be included in the discussion below.

As noted, In this formulation there is no distinction between cavity modes and modes of the Universe, every amplified mode exists inside and outside the cavity as a steady-state scattering solution. However, we now consider to the more challenging and subtle case of laser oscillation in the scattering picture.
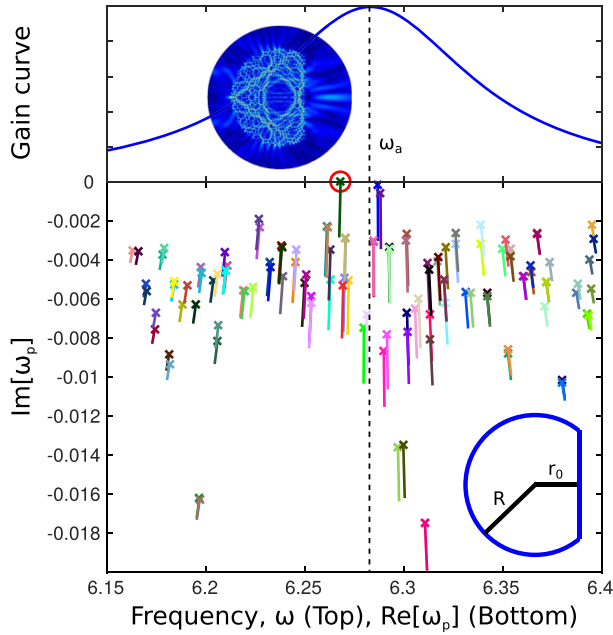
Figure 2. Simulations for a D-shaped laser cavity with $R = 4 \, \mu$m, $r_0 = 0.5R$, $n = 3.5$ as shown in the bottom schematic, with a uniform two-level atomic gain medium characterized by $\lambda_a = 1 \, \mu$m and $\gamma_\perp = 10$ nm. (Top) plot of the Lorentzian gain profile for the atomic medium. (Bottom) plot of the motion of the poles of the S-matrix as the D-shaped cavity is pumped to the first lasing threshold. The pole corresponding to the first lasing mode is seen to reach the real axis (marked with a red circle), and the spatial profile of the lasing mode at threshold is shown just above.

3. Quasi-modes, resonances and quasi-bound states

Any scattering region of finite extent with a real potential or dielectric function, and hence with a unitary S-matrix, will have associated with it resonant solutions; i.e. stationary solutions at discrete complex frequencies which are purely outgoing. It is easiest to think of these as evolving out of the bound states of the system at real frequencies, $\tilde{\omega}_n$, which would occur if the region were surrounded by a perfect lossless reflecting barrier. As the reflectivity of the fictitious barrier is reduced from unity to zero these real frequency solutions, corresponding to standing waves, evolve into complex frequency solutions with a substantial outgoing wave component (and no incoming wave from infinity). Due to the purely outgoing boundary conditions it can be shown that the imaginary part of the frequency must be negative, and the S-matrix, continued to complex frequencies, will become infinite with simple poles at $\omega_n - i\gamma_n/2$, $(\gamma_n > 0)$. Hence these purely outgoing solutions are discrete and countably infinite, in one to one correspondence with the bound states of the closed cavity.

In the asymptotic region the frequency and wavevector of the solutions are linked by the dispersion relation and the sign of γ_n will cause the corresponding spatial solution to grow exponentially at infinity and carry a diverging outgoing photon flux. Hence these solutions are not physically realizable, but have physical significance because the S-matrix will

have a resonant response near the real values ω_n, with a bandwidth given by γ_n (assuming well-separated high finesse resonances). The corresponding spatial solutions are referred to by various terms: quasi-modes, quasi-bound states and resonances. Historically the use of multiple terms seems to have arisen from the transition from 3D confinement, as e.g. in a microwave resonator, where the connection to bound states is more apparent, to open, axial mirror-based resonators, where the resonances perpendicular to the mirror axis are essentially absent, and the Fox–Li method was developed to determine the high Q resonances. However, there is no fundamental difference between the various types of outgoing solutions: they all correspond to poles of the passive-cavity S-matrix.

Such states also are very well-known from quantum scattering theory (where they are usually referred to as quasi-bound states), and are often considered to be physical states because if an initial condition is given corresponding to the quasi-bound state truncated to the 'cavity region', it will leak out to infinity with a rate given by γ_n (this property was explicitly demonstrated for the one-dimensional cavity treated by Lang *et al*). However, in this context we would like to emphasize that the true quasi-bound states, including the exponentially growing component at infinity, are not physically realizable steady-state solutions and cannot correspond to the lasing mode. Also, since they occur at the poles of the S-matrix for the passive cavity, they can only describe the properties of the passive cavity, independent of the gain medium and spectrum. Thus using 'quasi-mode' also as a term for the lasing mode in the presence of gain can, and in our view often has, led to confusion.

4. Threshold lasing modes and constant flux states

However, as described above, the properties of the laser/amplifier are not governed by a unitary S-matrix once the complex gain susceptibility is taken into account. This complex susceptibility violates the Hermiticity of the wave equation even for the closed cavity and invalidates the standard modal decomposition. But it makes possible purely outgoing solutions at real frequencies, which hence are physical solutions over all of space. The non-equilibrium steady-state gain susceptibility depends on the pump level, with its imaginary part changing sign from positive (absorbing) to negative (amplifying) at the transparency point, at which the upper and lower lasing levels are equally occupied and the susceptibility vanishes. At this pump value the poles of the cavity with gain coincide again with those of the passive cavity. As the pump is increased beyond the transparency point, all of the passive cavity poles continue to exist, but typically they move up in the complex plane towards the real axis [12, 50, 51], at a rate that depends on their position in the gain spectrum. The S-matrix of the amplifier can be measured for real frequencies as the pump increases, and the height of the resonances corresponding to each pole increases (the transmitted flux is greater than the incident flux due to amplification), while their width decreases (gain narrowing),

because the effective loss rate is proportional to the imaginary part of the pole frequency, which is decreasing. If there is no input to the amplifier then there is still no emitted field and no indication within semiclassical theory of the approach to lasing (although when quantum fluctuations are included there will be enhanced noise, amplified spontaneous emission, near the gain center). If lasing is achieved, one of the poles of the S-matrix reaches the real axis, gain completely balances loss, and we have a zero width resonance on the real axis. This process is demonstrated in figure 2 for a non-trivial two-dimensional cavity, a chaotic 'D-shaped' dielectric cavity [33], where each pole is seen to move towards the real axis an amount proportional to its distance from the center of the gain curve and one of the poles experiences enough gain to reach threshold.

It first may seem that we are back at the familiar situation of true bound states, which correspond to poles on the real axis, but this is not so. In the former case there is no photon gain and one simply has a decaying standing wave pattern, trapped in some region of space. Here we have a different type of object, which we call a TLM. This type of solution never occurs in Hermitian theories, a purely outgoing solution at real frequency; it reflects the fundamentally non-Hermitian nature of the electromagnetic wave equations with a complex susceptibility, which is an effective theory in which the matter degrees of freedom are hidden. Such a solution is characterized by a steady-state photon flux out to infinity, but with no unphysical growth of intensity *outside* the cavity and an emission pattern corresponding to the eigenvector of the S-matrix with a diverging eigenvalue. It is an example of a CF state: photons are generated within the cavity but are conserved outside the cavity, hence corresponding to a true physical lasing solution [9, 11, 12]. In appendix A we will show that each TLM is just one member of a complete biorthogonal basis set of CF states which forms a natural basis set for describing the nonlinear lasing solution above threshold.

However, we still have a conceptual problem: the linear S-matrix for this amplifier is now infinite at the lasing frequency. Within linear theory any input would be infinitely amplified, violating conservation of energy. The resolution of this dilemma is that infinitesimally above threshold the saturating nonlinearity kicks in and stabilizes the solution. In fact the steady-state lasing amplitude rises smoothly from zero and has no jump at threshold, just a discontinuous derivative. Even though formally the S-matrix has a pole fixed on the real axis above threshold, any input wave at the lasing frequency simply increases the gain saturation and does not lead to a runaway of the output. Strictly speaking, any such input actually forces the pole off the real axis, as we will show with an example in section 5. Hence the nonlinearity is essential for stablizing the linear CF solution. However, the actual solution near threshold is accurately described by the single TLM calculated from the linear wave equation with gain, since the saturation term is not strong enough to change the spatial form of the solution significantly.

Returning to the fundamental question of modes of the Universe versus quasi-modes of the cavity, we summarize the SALT perspective.

- If one treats lasers as unbounded scattering systems from the outset, all solutions are defined both inside and outside the cavity region and there is no distinction between cavity modes and modes of the Universe.
- If the laser is functioning as an amplifier, the solutions are simply amplified scattering states determined by the boundary conditions in the cavity region and the degree of gain at the input frequency.
- If the laser is functioning as a self-organized oscillator the solutions are purely outgoing states corresponding to an eigenvector of the linear scattering matrix with diverging eigenvalue due to the existence of a pole on the real axis. Such solutions require the presence of the saturating nonlinearity to be physically stable.
- For any given cavity and gain susceptibility such purely outgoing states can only occur at discrete frequencies because they arise from moving the discrete poles of the passive cavity onto the real axis. Within semiclassical theory (neglecting quantum and thermal fluctuations of the fields) these solutions have precisely zero linewidth.

We reserve the term 'lasing mode' for these special discrete solutions.

In summary then, our response to the question 'Why is the laser linewidth so narrow?' is that indeed this *is* due to the extreme limit of gain narrowing: only at discrete frequencies can the linear gain diverge and generate self-organized oscillation. As we will show in the next section, this zero-width resonance cannot be observed with an input signal, i.e. in amplification mode. However, these perfectly mono-chromatic purely outgoing solutions can be found by SALT and have been verified by brute force finite-difference time domain (FDTD) simulations of the laser equations for many examples [12, 14, 15]. Such infinitely sharp lines are not of course obervable in true laser emission due to quantum fluctuations as we will discuss in section 7. However, the smallness of these quantum effects lead to the extraordinary narrowness of the resulting lines, often many orders of magnitude narrower than the passive cavity resonance linewidths.

The effects of quantum fluctuations were intentionally not included in the work of Lang *et al*, who thus also found a delta function laser line within semiclassical theory. In their approach they find that this extreme narrowing of the passive cavity resonance can be regarded as a 'mode-locking effect' in which all of the modes of the Universe within the passive cavity resonance linewidth contribute to one lasing quasi-mode with a single frequency. This point of view emerges from a limiting procedure in which a finite system with quantized electromagnetic modes is studied and at the end the system size is taken to infinity in the presence of small fictitious absorption to damp out reflections from the boundaries as they tend to infinity. The conditions for this procedure to reproduce the correct infinite system behavior are discussed in section 6 below. We note that Lang *et al* expressed
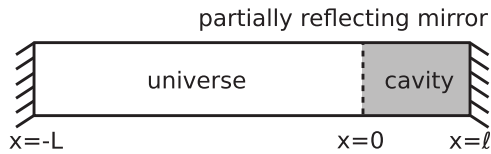
Figure 3. Schematic showing the system studied by Lang *et al* [7], which contains two regions, the cavity and the much larger 'Universe' (which also includes the cavity). Damping is added to the external region to prevent reflections from the boundary at $x = -L$, and as discussed below, this must be done by a particular limiting procedure in order to obtain the correct result for the laser.

dissatisfaction with invoking gain narrowing to explain the laser linewidth, because in their view such arguments (at the time) did not take into account the nonlinearity of the problem. However, the conceptual framework of SALT/I-SALT, outlined here, *does* take nonlinearity into account, essentially exactly, for single-mode lasing. We will now illustrate the previous concepts with a concrete one-dimensional example, very similar to the one studied by Lang *et al*, in the process making all of our statements above mathematically precise.

5. Lasing modes as zero-width resonances: a simple example

The example studied by Lang *et al* was a one-dimensional cavity of length $L + l$, with perfectly reflecting mirrors at $x = l, -L$ and a partially reflecting delta function mirror at $x = 0$, as shown in figure 3. The 'cavity' is the region $[0, l]$ and the 'Universe' is the region, $[-L, 0]$, where in treating lasing a small uniform absorption is added over all space, the limit $L \to \infty$ is taken, and lasing emission is in a single direction towards $x = -\infty$. For such a system the passive cavity will have a reflection coefficient of modulus unity at all frequencies, and resonant reflection will be distinguished only by a long delay time in reflection and an associated build-up of field intensity in the cavity. We prefer to make a small modification of the Lang *et al* example and study a symmetric two sided partially transmitting (Fabry–Perot) cavity with laser emission to $x = \pm\infty$, and with passive cavity resonances distinguished by enhanced transmission, reaching unity at the center, and with width, γ_c. As with the Lang *et al* example, the electromagnetic part of the problem will be described by the scalar Helmholtz equation with gain. However, unlike Lang *et al*, no limiting procedure is applied; our system is infinite from the start and no fictitious absorption is added to avoid reflections from perfect mirrors far away from the cavity. Hence, consistent with our scattering approach, we start from continuous solutions for the passive cavity + 'Universe', and will impose outgoing boundary conditions at infinity on the lasing solutions.

We begin by reviewing the properties of such a one-dimensional optical cavity. In figure 4 we plot the transmitted intensity of an incident signal through the passive cavity at frequency ω_{in}. As expected we observe a series of resonances of the cavity, spaced by the free spectral range, $\Delta = \pi/cL$,
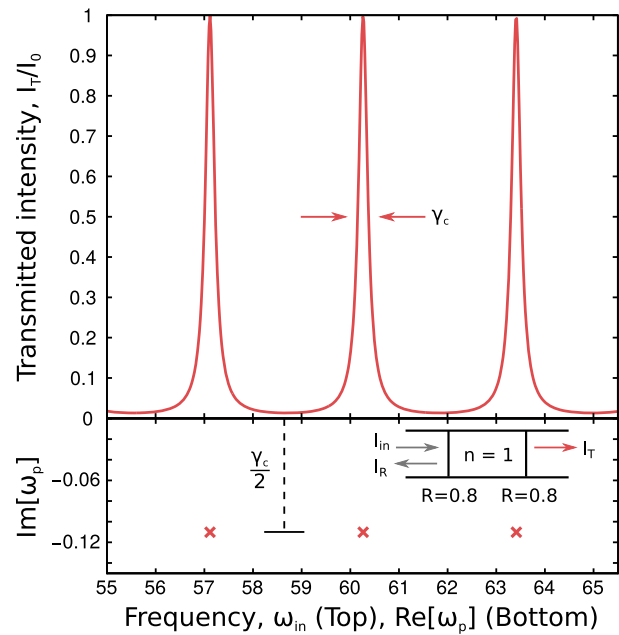

Figure 4. (Top) plot of the transmitted intensity of an injected signal with frequency ω_{in}, into a passive, two-sided cavity with partially reflecting mirrors ($R = 0.8$) on both ends, as indicated in the schematic. (Bottom) plot of the poles, ω_p of the scattering matrix of the same cavity, corresponding to purely outgoing solutions at complex frequency. The real frequency resonances of the cavity, where complete transmission is seen, are related to the poles of the cavity.

with a width given by the passive cavity decay rate $\gamma_c = -c/(2L)\ln(R^2)$, where L is the length of the cavity, and R is the reflectivity of the mirrors at the cavity boundary. The incident and reflected waves on the left side of the cavity satisfy a continuity boundary condition, with both incoming and outgoing components, whereas the transmitted wave leaving the right side of the cavity also is continuous but in addition is purely outgoing.

Any purely outgoing monochromatic wave is described mathematically by the Sommerfeld radiation condition [52],

$$\lim_{|x| \to \infty} |x|^{\frac{d-1}{2}} (\partial_{|x|} - ik)\psi(x) = 0, \qquad (5)$$

in which $\hat{x} = \vec{x}/|x|$, $\omega = ck$, and d is the dimensionality of the system. In one-dimension the external waves are simple plane waves (and not e.g. cylindrical or spherical waves) and the Sommerfeld condition can be imposed anywhere outside the boundary; thus at the boundary on the right side of the cavity the field must be an outgoing plane wave, satisfying

$$\left[(\partial_x - ik)\psi(x) \right]_{x=L} = 0. \qquad (6)$$

As noted, lasing would correspond to a solution satisfying a purely outgoing boundary condition on *both* sides of the cavity (or in all directions in higher dimension). If we attempt to impose this boundary condition on the passive cavity we find that no solutions are possible at real frequencies, but

there are discrete periodically spaced solutions at

$$\omega_m = \frac{\pi}{2cL} + m\Delta - i\frac{\gamma_c}{2} \qquad (7)$$

as shown in figure 4. The discreteness of the frequencies is a direct consequence of the Sommerfeld condition which can only be satisfied when there is constructive interference in the cavity. The fact that all the resonance frequencies have the same imaginary part is not generic, but is particular to one-dimensional cavities of this type. As noted above, the presence of these purely outgoing (pole) solutions leads to resonant transmission on the real axis centered at $\mathrm{Re}[\omega_m]$ with width γ_c. Applying the Sommerfeld condition to the outgoing solutions $e^{\pm ikx}$, and the relation $\omega = ck$ implies that $|\Psi(x)|^2 \sim e^{+\gamma_c|x|}$ at $\pm\infty$, the unphysical divergence we alluded to above.

In order to find normalizable physical solutions which are purely outgoing (i.e. lasing solutions) we need to add the effect of the gain medium to the cavity dielectric function in the form of χ_g as defined in (3) above. For this simple one-dimensional scalar model (similar to Lang *et al*), below threshold, the electric field satisfies

$$\left[\partial_x^2 + (nk)^2\right]\Psi(x) = 0, \qquad (8)$$

with the Sommerfeld boundary condition, where now the index of refraction $n = \sqrt{\varepsilon_c + 4\pi\chi_g} = n_1 - in_2$, and the gain nonlinearity is neglected, because we are below threshold and these are not yet physically allowed solutions (they are simply modified active cavity resonances). Note that the solutions inside the cavity here take the form $A\cos(nx) + B\sin(nx)$, i.e. they are trigonometric functions of a *complex* argument. This will remain true at threshold and is a key point which holds more generally: the electric fields for the physical system with gain cannot be purely real; if they were real they would conserve photon flux and describe true standing waves, whereas lasing or amplified modes must have running wave components. Any model which finds real spatial solutions for lasing or amplified modes has made an approximation. Such approximations have often been made in laser theory, including by Lang *et al* [7, 53–55] (see further discussion in section 6).

When we add the gain medium to the cavity and pump above the transparency point ($D_0 > 0$) the imaginary part of χ_g corresponds to amplification. This has an immediate effect, causing the poles to move upwards, as discussed earlier and as shown for this specific cavity in figure 5. While the pole solutions are still not physically realizable, we can observe a change in the cavity properties via the behavior of the transmission resonances at real frequencies. To calculate this effect accurately with a finite input wave we must now use the full saturable gain susceptibility of (3). The simulations shown here are performed using I-SALT which provides a nearly exact description of the steady-state single-frequency response of an active optical cavity to injected signals and any self-generated lasing signals [16], including the spatial degrees of freedom. SALT and I-SALT are reviewed in more detail in appendix A.
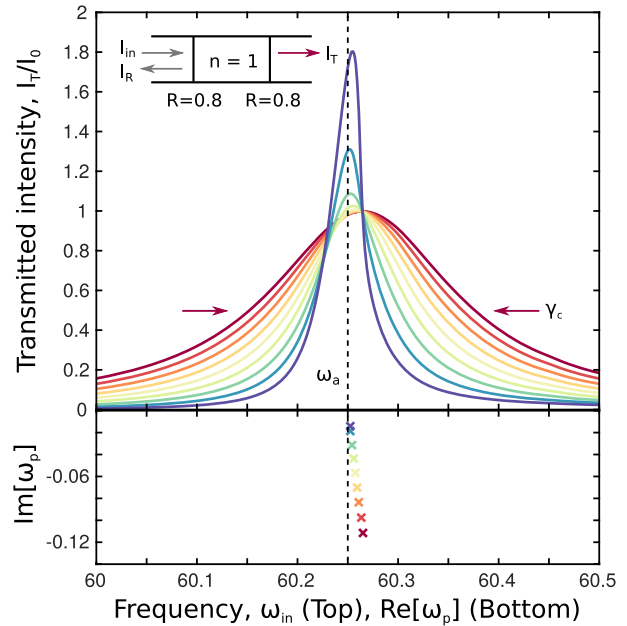

Figure 5. (Top) plot of the transmitted intensity of an injected signal as a function of frequency, ω_{in}, illustrating line narrowing with increased gain. $I_{in} = I_0 = 0.01$, and the two-sided cavity has partially reflecting mirrors ($R = 0.8$) at both ends, (see schematic). The pump on the cavity is increased from the transparency point, $D_0 = 0$ (red), to just above the first lasing threshold, $D_0 = 0.0040$ (blue), calculated using I-SALT [16]. As can be seen, the width of the cavity resonance narrows during this process. The slight shift of the resonance is due to line pulling from central atomic transition frequency, $\omega_a = 60$, with $\gamma_\perp = 2$. At the transparency point this system is identical to that shown in figure 4, and the outermost red curve seen here is identical to that from figure 4. (Bottom) plot of the motion of the pole of the associated cavity resonance, ω_p, through the complex plane as the pump is increased. This is also calculated using I-SALT, and thus includes the saturation of the gain medium due to the injected signal, which prevents the pole from reaching the real axis even though the final pump value is above the first non-interacting modal threshold. For the specific realization of the motion of the pole shown, the injected field is taken at $\omega_{in} = 60.25$, which corresponds to the peak in the blue curve. The pole is also seen to drift due to line pulling. Frequencies and rates are reported in units of c/L.

As the pole rises towards the real axis the amplified resonance narrows. Due to the frequency-dependent saturation of the gain medium, as already noted, injected signals with different frequencies will induce somewhat different trajectories for the pole in order to maximize the effective gain [16]. These are essentially line-pulling or pushing effects of the gain medium, which for a single frequency are always towards the atomic transition frequency. In the active cavity with dispersion, the precise width of the transmission resonance can no longer be defined as the distance of the pole from the real axis, as probing different frequencies will yield slightly different distances of the pole from the real axis. These line-pulling effects lead to the mild asymmetry in the series of curves shown in figure 5.

The bluest (and narrowest) curve shown in figure 5 corresponds to a pump value just *above* the non-interacting
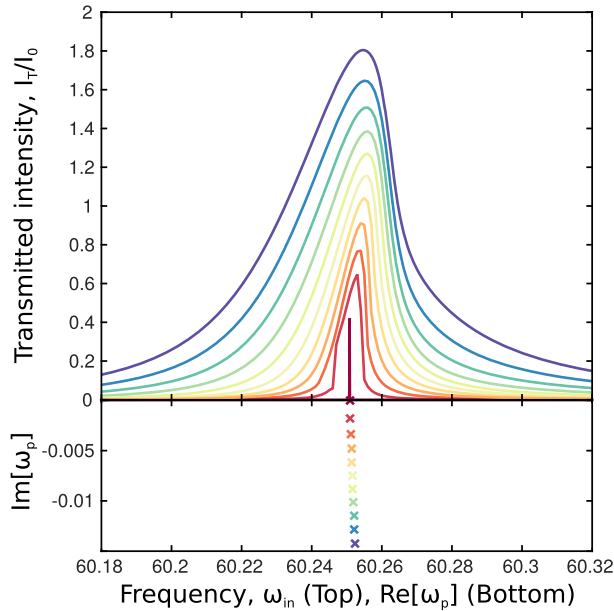
Figure 6. (Top) plot of the transmitted intensity of an injected signal as a function of frequency ω_{in}, into the active, two-sided cavity with partially reflecting mirrors ($R = 0.8$) on both ends, as shown in figure 5. The pump on the cavity is held fixed just above the first lasing threshold, $D_0 = 0.0040$, while the incident intensity is decreased from $I_{\text{in}} = I_0 = 0.01$ (blue), to $I_{\text{in}} = 0$ (red), where the cavity is operating as a free-running laser, calculated using I-SALT [16]. As can be seen, the width of the cavity resonance continues to narrow during this process, finally resulting in the lasing signal with infinitesimal width. Note, the outermost blue curve shown here is the same as that shown in figure 5. (Bottom) plot of the motion of the pole of the associated cavity resonance, ω_{p}, through the complex plane as the injected signal is turned off for the specific realization of $\omega_{\text{in}} = 60.25$. This is also calculated using I-SALT, and thus includes the saturation of the gain medium due to the injected signal, which prevents the pole from reaching the real axis until the injected signal is turned off. For different values of ω_{in} (not shown) the pole is able to reach the real axis before the injected signal is fully turned off.

first lasing threshold. This may seem initially puzzling; was not the lasing threshold the point at which the pole reaches the real axis? That definition only applies when there is no input signal. With an input signal, even at the lasing frequency, the saturation of the gain medium prevents the pole from reaching the real axis, keeping the response finite, and the amplifed resonance retains a non-zero width.

To see the true zero-width semiclassical lasing solution we slowly turn *off* the incident signal with the pump held above the first lasing threshold. With this protocol we see that the transmission resonance width continues to narrow, as shown in figure 6, while the pole continues to move towards the real axis. As the injected signal is turned off, the pole reaches the real axis and the cavity begins to undergo self-organized lasing action, shown as the reddest curve in the top panel of figure 6. This curve has precisely zero linewidth; no radiation comes out of the cavity except at this single frequency without input waves (assuming effectively zero temperature in the cavity and neglecting quantum fluctuations until section 7). This effect can never be fully achieved

through gain-narrowing of an amplified input signal, for the reasons we described above: gain saturation forces all poles off the real axis. Hence this effect does require correct treatment of the nonlinearity as stated by Lang *et al*. However, we have shown here that it is possible to treat the nonlinearity exactly at each stage and explain the zero width of the laser line quantitatively.

The model we have treated here is essentially the same as that treated by Lang *et al* except for the pedagogical difference of choosing a two-sided cavity in which one can study transmission resonances; in the next section we will study exactly the model of Lang *et al* by a different but equivalent method. As noted, in the standard SALT approach the system is infinite from the outset and there is no dissipation outside the cavity, which is in our view the most natural model for a real laser system. In contrast, Lang *et al* place their cavity in a much larger box which eventually will be taken to infinite size, and add dissipation external to the cavity, intended to damp unphysical reflections from the boundaries of the larger fictional box. We will discuss the effects of this difference in the next section.

6. Lasing modes and the limiting absorption principle (LAP)

While it is clear that lasing modes should satisfy the purely outgoing (Sommerfeld) boundary condition, this leads to a rather unfamiliar description of the spatial behavior of the lasing modes in terms of functions that are not power orthogonal, but instead satisfy a biorthogonality condition [25, 56, 57]. Basis functions of this type, the CF states [9–12], are central to the standard computational algorithm of SALT. The approach taken by Lang *et al* of beginning with a finite system and adding damping to remove the effect of the boundaries of the system is an alternative which has evolved into the well-developed computational technique of surrounding a cavity or region of emitters with a perfectly matched (damping) layer (PML). When done correctly this can reproduce the imposition of Sommerfeld conditions to arbitrary accuracy. In fact, newer computational methods for solving the SALT equations [58] make use of PMLs and do not impose the Sommerfeld boundary conditions in the straightforward manner of matching to outgoing solutions at a last scattering surface as did the earlier methods of satisfying the SALT equations [11]. The PML approach gives essentially perfect agreement with the Sommerfeld approach to SALT.

However, as noted above, the correct lasing solutions *must* be complex functions in space, so as to represent the non-zero photon flux, which grows inside the cavity and remains constant outside. For the simple 1D cavities studied here and by Lang *et al*, this means that the spatial solutions inside a constant-index cavity must be trigonometric functions of complex argument, corresponding to exponential growth towards the mirrors as well as oscillation. However, in the work of Lang *et al* they find a lasing mode of the form $E(x, t) \approx C_0 e^{(\alpha - i\omega)t} \sin[k_0(x - l)]$ for $0 < x < l$, where α
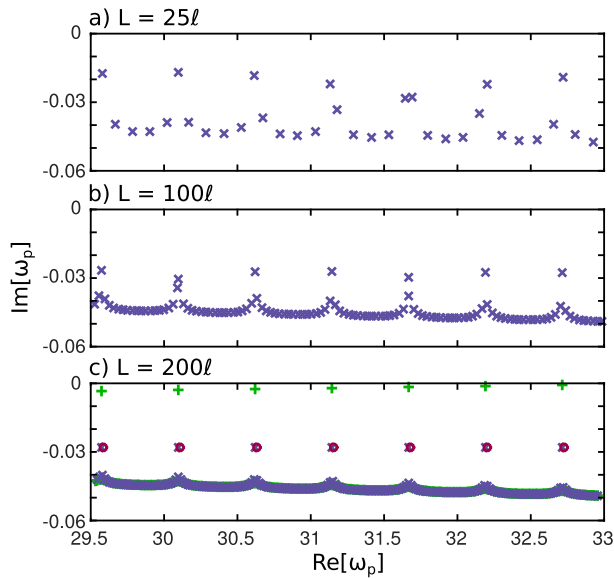
Figure 7. Plot of the resonances of the finite system with external damping depicted in figure 3 for three different values of L, $L = 25l$, $L = 100l$, and $L = 200l$, from top to bottom, except that the optical confinement is instead provided by a passive cavity dielectric of $n = 6$ in the cavity, rather than a partially reflecting mirror. In the external region an absorption of $\gamma = 0.003$ is present. In (c) the isolated solutions of the problem with damping are compared to the resonances of the same cavity subject to the Sommerfeld boundary condition on an infinite system without damping, given by (7) (red circles). Perfect agreement is found as discussed in the text. (Note that due to the presence of the passive cavity dielectric constant the right side of (7) must be multiplied by $1/n$ in making the comparison). The green crosses shown in (c) show the locations of the resonances when uniform gain is added to the cavity to find the TLM, $n = 6 - 0.005i$. As can be seen, the poles corresponding to resonances of the cavity have moved towards the real axis, with the rightmost one reaching the lasing threshold, while those corresponding to the resonances of the Universe have stayed relatively fixed, as discussed in the text. This is thus an alternative method to implement the Sommerfeld conditions for lasing modes.

is the difference between gain and loss, ω is the lasing frequency, C_0 is a constant, and $k_0 = \Omega/c$ is the real wavevector corresponding to the passive cavity resonance frequency, Ω. Since the spatial part of this function is real, it does not capture the photon flux through the partially transmitting mirror and is only an approximation to the complex sine wave solution which is correct for this one-sided cavity.

In our view the origin of this problem is the neglect of the LAP [59–61], which gives the correct mathematical procedure for recovering the Sommerfeld radiation condition from a finite system. Lang *et al* solved the spatial part of the problem for finite L, finding the quantized Hermitian basis states for the whole passive system ('modes of the Universe') and then addressing the lasing problem within this Hermitian basis set. Effectively the spatial part of the problem is fixed at finite L, and a small absorption, γ, is subsequently added everywhere in the Universe so as to eliminate the effect of reflection from the fictitious perfect mirror at $x = -L$. For this procedure to work one needs to solve the spatial problem

with the finite damping, and carefully take the limit $L \to \infty$ prior to $\gamma \to 0$; this requirement is known as the LAP. From a practical computational point of view the LAP means that one needs to solve the *non-Hermitian* spatial problem with Dirichlet (or other linear homogeous) boundary conditions *and* finite damping, while maintaining the condition $\gamma k L \gg 1$. The solutions so obtained have only exponentially small sensitivity to the presence of the boundary and reproduce to high accuracy those obtained by actually imposing directly the Sommerfeld conditions. Hence they are significantly different from the Hermitian solutions calculated before the damping is added (or when the limits are taken in the opposite order), which can carry no net flux as they are bound states. The modern PML techniques are a means for introducing *non-uniform* absorption which is impedance matched so that reflections are negligible even when L is *not* very large and one reproduces the Sommerfeld solutions with a much smaller and manageable computational cell. But the principle is basically the same: in order to get the correct spatial solutions one must treat the non-Hermitian problem with damping from the outset.

To illustrate this explicitly we briefly analyze the Lang *et al* system from this point of view. The system is the one-sided cavity between $[0, l]$ attached to the 'Universe' between $[-L, 0]$, as shown in figure 3, with perfect mirrors (Dirichlet boundary conditions), $E(l) = E(-L) = 0$, where $E(x, t)$ is the electric field of the system. Before adding the gain and considering lasing, we explicitly include the weakly absorbing material uniformly in the Universe, and focus on the region external to the cavity. The wave equation in this region is written as

$$\left[\partial_x^2 - \frac{1 + i\gamma}{c^2}\partial_t^2\right]E(x, t) = 0, \quad -L \leqslant x < 0. \quad (9)$$

We expand the field inside of the system in terms of a set of basis functions, $U_k(x)$, so that $E(x, t) = \sum_k U_k(x)e^{-i\omega t}$, and solve for the form of these basis functions in the Universe region as

$$U_k(x) = c_{1,k}e^{ikx - \frac{\gamma}{2}kx} + c_{2,k}e^{-ikx + \frac{\gamma}{2}kx},$$
$$-L \leqslant x \leqslant 0, \quad (10)$$

where we have expanded the exponent for $\gamma \ll 1$. The coefficients $c_{1,k}$ and $c_{2,k}$ and the quantized values of k are set by the boundary conditions $U_k(-L)$, $U_k(l) = 0$, and the matching conditions at $x = 0$, and will vary with L, l and the transparency of the mirror at the origin.

We first apply the LAP to obtain the Sommerfeld solution for this problem, taking the limit $L \to \infty$ prior to $\gamma \to 0$. Performing the limits in this order on the solution of $U_k(x)$, (10), one finds that to satisfy the Dirichlet boundary condition, $U_k(-L) = 0$, we must choose $c_{1,k} = 0$, as the term multiplying this coefficient is diverging as $-L \to -\infty$ for finite γ. Then, with the spatial limit taken, we can safely reduce the absorption in the problem to zero, recovering

$$U_k(x) = c_{2,k}e^{-ikx}, \quad -\infty \leqslant x \leqslant 0, \quad (11)$$

finding that all of the solutions of our problem are purely outgoing from the cavity, consistent with the Sommerfeld
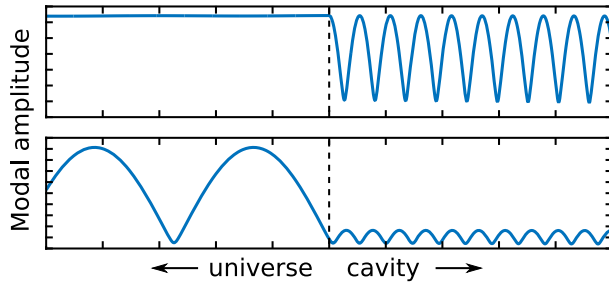
Figure 8. Plot of the modal amplitudes for a small region near the cavity-Universe boundary for two modes, one corresponding to a resonance of the cavity (top), and the other corresponding to a resonance of the external Universe (bottom). As can be seen, the cavity mode is a nearly pure running wave after it exits the cavity.

radiation condition, (5). However, if these two limits are taken in the opposite order, setting the absorption to zero before letting the length increase to infinity, we find that

$$\tilde{U}_k(x) = b_k \sin(k(x + L)), \quad -L \leqslant x \leqslant 0, \qquad (12)$$

which is exactly the result of Lang *et al* (outside the cavity region). With this result the matching to the cavity allows for real sine wave solutions, whereas with the outgoing solutions only complex sine waves are allowed in the cavity. The complex sines provide the outgoing flux at the partially transmitting mirror which is needed for continuity with the field outside the cavity.

These conclusions can be further reinforced by directly solving for the eigenfunctions of the finite system studied by Lang *et al*, shown in the three panels of figure 7. As L is increased, the modes of this finite system separate into two sets. One set consists of discrete isolated solutions with smaller imaginary frequency (damping); the number of such solutions does not change with increasing L. The second set has larger damping, and gets denser and denser, forming a psuedo-continuum as L is increased. The first set corresponds precisely to the complex resonance frequencies of the passive cavity, as can be confirmed by plotting the known resonances of the cavity given by (7), shown in figure 7(c). The spatial functions corresponding to the passive resonances are identical to the resonant solutions of the infinite system without damping *inside the cavity*, as there is no loss there, and the boundary condition at the mirror is the same. Outside the mirror these solutions experience a damping not present in the infinite system and hence decay slowly as they propagate outwards. On the scale shown in the top panel of figure 8 one sees only that the outgoing boundary condition is satisfied and not the weak decay. In contrast, the spatial solutions corresponding to the dense set of damped modes of the Universe can be seen in the bottom panel of figure 8; they form standing waves outside the cavity, with small leakage into the cavity, and do not satisfy the purely outgoing conditions. Hence simply discarding the latter and keeping the first set is equivalent to finding the Sommerfeld solutions of the passive problem without damping, in the vicinity of the cavity.

Finally, we can make contact with the SALT approach by adding gain (taken to be uniform for convenience) in the cavity region while leaving the damping present outside the cavity. The added gain has a much bigger effect on the cavity resonance solutions which have high amplitude there, than on the external standing wave solutions. For large enough gain one of the cavity resonances can reach the real axis and will then correspond to a TLM for uniform gain, as shown in figure 7(c) (green crosses). This is the essence of the SALT solution method adopted in [58] except in that work a PML damping layer is used for higher computational efficiency and the saturating nonlinearity is taken into account. Thus the Lang *et al* method agrees with the SALT approach if care is taken to apply the limits in the right order.

7. Quantum limited laser linewidth: an alternative form

Since this paper is devoted to the 'SALT perspective' on the laser linewidth, we include here a section on the finite laser linewidth due to quantum fluctuations, which has recently been addressed within two different generalizations of SALT, one based on input–output theory [50, 51] and another based on nonlinear perturbation theory around the SALT solution [17]. The fundamental origin of a finite quantum linewidth for the laser is the presence of spontaneous emission, which leads to random jumps in the time-dependent phase of the output field. The resulting phase diffusion is linear in time and the phase diffusion coefficient determines the intrinsic linewidth. The basic physical understanding of this goes back to the seminal work of Schawlow–Townes (ST) [62], which was subsequently substantially extended by Lax, Henry, Petermann and many others. Traditionally, the ST linewidth is written in terms of the cavity decay rate, γ_c, as

$$\delta\omega_{\text{ST}} = \frac{\hbar\omega_0 \gamma_c^2}{2P} n_{\text{sp}}, \qquad (13)$$

where ω_0 is the lasing frequency and P is the total output power. Here we have chosen to include in the ST formula the 'incomplete' inversion factor, n_{sp}, which is sometimes treated as a correction to it ($n_{\text{sp}} \geqslant 1$ measures the degree of inversion at the laser frequency and tends to unity when the lasing transition is completely inverted, the limit assumed by Schawlow and Townes in their original work) [63].

As discussed extensively above, semiclassical laser theory neglects quantum effects such as spontaneous emission and leads to a delta function linewidth. In order to go beyond this one needs to treat all three fields, the electromagnetic field, the polarization, and the inversion as quantum operators and consider their higher moments. The laser linewidth will be obtained from the correlation function of the electric field near the lasing frequency. It is well known that the effects of these quantum fluctuations can be captured by adding appropriate Langevin (random) noise terms to the equations of motion for the three fields [64]. These noise terms will be related to the dissipation and amplification in the system by the appropriate FD relations, and are then typically combined

to find the total effective noise upon the electric field

$$\left[\boldsymbol{\nabla} \times \nabla \times E(\mathbf{x}, t) + \varepsilon_{c}\ddot{\mathbf{E}}(\mathbf{x}, t) \right] = -4\pi\ddot{\mathbf{P}} + \tilde{\mathbf{F}}_{\text{tot}}, \quad (14)$$

in which $\tilde{\mathbf{F}}_{\text{tot}}(\mathbf{x}, t)$ is the total effective Langevin noise driving the field, and contains contributions from the noise in all three fields. This was one of the questions raised in the original work by Lang *et al*. It is now understood that semiclassical theory violates these FD relations but they are restored by including the Langevin terms

In this section we will first provide a new interpretation of the ST linewidth, arguing that it is more natural to express this quantity in terms of the rate of spontaneous emission rather than the passive cavity decay rate, as the former is directly connected to the physical process which yields the phase diffusion. Subsequently, we will provide an overview of the recent work by Pick *et al* [17], which provides the most complete treatment of the quantum limited laser linewidth known, and has been shown to agree quantitatively with first principles FDTD simulations [18].

7.1. Novel interpretation of ST linewidth

It seems initially puzzling that the noise due to spontaneous emission in the gain medium should lead to the standard ST formula, which is expressed only in terms of the passive cavity linewidth, γ_{c}, the total output power, and the lasing frequency. One would expect the spontaneous emission rate to appear as a fundamental quantity, not the cavity decay rate; and here we will show that the ST linewidth can (and perhaps should) be written in an equivalent form where this is the case.

The total Langevin force, $F_{\text{tot}}(t)$, in units of the quantized field amplitude operator, $\tilde{\mathbf{F}}_{\text{tot}} = i\sqrt{2\pi\hbar\omega} F_{\text{tot}}(t)\boldsymbol{\Psi}_{0}(\mathbf{x})$, has a direct contribution from the amplitude fluctuations of the field, $F_{a}(t)$, due to quantum fluctuations from outside the cavity, but also has indirect contributions from Langevin forces driving the atomic polarization and inversion, F_{p} and F_{d} respectively [64],

$$F_{\text{tot}}(t) = F_{a}(t) + \left(\frac{1}{\omega_{0} - \omega_{a} + i\gamma_{\perp}} \right) \sum_{\alpha} g_{\alpha} F_{p}(t), \quad (15)$$

where the summation runs over all of the atoms in the cavity at locations x_{α} and

$$g_{\alpha} = i\boldsymbol{\theta} \cdot \boldsymbol{\Psi}_{0}(\mathbf{x}_{\alpha})\sqrt{\frac{2\pi\omega_{0}}{\hbar}}, \quad (16)$$

is the dipole coupling element for the quantized electric field. $\boldsymbol{\theta}$ is the dipole matrix element for the lasing transition, and $\boldsymbol{\Psi}_{0}(x)$ is the normalized mode profile, $\int |\boldsymbol{\Psi}_{0}|^{2}\mathrm{d}x = 1$, of the electric field at frequency ω_{0}. The correlation function for the total effective Langevin force can then be calculated [64] as,

$$\langle F_{\text{tot}}^{\dagger}(t) F_{\text{tot}}(t') \rangle = \delta(t - t')(\gamma_{c} n_{\text{th}}$$
$$+ \sum_{\alpha} \frac{|g_{\alpha}|^{2}}{\gamma_{\perp}^{2}} \left[\gamma_{\perp}(1 + \langle d_{\alpha} \rangle) + \frac{\gamma_{\parallel}}{2}(d_{0} - \langle d_{\alpha} \rangle) \right]), \quad (17)$$

in which nth is the number of thermal photons at the lasing frequency, $\gamma_{\perp}$ is the dephasing rate of the polarization, and d_{α}

is the inversion of the αth atom, which would relax to the steady-state value of d_{0} in the absence of an electric field at the rate $\gamma_{\parallel}$. We have also simplified by assuming the lasing signal to be on resonance with the atomic transition, $\omega_{0} = \omega_{a}$, and that the dipole matrix element and field are co-linear, $\boldsymbol{\theta} \cdot \boldsymbol{\Psi}_{0} = \theta\Psi_{0}$.

The first term in (17) corresponds to the contribution from thermal and zero point fluctuations from outside the cavity to the field inside the cavity; as such it is controlled by the decay rate of the cavity, γ_{c}. The second and third terms stem from spontaneous emission events leading to the dephasing of the atomic polarization, and these terms depend on the parameters of the gain medium. For optical frequencies, the number of thermal photons present is negligible, so the direct field fluctuations do not have any effect on resulting linewidth. Moreover, for most lasers (class A and B) [65], $\gamma_{\parallel} \ll \gamma_{\perp}$, so that only the second term in (17) remains,

$$\langle F_{\text{tot}}^{\dagger}(t) F_{\text{tot}}(t') \rangle = \delta(t - t') \frac{1}{\gamma_{\perp}} \sum_{\alpha} |g_{\alpha}|^{2}(1 + \langle d_{\alpha} \rangle). \quad (18)$$

Assuming that the atoms are uniformly distributed over the cavity, the sum over the atoms in the cavity can be converted into an integral

$$\langle F_{\text{tot}}^{\dagger}(t) F_{\text{tot}}(t') \rangle = \delta(t - t') \frac{2|g'|^{2}}{\gamma_{\perp}}$$
$$\times \int \left(\frac{N_{2}(x)}{N_{2}(x) - N_{1}(x)} \right) D(x) |\Psi_{0}(x)|^{2}\mathrm{d}x, \quad (19)$$

where $N_{2}(x)$ ($N_{1}(x)$) is the density of atoms in the upper (lower) lasing level, and $g_{\alpha} = g'\Psi_{0}(x_{\alpha})$. Equation (19) now expresses the Langevin force in terms of the incomplete inversion factor [63],

$$n_{\text{sp}}(x) = \frac{N_{2}(x)}{N_{2}(x) - N_{1}(x)}. \quad (20)$$

Traditionally at this step the atomic populations are approximated to be uniform in space which clamps the inversion to its value at the lasing threshold, $D(x) \approx D_{0} = D_{0}^{(\text{thr})}$ [64]. In this approximation, the integral in (19) becomes just $D_{0}^{(\text{thr})} n_{\text{sp}}$, and can be related back to the cavity decay rate by

$$\frac{D_{0}^{(\text{thr})} |g'|^{2}}{\gamma_{\perp}} = \frac{\gamma_{c}}{2}, \quad (21)$$

which states that at the lasing threshold, gain must equal loss. Using this approximation contributes one factor of γ_{c} in the standard ST linewidth formula, (13). However, this is a rather crude approximation. Due to spatial hole-burning the atomic populations generally vary substantially in space, and as a consequence the overall inversion is not clamped at the threshold value.

Using (21) obscures the fact that the physical origin of the noise is spontaneous emission in the gain medium. It then is more natural to relate the linewidth to the rate of *spontaneous emission*, γ_{sp}, which can be calculated from the rate of

stimulated emission, γ_{SE} [15], as

$$\gamma_{\text{sp}}(x) = \frac{\gamma_{\text{SE}}(x)}{2n_{\text{p}}} = \frac{2|g'|^2}{\gamma_{\perp}} |\Psi_0(x)|^2, \tag{22}$$

where n_{p} is the number of photons in the steady-state lasing field, and is related to the intensity by $I_0 = 2\pi\hbar\omega_0 n_{\text{p}}$. Thus, we can rewrite (19) in terms of the rate of spontaneous emission, keeping the full space dependence of the problem, as

$$\langle F_{\text{tot}}^{\dagger}(t) F_{\text{tot}}(t') \rangle = \delta(t - t') \int \gamma_{\text{sp}}(x) D(x) n_{\text{sp}}(x) \mathrm{d}x. \tag{23}$$

As noted above, the coefficient of the Langevin noise determines the laser linewidth by standard manipulations. Using the more correct form of this coefficient above leads to a quantum-limited linewidth,

$$\delta\omega = \frac{\int \gamma_{\text{sp}}(x) D(x) n_{\text{sp}}(x) \mathrm{d}x}{2n_{\text{p}}}. \tag{24}$$

Typically, the next step is to relate this to the total output power of the cavity via

$$P_{\text{ST}} = \hbar\omega\gamma_{\text{c}} n_{\text{p}}, \tag{25}$$

which is also only an approximate relation, neglecting spatial variations of the intensity. Using both of the approximations mentioned above and some other simplifications valid for high Q cavities (and neglecting the amplitude contribution contained in the α factor), one then arrives at the standard ST linewidth form, (13).

However, the actual power emitted can be calculated using the true spatially varying fields using Poynting's theorem [66], as

$$P = \frac{I_0\omega_0}{2\pi} \int_{\text{c}} \text{Im}[-\varepsilon] |\Psi_0(x)|^2 \mathrm{d}x. \tag{26}$$

Using the specific form of the imaginary part of the dielectric function for a two-level gain medium (and using real units, not SALT units), the linewidth can then be written as

$$\delta\omega = \frac{\hbar\omega_0}{2P} \int \gamma_{\text{sp}}(x) D(x) n_{\text{sp}}(x) \mathrm{d}x \\ \times \int \gamma_{\text{sp}}(x) D(x) \mathrm{d}x. \tag{27}$$

Here we have derived a more physically intuitive and accurate form of the laser linewidth formula in which the critical role of spontaneous emission in the gain medium leaps out: the linewidth is determined by the spatially varying, mode-dependent spontaneous emission rate and by the steady-state saturated inversion in the presence of spatial hole-burning. Only if one makes the approximation of $\int \gamma_{\text{sp}} D \mathrm{d}x \approx \gamma_{\text{c}}$ does this reduce to the standard ST formula, and in making this substitution one tends to hide the true origin of the linewidth, which is spontaneous emission.

7.2. N-SALT laser linewidth

The previous derivation of the laser linewidth is useful for its relative simplicity, but it fails to take into account a few well known corrections, such as those stemming from relaxation oscillations of the gain medium (α factor) [23, 67], the openness of the cavity (Petermann factor) [56, 57, 68–70], and dispersion (bad-cavity factor) [63, 71–74]. To include these effects, Pick *et al* instead begin with the wave equation, (14), and assume that the known SALT lasing mode solutions to the semiclassical problem remains relatively unchanged in the presence of noise [17], expanding the electric field as

$$\mathbf{E}(\mathbf{x}, t) = \sum_{\mu} a_{\mu}(t) \mathbf{\Psi}_{\mu}(\mathbf{x}) \mathrm{e}^{-\mathrm{i}\omega_{\mu}t}. \tag{28}$$

Although the total noise source $\tilde{\mathbf{F}}_{\text{tot}}(\mathbf{x}, t)$ is broadband, only the noise near the lasing frequency is strongly amplified, and thus the total noise source is also expanded as

$$\tilde{\mathbf{F}}_{\text{tot}}(\mathbf{x}, t) = \sum_{\mu} \mathbf{F}_{\mu}(\mathbf{x}, t) \mathrm{e}^{-\mathrm{i}\omega_{\mu}t}. \tag{29}$$

This allows for one to write down a set of coupled mode equations for the modal amplitudes

$$\dot{a}_{\mu} = \sum_{\nu} C_{\mu\nu}\left(\left|a_{\nu}(t)\right|^2 - \left|a_{\nu 0}\right|^2\right) a_{\mu}(t) + f_{\mu}(t), \tag{30}$$

that are dependent upon the projections of spontaneous emission noise onto the corresponding lasing modes, f_{μ}, and are coupled through the self-saturation and cross-saturation coefficients $C_{\mu\nu}$. These coefficients are related to the spatially dependent parameters of the system through the Green's function of (14), such that for class A lasers

$$C_{\mu\nu} = \frac{-\mathrm{i}\omega_{\mu}^2 \int_C \frac{\partial \varepsilon(\omega_{\mu})}{\partial |a_{\nu}|^2} \mathbf{\Psi}_{\mu}^2 \mathrm{d}\mathbf{x}}{\int_C \frac{\partial}{\partial\omega}\left[\omega^2\varepsilon(\omega)\right]_{\omega_{\mu}} \mathbf{\Psi}_{\mu}^2 \mathrm{d}\mathbf{x}}, \tag{31}$$

$$f_{\mu}(t) = \frac{\mathrm{i}\int_C \mathbf{\Psi}_{\mu} \cdot \mathbf{F}_{\mu}(t) \mathrm{d}\mathbf{x}}{\int_C \frac{\partial}{\partial\omega}\left[\omega^2\varepsilon(\omega)\right]_{\omega_{\mu}} \mathbf{\Psi}_{\mu}^2 \mathrm{d}\mathbf{x}}. \tag{32}$$

This same method is also applicable to class B lasers, however the modal response is no longer instantaneous and (32) becomes an integro-differential equation in time with contributions from the recent history of the medium. The main effect of the delayed response is to introduce relaxation oscillations into the spectrum.

The total phase diffusion of the modal amplitudes can then be calculated by considering $\langle f_{\mu}(t) f_{\nu}^*(t') \rangle \sim \delta_{\mu\nu}\delta(t - t')$, which can be determined from the auto-correlation of the total Langevin noise, and is given by the FD theorem in terms of Im[χ]. The resulting N-SALT linewidth for a single mode laser has the same general features as the simpler result of (27), but it now contains all of the known corrections discussed

above, and can be written as [17],

$$\delta\omega_{N-SALT} = \frac{\hbar\omega_0}{2P}\left(1 + \tilde{\alpha}^2\right)$$

$$\times \frac{\omega_0^2 \int \operatorname{Im}[\varepsilon]\left|\Psi_0\right|^2 d\mathbf{x} \int \operatorname{Im}[\varepsilon]\frac{N_2(\mathbf{x})}{D(\mathbf{x})}\left|\Psi_0\right|^2 d\mathbf{x}}{\left|\int \Psi_0^2\left(\varepsilon + \frac{\omega_0}{2}\frac{d\varepsilon}{d\omega}\big|_{\omega_0}\right)d\mathbf{x}\right|^2}, \quad (33)$$

where Ψ_0 is the normalized SALT lasing mode, and $\tilde{\alpha} = \operatorname{Im}[C]/\operatorname{Re}[C]$ is a generalization of the Lax–Henry α factor. Very recently this formula has been quantitatively confirmed using first-principles numerical simulations [18] for laser parameters where it predicts significant departures from the standard (corrected) form of the ST formula. Similar to the discussion in section 7.1, note that in (33) the cavity decay rate does not naturally appear, while the imaginary part of the response of the gain medium plays a critical role.

The most striking feature of the N-SALT linewidth formula is that it contains all of the different linewidth corrections due openness, incomplete inversion and gain dispersion in an interconnected form. Prior to the development of SALT-based linewidth theories [17, 50, 51], the most general form of the laser linewidth known was based upon the original ST result multiplied by all of these corrections independently (the incomplete inversion factor, n_{sp}, Petermann factor, K, the bad-cavity factor, B), as well as the α factor correction:

$$\delta\omega_{ST} = \frac{\hbar\omega_0\gamma_c^2}{2P}n_{sp}KB\left(1 + \alpha^2\right). \quad (34)$$

To illustrate this difference, consider the Petermann factor originally derived as [68],

$$K = \left|\frac{\int_C \left|\phi_0(\mathbf{x})\right|^2 d\mathbf{x}}{\int_C \phi_0^2(\mathbf{x}) d\mathbf{x}}\right|^2, \quad (35)$$

in which ϕ_0 is the passive cavity mode profile. These factors of the mode profile are all present in (33) (corrected to the lasing mode profiles Ψ_0), but do not exist in isolation. Rather, the field profiles serve as the weighting function in the integrals over the other correction factors within the cavity, such as the bad-cavity factor [63, 71–74],

$$B = \left(\frac{\gamma_\perp}{\gamma_\perp + \frac{\gamma_c}{2}}\right)^2 = \left|\frac{1}{1 + \frac{\omega_0}{2}\frac{\partial\varepsilon}{\partial\omega}\big|_{\omega_0}}\right|^2, \quad (36)$$

as well as what would otherwise be the cavity decay rate, $\gamma_c \approx \omega_0 \operatorname{Im}[\varepsilon]/\varepsilon$. Hence all the space dependence which arises from nonlinear spatial hole-burning is taken into account within the line width formula for the first time. The interconnected nature of these effects means that except for certain limits, we cannot consider these effects to be separate.

The second improvement that N-SALT provides over previous theories is its use of the lasing mode profile rather than the passive cavity mode profile. While the difference between these two profiles is negligible for high-Q cavities,

for systems such as random lasers the difference can be pronounced due to strong line-pulling effects from the gain medium in such a low Q cavity. Using the proper lasing mode profile also affects the α factor due to its dependence upon C. Thus, despite the fact that the N-SALT linewidth prediction and numerical confirmation have yielded no qualitatively new physical effects in the behavior of the laser linewidth, the N-SALT result represents a significant development in laser physics, unifying these previously disparate theories.

8. Conclusion

In summary we have presented a scattering framework for semiclassical laser theory and showed how precisely it leads to zero laser linewidth, due to the discreteness of the poles of the scattering matrix. This approach treats the full system as infinite from the beginning and does not introduce a finite 'Universe' and impose a limiting procedure and damping to recover the infinite system result, although such an approach is a valid alternative method if the LAP is observed. We were inspired to analyze this question by the classic work of Lang *et al* which introduced the notion of 'modes of the Universe'. Our scattering point of view underlies recently developed *ab initio* laser theory (SALT), which leads to efficient computational tools for calculating steady-state lasing properties. A generalization of the *ab initio* theory to treat noise allows accurate calculation of the intrinsic linewidth due to quantum fluctuations when one goes beyond semiclassical laser theory. We emphasize that the quantum-limited linewidth depends directly on the properties of the gain medium in the active lasing state, and it may be useful to write it in a form which does not involve the passive cavity decay rate, but rather emphasizes the spontaneous emission rate.

Acknowledgments

Appendix A. Review of SALT

In this appendix we will provide a brief review of SALT to assist in understanding the discussion in section 4, and the results presented in section 5. However, the interested reader is referred to works dedicated to the SALT and the CF basis [9, 11, 12], and its elaborations to treat amplified modes [16], and complex gain media [14, 15]. The SALT equations find the steady-state solutions to the semiclassical Maxwell–Bloch equations, for which the wave equation was given above in

(2). For this brief review we will only consider the simplest case of two-level atomic gain media, which leads to the Bloch equations for the evolution of the atomic polarization: inversion are

$$\partial_t \mathbf{P}_g^+(\mathbf{x}, t) = -\left(\gamma_\perp + i\omega_a\right)\mathbf{P}_g^+ - \frac{id}{\hbar}(\boldsymbol{\theta} \cdot \mathbf{E})\boldsymbol{\theta}^*, \quad \text{(A.1)}$$

$$\partial_t D(\mathbf{x}, t) = -\gamma_\parallel\big(D - D_0(\mathbf{x})\big) - \frac{2}{i\hbar}\big(\mathbf{P}_g^- - \mathbf{P}_g^+\big) \cdot \mathbf{E}. \quad \text{(A.2)}$$

For an electric field with a single frequency component, or operating in the stable multi-mode regime, we can expand the electric field over its N constituent frequency components as

$$\mathbf{E}^+(\mathbf{x}, t) = \sum_\sigma^N \boldsymbol{\Psi}_\sigma(\mathbf{x})e^{-i\omega_\sigma t}, \quad \text{(A.3)}$$

where $\mathbf{E}^+$ is the positive frequency component of the electric field, $\mathbf{E} = 2\,\mathrm{Re}[\mathbf{E}^+]$. In this expansion we make no distinction between lasing modes, which are associated with the poles of the scattering matrix, and amplified modes due to injected signals, which are externally generated.

The critical assumption in SALT is the stationary inversion approximation (SIA), which states that $\partial_t D = 0$, even when multiple frequencies in the electric field are present; i.e. the populations do not respond to beating terms between the modes, that would lead to four wave mixing and frequency comb generation. This assumption requires that the non-radiative relaxation time of the gain medium is much slower than these beating terms, $\gamma_\parallel \ll \Delta$, where Δ is the minimum frequency difference $\Delta = \omega_\sigma - \omega_\nu$ present in the set of frequencies comprising the electric field, $\{\omega_\sigma\}$, and is satisfied by nearly all microcavity lasers. In the parlance developed for the study of temporal chaotic laser behavior, the SIA is valid for class A and B lasers, but not class C [65, 75]. For class B lasers, which include most semiconductor based devices, the mode spacing must also be well separated from the relaxation oscillation frequency, so as to avoid resonantly driving relaxation fluctuations which could destabilize the multimode solution. The relaxation oscillation frequency is $\omega_{RO} \sim \sqrt{\kappa\gamma_\parallel}$ [75], where $\kappa = \gamma_c/2$ is the field decay rate of the cavity. For microcavities without an injected signal, $\kappa \leqslant \Delta_{FSR}$, the mode spacing from the free spectral range of the cavity, so $\omega_{RO} \leqslant \sqrt{\Delta_{FSR}\gamma_\parallel} < \Delta_{FSR}$, and as such, the SIA is still valid when $\gamma_\parallel < \Delta_{FSR}$ [16]. Furthermore, SALT has been rigorously tested using direct FDTD simulations of the Maxwell–Bloch equations, with excellent quantitative agreement found [13–16].

As the SIA removes any time dependence from the inversion, the polarization must now contain exactly the same frequency components as the electric field

$$\mathbf{P}_g^+(\mathbf{x}, t) = \sum_\sigma^N \mathbf{p}_\sigma(\mathbf{x})e^{-i\omega_\sigma t}. \quad \text{(A.4)}$$

Using the multimode expansion for the electric field, (A.3), the polarization equation, (A.1), can be solved independently for each frequency component, providing an expression for $\mathbf{p}_\sigma$

in terms of D and Ψ_σ, and inserted into the wave equation as

$$\left[\nabla \times \nabla \times - \left(\varepsilon_c(\mathbf{x}) + \frac{4\pi\,|\theta|^2 D(\mathbf{x})}{\hbar\left(\omega_\sigma - \omega_a + i\gamma_\perp\right)}\right)k_\sigma^2\right]\Psi_\sigma(\mathbf{x})$$
$$= 0, \quad \text{(A.5)}$$

which explicitly identifies the gain due to the inversion of the atomic medium, and in which θ has been taken to be collinear to the electric field. Likewise, using (A.2), we can solve directly for the inversion as

$$D(\mathbf{x}) = \frac{D_0(\mathbf{x})}{1 + \frac{4\,|\theta|^2}{\hbar^2\gamma_\perp\gamma_\parallel}\sum_\sigma^N \Gamma_\sigma\,|\Psi_\sigma|^2}, \quad \text{(A.6)}$$

in which the denominator contains the effects of gain-saturation due to the electric field within the cavity, and $\Gamma_\sigma = \gamma_\perp^2/((\omega_\sigma - \omega_a)^2 + \gamma_\perp^2)$ is the Lorentzian gain curve. Together, (A.5) and (A.6) form the fundamental SALT equations (for the two level gain case): N self-consistent, differential equations for the unknown spatial profiles of the N modes in the electric field at N unknown frequencies, coupled through the nonlinear spatial hole-burning in the inversion. These are to be solved with outgoing Sommerfeld boundary conditions at infinity, which can only be satisfied at discrete frequencies. The SALT equations also reveal the natural units for the electric field and atomic inversion, $E_{SALT} = 2|\theta|/(\hbar\sqrt{\gamma_\perp\gamma_\parallel})E$, and $D_{SALT} = (4\pi|\theta|^2/\hbar\gamma_\perp)D$.

The numerical solution of the SALT equations is performed by expanding the mode profile functions over a basis

$$\Psi_\sigma(\mathbf{x}) = \sum_n a_n^{(\sigma)}\mathbf{f}_n(\mathbf{x}; \omega_\sigma), \quad \text{(A.7)}$$

and then using a nonlinear solver to find the expansion coefficients $a_n^{(\sigma)}$. There are two different basis sets that have been used to solve the SALT equations. The first, and more developed, is based on a set of CF states [9, 12], which are defined as

$$\left[-\nabla \times \nabla \times + \left(\varepsilon_c(\mathbf{x}) + \eta_n F(\mathbf{x})\right)\mathbf{k}^2\right]\mathbf{u}_n(\mathbf{x}; \omega) = 0, \quad \text{(A.8)}$$

and satisfy the outgoing boundary condition at frequency ω at the surface of last scattering. Here, $\eta_n(\omega)$ and $\mathbf{u}_n(\mathbf{x}; \omega)$ are the eigenvalue and eigenstates of the CF equation, and $D_0(\mathbf{x}) = D_0 F(\mathbf{x})$ is the profile of the pump. A key aspect of the SALT approach is that for any given frequency (wavevector) the CF states form a complete biorthogonal basis for an arbitrary purely outgoing solution; hence they can expand any solution to the SALT equations.

Moreover, by observing the similarities between (A.8) and (A.5), it is clear that just above the first lasing threshold where the spatial hole-burning in the inversion is negligible, the TLM is given by a single CF basis state at the lasing frequency satisfying

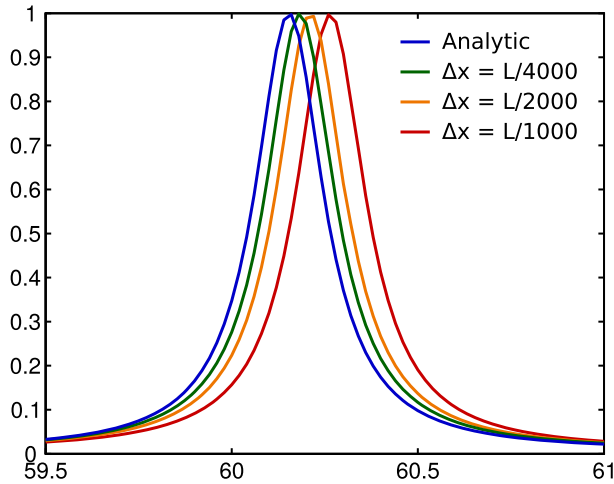$$\eta(\omega) = \frac{\gamma_\perp D_0}{\omega - \omega_a + i\gamma_\perp}. \quad \text{(A.9)}$$

Figure B1. Plot of the transmitted intensity of an injected signal with frequency ω_{in}, into the passive, two-sided cavity with partially reflecting mirrors ($R = 0.8$) on both ends studied in figure 4. Different colors represent different choices of resolution, Δx (red, orange, and green), or the analytic solution (blue).

Hence above threshold, the CF basis represents a numerically efficient basis for solving the SALT equations, as each lasing mode can be represented by only a few CF states. The Sommerfeld condition is implemented by matching conditions on the superposition at the last scattering surface to purely outgoing free hermitian solutions. The CF equation also facilitates solving the SALT equations for amplified modes by yielding an additional basis, $\mathbf{v}_m(\mathbf{x}; \omega)$, $\beta_m(\omega)$, which are also defined by (A.8), but satisfy a purely incoming boundary condition instead, thus representing a signal with a CF into the cavity.

Recently an alternative approach for solving the SALT equations has been developed, using a position space basis, and a truncated simulation region bordered by a perfectly matched layer [58]. This follows the conceptual approach discussed in detail at the end of section 6. The Sommerfeld condition is not imposed by matching at a last scattering surface, but rather by the presence of the PML. In general, the CF methods allow for more elegant theoretical analysis, while the position space basis is ultimately expected to be more efficient computationally. In both approaches a nonlinear iteration scheme will need to be implemented to find the correct self-consistent solution in the basis of choice.

The SALT modal wave equation, (A.5) is satisfied by both lasing modes and injected signals, which interact and compete with each other for gain through the nonlinear spatial hole-burning denominator of (A.6). There are two major differences between lasing modes and injected signals though; the former must satisfy the Sommerfeld radiation condition, while the latter satisfies a mixed boundary condition with both incoming and outgoing terms. As a direct consequence of this, the frequency of a lasing mode needs to be determined self-consistently from the outgoing boundary condition itself, while the frequency of an injected signal is externally fixed, and the mixed boundary condition can be satisfied at any frequency.

Appendix B. Partially reflecting mirrors in SALT

Traditionally, SALT has been used to simulate semiconductor based microlasers, in which the optical confinement for one-dimensional systems is provided by Fresnel reflection due to the dielectric mismatch across the cleaved facets of the device. However, for even the materials with the largest index of refraction, the optical confinement due to Fresnel reflection is weak, $R < 0.36$ for $n < 4$. Thus to simulate high finesse systems, one must either introduce a geometric solution, such as a Bragg reflector, or a partially reflecting mirror. In the results presented in section 5 we chose to use a partially reflecting mirror, as we have the additional motivation of simulating a similar system as described by Lang *et al* and others [7, 53–55]. In this appendix we briefly detail how to implement a partially reflecting mirror in SALT simulations.

Mathematically, a partially reflecting mirror can be described by a delta function in the passive cavity dielectric function

$$\varepsilon_{\text{c,total}}(x) = \varepsilon_{\text{c}}(x) + \sum_m \Lambda_m \delta(x - x_m), \qquad \text{(B.1)}$$

where Λ_m represents the strength of the mth mirror at x_m. For a single mirror, it is straightforward to show that the reflection and transmission coefficients are

$$R = \frac{\left(\frac{k\Lambda}{2}\right)^2}{1 + \left(\frac{k\Lambda}{2}\right)^2}, \qquad \text{(B.2)}$$

$$T = \frac{1}{1 + \left(\frac{k\Lambda}{2}\right)^2}, \qquad \text{(B.3)}$$

which demonstrates that the figure of merit for the reflectivity of a partially reflecting mirror is $k\Lambda$. There are two different ways to implement this delta-function in SALT, either by incorporating the partially reflecting mirrors directly into the passive cavity dielectric, as indicated above, or for mirrors at the cavity boundary, adjusting the outgoing boundary condition [76]. The former method is numerically more stable, and is implemented in a discretized system as

$$\varepsilon_{\text{c,total}}(x_n) = \varepsilon_{\text{c}}(x_n) + \sum_m \frac{\Lambda_m}{\Delta x} \delta_{x_n, x_m}, \qquad \text{(B.4)}$$

where δ_{ij} is the Kronecker delta, x_n is the spatial position of the nth pixel in the system, and Δx is the resolution. Despite the fact that the reflectivity of such a mirror is dependent upon $\zeta = k\Lambda$, when using the CF basis to solve the SALT equations, one can choose ζ to be a constant, and then adjust Λ accordingly to achieve a system with reflectivity independent of the frequency, something which is not possible in time-domain simulation techniques.

We can confirm that the discretized representation of a partially reflecting mirror used in conjunction with the SALT algorithm is correct by comparing it against the analytic

result. In figure B1 we show the transmission resonance for the same passive cavity shown in figure 4 when calculated analytically (blue), and simulated using I-SALT with increasing resolution (red, orange, and green). The effect of discretizing the system is to shift the cavity resonances slightly, and as the resolution is increased, the location of the cavity resonance converges to the analytic value, as shown. However, regardless of the spatial resolution, the Lorentzian shape of the cavity resonance remains unchanged.

References

[1] Haken H and Sauermann H 1963 Nonlinear interaction of laser modes *Z. Phys.* **173** 261–&

[2] Lamb W E 1964 Theory of an optical maser *Phys. Rev.* **134** A1429–50

[3] Lax M 1966 Quantum noise: IV. Quantum theory of noise sources *Phys. Rev.* **145** 110–29

[4] Scully M O and Lamb W E 1967 Quantum theory of an optical maser: I. General theory *Phys. Rev.* **159** 208–26

[5] Graham R and Haken H 1968 Quantum theory of light propagation in a fluctuating laser-active medium *Z. Phys.* **213** 420–&

[6] Graham R and Haken H 1970 Functional quantum statistics of light propagation in a 2-level system *Z. Phys.* **235** 166–&

[7] Lang R, Scully M O and Lamb W E 1973 Why is the laser line so narrow? A theory of single-quasimode laser operation *Phys. Rev.* A **7** 1788–97

[8] Fox A and Li T 1961 Resonant modes in a maser interferometer *Bell Syst. Tech. J.* **40** 453–+

[9] Türeci H E, Stone A D and Collier B 2006 Self-consistent multimode lasing theory for complex or random lasing media *Phys. Rev.* A **74** 043822

[10] Türeci H E, Stone A D and Ge L 2007 Theory of the spatial structure of nonlinear lasing modes *Phys. Rev.* A **76** 013813

[11] Türeci H E, Ge L, Rotter S and Stone A D 2008 Strong interactions in multimode random lasers *Science* **320** 643–6

[12] Ge L, Chong Y D and Stone A D 2010 Steady-state *ab initio* laser theory: generalizations and analytic results *Phys. Rev.* A **82** 063824

[13] Ge L, Tandy R J, Stone A D and Türeci H E 2008 Quantitative verification of *ab initio* self-consistent laser theory *Opt. Express* **16** 16895–902

[14] Cerjan A, Chong Y D, Ge L and Stone A D 2012 Steady-state *ab initio* laser theory for *n*-level lasers *Opt. Express* **20** 474–88

[15] Cerjan A, Chong Y D and Stone A D 2015 Steady-state *ab initio* laser theory for complex gain media *Opt. Express* **23** 6455–77

[16] Cerjan A and Stone A D 2014 Steady-state *ab initio* theory of lasers with injected signals *Phys. Rev.* A **90** 013840

[17] Pick A, Cerjan A, Liu D, Rodriguez A, Stone A D, Chong Y D and Johnson S 2015 *Ab initio* multimode linewidth theory for arbitrary inhomogeneous laser cavities *Phys. Rev.* A **91** 063806

[18] Cerjan A, Pick A, Chong Y D, Johnson S G and Stone A D 2015 Quantitative test of general theories of the intrinsic laser linewidth *Opt. Express* **23** 28316–40

[19] Spencer M B and Lamb W E 1972 Laser with a transmitting window *Phys. Rev.* A **5** 884–92

[20] Fu H and Haken H 1987 Semiclassical dye-laser equations and the unidirectional single frequency operation *Phys. Rev.* A **36** 4802–16

[21] Fu H and Haken H 1988 Semiclassical theory of dye lasers: the single-frequency and multifrequency steady states of operation *J. Opt. Soc. Am.* B **5** 899–908

[22] Fu H and Haken H 1991 Multifrequency operations in a short-cavity standing-wave laser *Phys. Rev.* A **43** 2446–54

[23] Henry C 1986 Theory of spontaneous emission noise in open resonators and its application to lasers and optical amplifiers *J. Lightwave Technol.* **4** 288–97

[24] Haken H 1984 *Laser Theory* (Berlin: Springer)

[25] Siegman A E 1986 *Lasers* (Mill Valley, CA: University Science Books)

[26] Sargent M, Scully M O and Lamb W E 1978 *Laser Physics* (Cambridge, MA: Westview Press)

[27] Levi A F J, Slusher R E, McCall S L, Tanbun-Ek T, Coblentz D L and Pearton S J 1992 Room temperature operation of microdisc lasers with submilliamp threshold current *Electron. Lett.* **28** 1010–2

[28] McCall S L, Levi A F J, Slusher R E, Pearton S J and Logan R A 1992 Whispering gallery mode microdisk lasers *Appl. Phys. Lett.* **60** 289–91

[29] Collot L, Lefèvre-Seguin V, Brune M, Raimond J M and Haroche S 1993 Very high-q whispering-gallery mode resonances observed on fused silica microspheres *Europhys. Lett.* **23** 327

[30] Armani D K, Kippenberg T J, Spillane S M and Vahala K J 2003 Ultra-high-q toroid microcavity on a chip *Nature* **421** 925–8

[31] Gmachl C, Capasso F, Narimanov E E, Nöckel J U, Stone A D, Faist J, Sivco D L and Cho A Y 1998 Highpower directional emission from microlasers with chaotic resonators *Science* **280** 1556–64

[32] Lacey S, Wang H, Foster D H and Nöckel J U 2003 Directional tunneling escape from nearly spherical optical resonators *Phys. Rev. Lett.* **91** 033902

[33] Redding B, Cerjan A, Huang X, Lee M L, Stone A D, Choma M A and Cao H 2015 Low spatial coherence electrically pumped semiconductor laser for speckle-free full-field imaging *Proc. Natl Acad. Sci. USA* **112** 1304–9

[34] Painter O, Lee R K, Scherer A, Yariv A, O'Brien J D, Dapkus P D and Kim I 1999 Two-dimensional photonic band-gap defect mode laser *Science* **284** 1819–21

[35] Meier M, Mekis A, Dodabalapur A, Timko A, Slusher R E, Joannopoulos J D and Nalamasu O 1999 Laser action from two-dimensional distributed feedback in photonic crystals *Appl. Phys. Lett.* **74** 7–9

[36] Imada M, Noda S, Chutinan A, Tokuda T, Murata M and Sasaki G 1999 Coherent two-dimensional lasing action in surface-emitting laser with triangular-lattice photonic crystal structure *Appl. Phys. Lett.* **75** 316–8

[37] Cao H, Zhao Y G, Ho S T, Seelig E W, Wang Q H and Chang R P H 1999 Random laser action in semiconductor powder *Phys. Rev. Lett.* **82** 2278–81

[38] Cao H, Xu J Y, Seelig E W and Chang R P H 2000 Microlaser made of disordered media *Appl. Phys. Lett.* **76** 2997–9

[39] Polson R C, Chipouline A and Vardeny Z V 2001 Random lasing in -conjugated films and infiltrated opals *Adv. Mater.* **13** 760–4

[40] Bahoura M and Noginov M A 2003 Determination of the transport mean free path in a solid-state random laser *J. Opt. Soc. Am.* B **20** 2389

[41] van der Molen K L, Tjerkstra R W, Mosk A P and Lagendijk Ad 2007 Spatial extent of random laser modes *Phys. Rev. Lett.* **98** 143901

[42] Wiersma D S and Cavalieri S 2001 Light emission: a temperature-tunable random laser *Nature* **414** 708–9

[43] Wu X, Fang W, Yamilov A, Chabanov A A, Asatryan A A, Botten L C and Cao H 2006 Random lasing in weakly scattering systems *Phys. Rev.* A **74** 053812

[44] Collett M J and Gardiner C W 1984 Squeezing of intracavity and traveling-wave light fields produced in parametric amplification *Phys. Rev.* A **30** 1386–91

[45] Gardiner C W and Collett M J 1985 Input and output in damped quantum systems: quantum stochastic differential equations and the master equation *Phys. Rev.* A **31** 3761–74

[46] Glauber R J and Lewenstein M 1991 Quantum optics of dielectric media *Phys. Rev.* A **43** 467–91

[47] Hackenbroich G, Viviescas C and Haake F 2002 Field quantization for chaotic resonators with overlapping modes *Phys. Rev. Lett.* **89** 083902

[48] Viviescas C and Hackenbroich G 2003 Field quantization for open optical cavities *Phys. Rev.* A **67** 013805

[49] Hackenbroich G, Viviescas C and Haake F 2003 Quantum statistics of overlapping modes in open resonators *Phys. Rev.* A **68** 063805

[50] Chong Y D and Stone A D 2012 General linewidth formula for steady-state multimode lasing in arbitrary cavities *Phys. Rev. Lett.* **109** 063902

[51] Pillay J C, Natsume Y, Stone A D and Chong Y D 2014 Generalized sub-Schawlow–Townes laser linewidths via material dispersion *Phys. Rev.* A **89** 033840

[52] Sommerfeld A 1949 *Partial Differential Equations in Physics* (New York: Academic)

[53] Baseia B and Nussenzveig H M 1984 Semiclassical theory of laser transmission loss *Opt. Acta* **31** 39–62

[54] Penaforte J C and Baseia B 1984 Quantum theory of a one-dimensional laser with output coupling: linear approximation *Phys. Rev.* A **30** 1401–6

[55] Dutra S M and Knight P L 1996 Spontaneous emission in a planar Fabry–Pérot microcavity *Phys. Rev.* A **53** 3587–605

[56] Hamel W A and Woerdman J A 1989 Nonorthogonality of the longitudinal eigenmodes of a laser *Phys. Rev.* A **40** 2785–7

[57] Hamel W A and Woerdman J A 1990 Observation of enhanced fundamental linewidth of a laser due to nonorthogonality of its longitudinal eigenmodes *Phys. Rev. Lett.* **64** 1506–9

[58] Esterhazy S, Liu D, Liertzer M, Cerjan A, Ge L, Makris K G, Stone A D, Melenk J M, Johnson S G and Rotter S 2014 Scalable numerical approach for the steady-state *ab initio* laser theory *Phys. Rev.* A **90** 023816

[59] Ignatowsky W 1905 Reflexion elektromagnetisches Wellen an einem Draht *Ann. Phys.* **323** 495–522

[60] Èidus D M 1962 On the principle of limiting absorption *Mat. Sb. (N.S.)* **57** 13–44

[61] Schulenberger J R and Wilcox C H 1971 The limiting absorption principle and spectral theory for steady-state wave propagation in inhomogeneous anisotropic media *Arch. Ration. Mech. Anal.* **41** 46–65

[62] Schawlow A L and Townes C H 1958 Infrared and optical masers *Phys. Rev.* **112** 1940–9

[63] Lax M 1966 Quantum noise: V. Phase noise in a homogeneously broadened maser *Physics of Quantum Electronics* ed P L Kelley *et al* (New York: McGraw-Hill)

[64] Haken H 1985 *Light: Laser Dynamics* vol 2 (New York: North-Holland Phys. Publishing)

[65] Ohtsubo J 2007 *Semiconductor Lasers: Stability, Instability, and Chaos* (New York: Springer)

[66] Jackson J D 1998 *Classical Electrodynamics* 3rd edn (New York: Wiley)

[67] Henry C 1982 Theory of the linewidth of semiconductor-lasers *IEEE J. Quantum Electron.* **18** 259–64

[68] Petermann K 1979 Calculated spontaneous emission factor for double-heterostructure injection-lasers with gain-induced waveguiding *IEEE J. Quantum Electron.* **15** 566–70

[69] Haus H A and Kawakami S 1985 On the excess spontaneous emission factor in gain-guided laser-amplifiers *IEEE J. Quantum Electron.* **21** 63–9

[70] Siegman A E 1989 Excess spontaneous emission in non-hermitian optical-systems: II. Laser-oscillators *Phys. Rev.* A **39** 1264–8

[71] Kolobov M I, Davidovich L, Giacobino E and Fabre C 1993 Role of pumping statistics and dynamics of atomic polarization in quantum fluctuations of laser sources *Phys. Rev.* A **47** 1431–46

[72] Kuppens S J M, van Exter M P and Woerdman J P 1994 Quantum-limited linewidth of a bad-cavity laser *Phys. Rev. Lett.* **72** 3815–8

[73] van Exter M P, Kuppens S J M and Woerdman J P 1995 Theory for the linewidth of a bad-cavity laser *Phys. Rev.* A **51** 809–16

[74] Kuppens S J M, van Exter M P, Vanduin M and Woerdman J P 1995 Evidence of nonuniform phase-diffusion in a bad-cavity laser *IEEE J. Quantum Electron.* **31** 1237–41

[75] Arecchi F T, Lippi G L, Puccioni G P and Tredicce J R 1984 Deterministic chaos in laser with injected signal *Opt. Commun.* **51** 308–14

[76] Cerjan A 2015 Fundamental physics and device design using the steady-state *ab initio* laser theory *PhD Thesis* Yale University