The ultimate quantum limit to the linewidth of lasers

H.M. Wiseman
Centre for Laser Science, Department of Physics, The University of Queensland, Queensland 4072 Australia.

The standard quantum limit to the linewidth of a laser for which the gain medium can be adiabatically eliminated is \( \ell_0 = \kappa/2\bar{n} \). Here \( \kappa \) is the intensity damping rate and \( \bar{n} \) the mean photon number. This contains equal contributions from the loss and gain processes, so that simple arguments which attribute the linewidth wholly to phase noise from spontaneous gain are wrong. I show that an unstimulated gain process actually introduces no phase noise, so that the ultimate quantum limit to the laser linewidth comes from the loss alone and is equal to \( \ell_{\text{ult}} = \kappa/4\bar{n} \). I investigate a number of physical gain mechanisms which attempt to achieve gain without phase noise: a linear atom-field coupling with finite interaction time; a nonlinear atom-field coupling; and adiabatic photon transfer using a counterintuitive pulse sequence. The first at best reaches the standard limit \( \ell_0 \), the second reaches \( 4\ell_0 \), while the third reaches the ultimate limit of \( \ell_{\text{ult}} = \frac{1}{2}\ell_0 \).

I. INTRODUCTION

It is more than 40 years since Schawlow and Townes introduced the idea of an “optical maser” [1], now known of course as a laser. Probably the most famous result from this paper is the expression for the quantum-limited laser linewidth, their Eq. (17),

\[
\Delta \omega_{\text{osc}} = \frac{2\hbar \omega}{P_{\text{out}}} (\Delta \omega)^2.
\] (1.1)

Here \( \Delta \omega_{\text{osc}} \) is the half-width at half maximum (HWHM) of the laser, \( \Delta \omega \) is the HWHM of the relevant atomic transition, \( P_{\text{out}} \) is the output power and \( \omega \) the frequency of the laser. Defining \( \ell = 2\Delta \omega_{\text{osc}} \) and \( \gamma = 2\Delta \omega \), this expression can be rewritten

\[
\ell_{\text{ST}} = \frac{\hbar \omega}{P_{\text{out}}} \gamma^2;
\] (1.2)

where ST stands for Schawlow-Townes. The derivation of this expression assumes reabsorption of photons by atoms in the ground state of the relevant transition is negligible, and also ignores thermal photons and other extraneous noise sources.

To accurately describe lasers a number of refinements must be made to the Schawlow-Townes expression [2]. These are discussed in the Appendix. This discussion, I believe, helps to put in perspective some of the past work on quantum limits to the laser linewidth. The end result is that a better expression for the standard quantum limit to the laser linewidth is

\[
\ell_{\text{st}} = \frac{\ell_{\text{bare}}}{2N} \leq \frac{\hbar \omega}{2P_{\text{out}}} \frac{\gamma^2 \kappa^2}{(\gamma + \kappa)^2}.
\] (1.3)

Here st stands for standard (quantum limit). As explained in the Appendix, \( \ell_{\text{bare}} \) is the bare linewidth, \( N \) is the number of coherent excitations stored in the laser mode and its gain medium and \( \kappa \) is the cavity linewidth. The inequality is an equality only for perfectly efficient output coupling.

In the limit \( \gamma \gg \kappa \) the gain medium can be adiabatically eliminated, resulting in Markovian evolution for the laser mode. This means that \( N \) can be replaced by \( \bar{n} \) (the mean photon number), and \( \ell_{\text{bare}} \) by \( \kappa \), to give the standard Markovian limit as

\[
\ell_0 = \frac{\kappa}{2\bar{n}}.
\] (1.4)

For the remainder of this paper I will assume the Markovian limit, and drop the adjective “Markovian” distinguishing \( \ell_{\text{st}} \) from \( \ell_{\text{st}} \) when no confusion is likely to arise.

Most older textbooks [3–5] quote the result in Eq. (1.4), or one which reduces to it in the appropriate limit of neither reabsorption nor thermal photons. The first two of these [3,4] derive this result rigorously using quantum Langevin equations. All three attempt to explain it in terms of the noise added by the spontaneous contribution to the (mostly stimulated) gain of photons from the atomic medium. Loudon [5] even recommends the argument based on the uncertainty principle given by Weichel [6].

The argument of Weichel is as follows (in my notation). In a laser at steady state, the rate of spontaneous emissions to total (spontaneous and stimulated) emissions is \( 1 : \bar{n} + 1 \). Since the total gain rate must equal the total loss rate \( \kappa \bar{n} \), the rate of spontaneous emissions is

\[
A = \frac{\kappa \bar{n}}{1 + \bar{n}} \simeq \kappa,
\] (1.5)

where it is assumed that \( \bar{n} \gg 1 \). Now the reciprocal of this, \( \Delta t = 1/\kappa \), is [6] “the average time between phase fluctuations caused by spontaneous emissions into the mode.” Invoking the uncertainty principle

\[
\Delta E \Delta t \geq \hbar/2,
\] (1.6)

with the energy uncertainty of the mode being \( \Delta E \simeq \bar{n} \hbar \Delta \omega_{\text{osc}} \) gives
which agrees with the Schawlow-Townes result \( \ell = 2 \Delta \omega_{\text{osc}} = \frac{1}{\bar{n} \Delta t} = \frac{\kappa}{\bar{n}} \), \( \text{(1.7)} \)

which is a representation based on anti-normally ordered operator products. That is, it results from using (implicitly in most cases) the Glauber-Sudarshan P function \( \overline{\rho} \) as a true representation of the the fluctuations in the laser mode field. The \( \overline{P} \) function is of course no more fundamental than the \( \overline{Q} \) function \( \overline{\rho} \), which is a representation based on anti-normally ordered statistics. If one were to use the \( \overline{Q} \) function as an aid to intuition, one would find that it is the loss process that is wholly responsible for the phase noise. Of course the rate of phase diffusion would agree with that from the \( \overline{P} \) function, at least in steady-state where loss and gain balance.

If one asks a question about phase diffusion, the only objective answer will come from using the phase basis itself. This is far more difficult than using the more familiar phase-space representations, but some approximate results have been obtained \( \overline{\mathcal{D}} \). These show that, at steady state, the phase diffusion has equal contributions from the loss and gain process. The same result occurs from a Wigner function calculation \( \overline{\mathcal{D}} \). This is not surprising since symmetrically ordered moments are known to closely approximate the true moments for the phase operator for states with well-defined amplitude \( \overline{\mathcal{D}} \).

The fact that phase diffusion comes equally from the loss and gain processes suggests that the standard quantum limit to the laser linewidth, \( \ell_0 \) of Eq. \( \overline{\mathcal{D}} \), may not be the ultimate quantum limit. The contribution from the loss mechanism is unavoidable. A laser, at least in useful definitions \( \overline{\mathcal{D}} \), requires linear damping of the laser mode in order to form an output beam. However, it may be that the standard gain mechanism could be replaced by some other gain mechanism that causes less phase diffusion. The ultimate quantum limit to the laser linewidth could thus be as small as one half of the standard limit.

In this paper I investigate various gain mechanisms in an attempt to find one which causes less phase diffusion than the standard gain mechanism. First, in Sec. II, I review the standard (ideal) model for a laser, giving rise to the standard quantum limit \( \ell_0 \). Next, in Sec. III, I present gain without stimulated emission, which produces a linewidth of \( \ell_0/2 \), and discuss how this can be physically realized. In Secs. IV and V I present models which attempt to approximate gain without stimulated emission, using a micromaser-like interaction and a nonlinear field-atom interaction respectively, and discuss their success. After a comparison of these results in Sec. VI, I conclude in Sec. VII by returning to a derivation of the quantum limits to the laser linewidth using an uncertainty relation.
Here I have included linear loss (allowing the laser output) at rate $\kappa$, and I am using the notation

$$D[c] \equiv J[c] - A[c].$$

As long as $\Gamma^2 < \kappa$, this master equation has a steady state. However it is not an appropriate steady state for the device to be considered a laser. As discussed in Ref. [12], it is necessary to have $\bar{n} \gg 1$ for the output of the device to be coherent (in a quantum statistical sense).

But in this limit, the stationary state of the master equation (2.7) has a photon number uncertainty $\sigma(n) \sim \bar{n}$. This leads to enormous low-frequency ($\sim \kappa/\bar{n}$) fluctuations in the intensity of the output beam. This ruins the second-order coherence of the device, which is why it could not be considered a laser [12].

The origin of the problem with Eq. (2.7) is stimulated emission. Since this is part of the acronym l.a.s.e.r. it could not be considered a laser [12].

It might be thought that this is a good thing, but it actually occurs at Poisson-distributed times, with a rate $\Gamma \ll \Omega \bar{n}$, then a Markovian master equation for the field results. If one also includes linear damping at rate $\kappa$ as above, and lets the gain (the rate of photon addition) be $\Gamma = \kappa \mu$, then one gets

$$\kappa^{-1} \hat{\rho} = \mu \left( J[a]^{\dagger} A[a]^{\dagger} - 1 \right) \rho + D[a] \rho$$

$$= \mu D[a]^{\dagger} A[a]^{\dagger} - 1 \rho + D[a] \rho. \quad (2.13)$$

From Eq. (2.11) and the identity

$$1 = \int_0^\infty du \exp(-uaa^\dagger/2)a^\dagger a \exp(-uaa^\dagger/2) \quad (2.14)$$

it is easy to see that the ideal laser master equation (2.13), first derived in Ref. [13], is of the required Lindblad form [13].

### B. Stationary State

In the fock basis the laser master equation (2.13) is

$$\dot{\rho}_{n,m} = \mu \left( \frac{2\sqrt{nm}}{n+m} \rho_{n-1,m-1} - \rho_{n,m} \right) - \frac{n+m}{2} \rho_{n,m} + \sqrt{(n+1)(m+1)} \rho_{n+1,m+1}. \quad (2.15)$$

Here, as in the remainder of the paper, I have set $\kappa = 1$. Clearly the stationary state will be of the form $\rho_{n,m} = \delta_{n,m} P_n$. The equation of motion for $P_n$ is

$$\kappa^{-1} \dot{P}_n = \mu (P_{n-1} - P_n) + (n+1)P_{n+1} - nP_n. \quad (2.16)$$

This has the stationary solution $P_n = e^{-\mu n}/n!$. That is, the intracavity photon statistics are exactly Poissonian.

The stationary state matrix can therefore be written

$$\rho_{ss} = \sum_n e^{-\mu n} \frac{n!}{n!} |n\rangle \langle n|. \quad (2.17)$$

Equivalently, it can be written

$$\rho_{ss} = \int_0^{2\pi} d\phi |\alpha e^{i\phi}\rangle \langle \alpha e^{i\phi}|, \quad (2.18)$$

where $|\alpha| = \sqrt{\mu}$ and $|\alpha e^{i\phi}\rangle$ is a coherent state of amplitude $|\alpha e^{i\phi}|$. From either expression it is easy to verify that the mean number is $\text{Tr}[a^\dagger a \rho_{ss}] = \mu$ and the mean amplitude $\text{Tr}[a \rho_{ss}] = 0$. 

3
C. Calculating the linewidth

There are many different ways of calculating the linewidth of a laser from its master equation. One way is to covert the master equation into a Fokker-Planck equation for a quasiprobability distribution function such as the $P$, $Q$ or $W$ function [10]. This is relatively straightforward for a master equation of the form (2.13), despite the apparent awkwardness of the inverse superoperator $\mathcal{A}[a^1]^{-1}$ [13]. However, for other master equations as I will consider later in this paper, the conversion is not so simple. Therefore I will adopt a method using the Fock basis. The method is essentially a more rigorous version of that used by Sargent, Scully, and Lamb [4].

The linewidth $\ell$ of a laser I have taken to be the FWHM of the Power spectrum

$$P(\omega) \propto \int_0^{\infty} d\tau g^{(1)}(\tau) \cos \omega \tau, \quad (2.19)$$

where the normalized first-order coherence function is

$$g^{(1)}(\tau) = \langle a^\dagger(t + \tau)a(t) \rangle_{ss}/\langle a^\dagger a \rangle_{ss}. \quad (2.20)$$

If one represents the master equation (2.13) as $\dot{\rho} = \mathcal{L}\rho$ then one can write

$$g^{(1)}(\tau) = \frac{\text{Tr}[a^\dagger e^{\mathcal{L}\tau}(a \rho_{ss})]}{\mu}. \quad (2.21)$$

Note above that the stationary state matrix $\rho_{ss}$ is a mixture of coherent states, as in Eq. (2.18). Since $g^{(1)}(\tau)$ is invariant under a phase shift, Eq. (2.18) implies that in Eq. (2.21) one can take $\rho_{ss} = |\alpha\rangle\langle\alpha|$, with $|\alpha|^2 = \mu$. Then Eq. (2.21) becomes

$$g^{(1)}(\tau) = \frac{\text{Tr}[a^\dagger \alpha \rho(\tau)]}{\mu}, \quad (2.22)$$

where $\rho(t)$ obeys the master equation (2.13) and

$$\rho(0) = |\alpha\rangle\langle\alpha|. \quad (2.23)$$

If one defines

$$f_n(t) = \sqrt{n} \rho_{n-1,n}(t)/\alpha^* \quad (2.24)$$

then one can write

$$g^{(1)}(t) = \sum_n f_n(t). \quad (2.25)$$

Clearly if one can determine the evolution of $f_n(t)$, one can find $g^{(1)}(t)$ and hence the linewidth of the laser. From Eq. (2.13) one finds

$$\dot{f}_n = \mu - \frac{2n}{2n-1} f_{n-1} - \frac{2n-1}{2} f_n. \quad (2.26)$$

Defining

$$r_n(t) = \frac{\mu f_n(t)}{n f_{n+1}(t)}, \quad (2.27)$$

one obtains

$$\dot{f}_n = \left[ \frac{2n(n-1)}{2n-1} r_{n-1} - \mu + \frac{\mu}{r_n - \frac{2n-1}{2}} \right] f_n. \quad (2.28)$$

Now from the definition (2.27), $r_n(0) \equiv 1$. Assuming that this ratio remains unity, expand Eq. (2.28) to leading order in $1/\mu$ to get

$$\dot{f}_n \approx -\frac{1}{4n} f_n. \quad (2.29)$$

Solving this and substituting into Eq. (2.27) gives, to leading order,

$$r_n(t) \approx \exp \left( -\frac{t}{4n^2} \right) \approx 1 - \frac{t}{4n^2}. \quad (2.30)$$

where the expansion to first order is valid for times much less than $\mu^2$. Since, as will be shown, the coherence time $\sim 2/\ell$ is of order $\mu$, it is quite safe to make this expansion even for times long compared to the coherence time.

Substituting this expression for $r_n(t)$ into Eq. (2.28) gives the more accurate expression

$$\dot{f}_n \approx -\frac{1}{4n} \left[ 1 + \frac{n - \mu}{n^2} \right] f_n. \quad (2.31)$$

Since the initial condition is

$$f_n(0) = e^{-\mu} \frac{\mu^{n-1}}{(n-1)!}, \quad (2.32)$$

the only significant contribution to the sum (2.21) comes from $n$ such that $|n - \mu| \lesssim \sqrt{\mu}$. Also, as noted above, one can assume $t \lesssim n$. Then the correction term in Eq. (2.31) is of order $\mu^{-1/2}$ and can be ignored. One can thus return to the expression Eq. (2.28), which becomes (again ignoring corrections of order $\mu^{-1/2}$),

$$\dot{f}_n \approx -\frac{1}{4n} f_n. \quad (2.33)$$

The first order coherence function is thus

$$g^{(1)}(\tau) = \exp(-\tau/4\mu), \quad (2.34)$$

so that the coherence time is $4\mu$ (which is of order $\mu$ as promised). The Fourier transform of this expression is a Lorentzian with FWHM

$$\ell = \frac{1}{2\mu}. \quad (2.35)$$

This is the standard quantum limit $\ell_0$ of the linewidth for an ideal laser.
III. GAIN WITHOUT STIMULATED EMISSION

Since the “stimulated emission of radiation” is part of the acronym for laser, it might be thought that stimulated emission is essential to produce a laser. While a typical laser does rely upon stimulated emission to ensure that it runs single-mode, the fact that the model of Section 1A adds photons one by one suggests that it is not strictly necessary. I will now show that stimulated emission is indeed not necessary for laser action, and in fact that eliminating stimulated emission eliminates the phase diffusion caused by the gain process.

Stimulated emission is a simple consequence of the linear coupling of the laser field to its source, as in Eq. (2.1). That is to say, in Eq. (2.1), replace the $c$-number $\sqrt{\hat{n}}e^{i\phi}$. Whenever a rate is calculated in quantum theory it depends on the square of the Hamiltonian. Hence the fundamental gain rate from a linear coupling will vary as $n$, which is the so-called stimulated emission or Bose-enhancement factor. A fully quantum calculation of course gives spontaneous emission as well, and hence a gain rate proportional to $n+1$.

Since stimulated emission can be traced to the presence of $a$ in the coupling Hamiltonian, the only way to totally remove it is to substitute for $a$ a different lowering operator, one whose classical analogue does not increase with $n$. That is to say, in Eq. (2.1), replace

$$a = \sum_{n=1}^{\infty} \sqrt{n} |n-1\rangle \langle n|$$

(3.1)

by the Susskind-Glogower\,\cite{17} $e \equiv e^{i\phi}$ operator

$$e = (aa^\dagger)^{-1/2}a = \sum_{n=1}^{\infty} |n-1\rangle \langle n|.$$  \hspace{1cm} (3.2)

The new Hamiltonian would be extremely nonlinear if expressed as a power series in $a$ and $a^\dagger$, but it cannot be denied that it will not exhibit any stimulated emission.

Replacing $a$ by $e$ in the Hamiltonian (2.1) presents no problems in the rest of the derivation. Moreover, it is not even necessary to assume that $\epsilon = \Omega \tau$ is very small. Instead, the result is independent of $\epsilon$, due to the fact that $ee^\dagger = 1$. In particular, if one chooses $\epsilon = \pi/2$, the transformation effected on the field by one transit of the atom is semi-unitary:

$$\exp\left[\frac{\pi}{2}(e^\dagger \sigma - \sigma^\dagger e)\right]|u\rangle\langle \psi| = |l\rangle S|\psi\rangle.$$  \hspace{1cm} (3.3)

Here $|\psi\rangle$ is the state of the field and

$$S = e^\dagger = \sum_{n=0}^{\infty} |n+1\rangle \langle n|.$$  \hspace{1cm} (3.4)

The operator $S$ is semi-unitary rather than unitary because $S^\dagger S = 1$, but $SS^\dagger = 1 - |0\rangle \langle 0|$. Surprisingly, this transformation can be achieved physically using only the usual electric-dipole coupling \cite{18}. The trick is to use a three-level $\Lambda$ atom and another, classical field \cite{19}. Then, using a using a counter-intuitive pulse sequence, the atom is transferred from one lower state to the other, and one photon is created in the cavity field (with the energy lost from the classical field). Like the gain process in Sec. II, this adds precisely one photon to the field. The difference is that it does this without entangling the state of the field and the atom, and hence leaves the state of the field pure.

Taking the rate of addition of photons to the field to be $\Gamma$ as before, in place of Eq. (2.13) one obtains

$$\dot{\rho} = \mu D[e^\dagger]\rho + D[a]\rho.$$  \hspace{1cm} (3.5)

In the fock basis this becomes

$$\dot{\rho}_{m,n} = \mu (\rho_{n-1,m-1} - \rho_{m,n}) - (n + m)\rho_{n,m}/2 - \sqrt{(n+1)(m+1)}\rho_{n+1,m+1}.$$  \hspace{1cm} (3.6)

This yields exactly the same equation for the diagonal elements (the photon number populations). Hence the unstimulated master equation produces exactly the same photon number statistics as does the standard laser master equation (2.13).

To calculate the linewidth, proceed as before. One finds the following equation for $f_n$, defined as in Sec. II:

$$\dot{f}_n = \mu \left( \sqrt{\frac{n}{n-1}} f_{n-1} - f_n + nf_n - \frac{n-1}{2} f_n \right)$$  \hspace{1cm} (3.7)

$$= \left[ \sqrt{n(n-1)} r_{n-1} - \mu + \frac{\mu}{r_n} - \frac{n-1}{2} \right] f_n.$$  \hspace{1cm} (3.8)

Assuming $r_n \approx 1$ yields, as above, the self-consistent solution

$$f_n \approx -\frac{1}{8\mu} f_n.$$  \hspace{1cm} (3.9)

The first order coherence function is therefore

$$g^{(1)}(\tau) = \exp(-\tau/8\mu),$$  \hspace{1cm} (3.10)

so that the linewidth is

$$\ell = \frac{1}{4\mu}.$$  \hspace{1cm} (3.11)

This is half the standard quantum limit $\ell_0$ of Eq. (1.4). As explained in the introduction, the standard quantum limit for the phase diffusion rate contains equal contributions from the gain and loss processes. The gain process considered in this section does not introduce any phase noise; the operator $e^\dagger$ is more or less the exponentiation of the phase operator and so increases the photon number without affecting the phase distribution at all. Thus the phase diffusion in this model comes wholly from the loss process, and the rate is half the standard rate.
IV. FINITE ATOM-FIELD INTERACTION TIME

The preceding section showed that an interaction in which the atom is sure to give up its quantum of energy to the field from a single pass results in a linewidth a factor of two smaller than the standard limit. It was noted there that this could be achieved using an adiabatic passage, but this has yet to be done experimentally. This suggests that it would be worth exploring other ways to mimic the unstimulated gain process.

In this section I investigate one idea, based upon the gain mechanism of a micromaser \[20,21\]. This utilizes the same Jaynes-Cummings coupling \[2,4\] as in Sec. II. The difference is that the scaled interaction time \(\epsilon = \Omega\tau\) is not assumed to be small. This modifies the results of Sec. II as follows. The state of the field conditioned on the detection of an atom in the lower state is \[20\]

\[
\rho_l = \mathcal{J}_l \rho,
\]

where

\[
\mathcal{J}_l = \mathcal{J} \left[ e^{i \sin \left( \epsilon \sqrt{\mu} a \right)} \right].
\]

The field state conditioned on an atom passing through and remaining in the upper state is \[20\]

\[
\rho_u = \mathcal{J}_u \rho,
\]

where

\[
\mathcal{J}_u = \mathcal{J} \left[ \cos \left( \epsilon \sqrt{\mu} a \right) \right].
\]

For states having a photon distribution localized around \(\bar{n}\), if \(\epsilon\) is such that \(\epsilon \sqrt{\bar{n}} \approx \pi/2\), then it would seem that the action of the above superoperators could be approximated by

\[
\mathcal{J}_l \approx \mathcal{J} \left[ e^{i} \right],
\]

\[
\mathcal{J}_u \approx 0.
\]

That is, the atom would almost certainly come out in the lower state, having given up its quantum of energy to the field. This is the same situation as for the unstimulated gain as shown in Sec. III. This is why a finite interaction time \(\epsilon\) might be expected to lead to a linewidth below the standard limit.

If atoms are injected at a Poissonian rate \(\mu\) then the total master equation is the usual micromaser master equation

\[
\dot{\rho} = \left\{ \mu (\mathcal{J}_u + \mathcal{J}_l - 1) + \mathcal{D}[a] \right\} \rho.
\]

Here linear damping at rate unity has been included also.

This master equation has very complicated dynamics. For some values of \(\epsilon\) and \(\mu\) the stationary state does not have a well-defined intensity. That is, it is not the case that \(\sigma(n) \ll \bar{n}\). Hence the device is not necessarily a true laser in the sense of Ref. \[12\].

To ensure that the a well-defined photon number distribution is produced, the same technique as in Sec. II can be used. That is, if an atom is detected still in the upper state it is sent through again until it is detected in the lower state. The resulting master equation is

\[
\dot{\rho} = \left\{ \mu \mathcal{J}_l \sum_{k=0}^{\infty} \mathcal{J}_k^\dagger + \mathcal{D}[a] \right\} \rho
\]

\[
= \left\{ \mu \mathcal{J}_l (1 - \mathcal{J}_u)^{-1} + \mathcal{D}[a] \right\} \rho.
\]

In the photon number basis

\[
\dot{\rho}_{n,m} = \mu \frac{\sin(\epsilon \sqrt{n}) \sin(\epsilon \sqrt{m})}{\sqrt{n} - 1} \left[ 1 - \cos(\epsilon \sqrt{n} - 1) \cos(\epsilon \sqrt{m}) \right] \rho_{n-1,m-1}
\]

\[
- \mu \rho_{n,m} - (n + m) \rho_{n,m}/2 + \sqrt{(n + 1)(m + 1)} \rho_{n+1,m+1}.
\]

To find the linewidth, one proceeds as before to get the following equation for \(f_n\):

\[
\dot{f}_n = \mu \frac{\sqrt{n} \sin(\epsilon \sqrt{n} - 1) \sin(\epsilon \sqrt{m})}{\sqrt{n} - 1} \left[ 1 - \cos(\epsilon \sqrt{n} - 1) \cos(\epsilon \sqrt{m}) \right] f_{n-1}
\]

\[
- \mu f_n + n f_{n+1} - \frac{2n - 1}{2} f_n.
\]

Using the parameter

\[
\phi \equiv \epsilon \sqrt{\mu},
\]

one can continue the analysis as before and find eventually

\[
\dot{f}_n \approx - \frac{1}{8 \mu} \left[ 1 + \frac{\mu^2 \sin^2(\phi/\mu)}{\sin^2\phi} \right] f_n.
\]

That is, the linewidth of the laser is found to be

\[
\ell = \frac{1}{4 \mu} \left[ 1 + \left( \frac{\sin(\phi/\mu)}{\sin(\phi/\mu)} \right)^2 \right].
\]

It is easy to verify that this expression has a global minimum

\[
\ell = \lim_{\phi \to 0} \frac{1}{4 \mu} \left[ 1 + \left( \frac{\sin(\phi/\mu)}{\sin(\phi/\mu)} \right)^2 \right] = \frac{1}{2 \mu}.
\]

The limit \(\phi \to 0\) is the limit of short interaction times in which the original model of Sec. II is recovered, and the original linewidth \(\ell_0\) also. That is, no linewidth narrowing is possible using a finite interaction time in preference to an infinitesimal interaction time, despite the fact that the former can deposit a photon in the cavity in a single pass of the atom with very high probability.

This line-broadening is definitely not an artifact of the assumption that the atom is always put through again if it is detected still in its upper state; a similar result is obtained for the usual master micromaser equation with
a single pass per atom [23]. The approach to calculating the linewidth used in Ref. [22] was similar to the one used here. A more accurate estimation of the linewidth for the usual micromaser has to take into account the fact that the intensity is not always well defined [23]. This yields some deviations from the simple theory of Ref. [22], but still never shows any line-narrowing.

The reason that no linewidth narrowing occurs can be seen from the method of calculation I have employed. What turns out to be crucial is not to try to mimic the two terms in the ideal unstimulated gain, namely

\[ \mathcal{D}[a^\dagger] = \mathcal{J}[a^\dagger] - \mathcal{A}[a^\dagger] = \mathcal{J}[a^\dagger] - 1, \quad (4.16) \]

but rather to mimic the following ratio of matrix elements involving these two terms:

\[ \frac{\langle n+1 \mid \mathcal{J}[a^\dagger] \mid n \rangle}{\langle n \mid \mathcal{A}[a^\dagger] \mid n \rangle} = 1. \quad (4.17) \]

In the unstimulated case the ratio is unity, and the difference from unity in other cases is proportional to the contribution to the linewidth from the gain process. For the standard ideal laser,

\[ \frac{\langle n+1 \mid \mathcal{J}[a^\dagger] \mid n \rangle}{\langle n \mid \mathcal{A}[a^\dagger] \mid n \rangle} \approx 1 - \frac{1}{8n^2}. \quad (4.18) \]

Multiplying the deviation from unity by the gain constant \( \mu \) and replacing \( n \) by the mean photon number \( \mu \) gives \( 1/8\mu \). This is the standard contribution to the linewidth from the gain. For the micromaser,

\[ \frac{\langle n+1 \mid \mathcal{J}_\mu[n] \mid n \rangle}{\langle n \mid [1 - \mathcal{J}_\mu] \mid n \rangle} \approx 1 - \frac{\sin^2(\phi/\mu)}{8\sin^2 \phi}, \quad (4.19) \]

which again explains the result in Eq. (4.14).

V. NONLINEAR ATOM-FIELD INTERACTION

With now a better understanding of how to reduce the gain-induced phase diffusion, I turn to a second method for trying to mimic unstimulated gain. As noted in Sec. III, the operator \( e^1 \) would require an infinite series to be expressed in terms of powers of \( a \) and \( a^\dagger \). Any Hamiltonian containing infinite powers of the field is unphysical. However, nonlinear optical processes containing field powers greater than unity do occur. This suggests that it is worth considering the following approximation

\[ e^1 = a^\dagger (aa^\dagger/\mu)^{-1/2} = a^\dagger [\mu + (aa^\dagger - \mu)]^{-1/2} \approx a^\dagger \sqrt{\frac{3}{2}} \frac{1}{\mu} (\frac{1}{\mu} a^\dagger). \quad (5.1) \]

That is, I wish to consider a nonlinear Jaynes-Cummings Hamiltonian of the form

\[ H = i\Omega \left[ \sigma a^\dagger (3 - aa^\dagger/\mu) + (3 - aa^\dagger/\mu) a^\dagger \right], \quad (5.3) \]

which I expect to be useful when the photon number is approximately \( \mu \).

Physically, this Hamiltonian means that there are two processes which can excite the atom. The first is the usual linear dipole coupling to the field. The second is a three-photon process whereby a photon is virtually absorbed and re-emitted before finally being absorbed by the atom. The Hamiltonian matrix element for the second process is much smaller (for \( \mu \gg 1 \)), which is physically reasonable, and is of the opposite sign. It is very doubtful that such a Hamiltonian could be achieved simply using a two-level atom. However, it is possible that an effective Hamiltonian of this form could be achieved using a multilevel atom, and other fields. I will not further discuss the feasibility of producing this Hamiltonian, as my chief concern is with the question of principle: how well can the nonlinear Hamiltonian (5.3) reproduce the results of the unstimulated laser?

Assuming, as in previous sections, that the atoms are initially in the upper state and that any atom which exits the cavity still in the upper state is put through again, one derives, following the method of Sec. II, the following master equation for the cavity mode.

\[ \dot{\rho} = \mu [a^\dagger (3 - aa^\dagger/\mu)] A [a^\dagger (3 - aa^\dagger/\mu)]^{-1} \rho + \mathcal{D}[a] \rho. \quad (5.4) \]

This has the same Poissonian mixture of number states as in the standard laser, and is amenable to the same method of calculating the linewidth. The result is

\[ \ell = \frac{3}{8\mu}. \quad (5.5) \]

That is, the contribution from the gain is \( 1/8\mu \), which is half the standard result and half the contribution of \( 1/4\mu \) from the loss (which is of course unchanged). This result can again be understood from the ratio

\[ \frac{\langle n+1 \mid \{ \mathcal{J}[a^\dagger (3 - aa^\dagger/\mu)] \} \mid n \rangle}{\langle n \mid \{ A[a^\dagger (3 - aa^\dagger/\mu)] \} \mid n \rangle} \approx 1 - \frac{1}{16n^2}. \quad (5.6) \]

VI. DISCUSSION

The standard quantum limit to the laser linewidth is not the ultimate quantum limit, even for the Markovian case in which the gain medium is eliminated from the equations of motion of the laser mode. Hidden within the standard Markovian expression

\[ \ell_0 = \frac{\kappa}{2n}, \quad (6.1) \]
are equal contributions of $\kappa/4\bar{n}$ from the gain and loss mechanisms for the laser. The latter contribution is a fundamental limit because linear loss is necessary for a coherent output beam to form. However the former results from a particular (extremely reasonable) assumption about the gain mechanism for laser action, namely that it comes from a weak linear coupling between the field and the gain medium.

These arguments suggest that a different sort of gain mechanism could produce a laser with a linewidth up to 50% below the standard quantum limit. As I have shown above, this ultimate Markovian limit

\[ \ell_{\text{ult}} = \frac{\kappa}{4\bar{n}} \]  

(6.2)

can be achieved with a gain mechanism in which stimulated emission into the laser mode is eliminated. This requires that the matrix element for the addition of a photon to the laser mode be independent of the number of photons in the mode. As discussed, this could be physically achieved with adiabatic transfer of photons from another field using a counter-intuitive pulse sequence.

I also examined two other gain mechanisms with similarities to the non-stimulated gain, to see if they also produced linewidth narrowing. The first, using the usual Jaynes-Cummings Hamiltonian but with a finite interaction time (as in the micromaser) did not. The second, using a nonlinear Jaynes-Cummings Hamiltonian involving three-photon as well as one-photon processes, produced a linewidth of

\[ \ell = \frac{3\kappa}{8\bar{n}} \]  

(6.3)

That is, the phase diffusion due to the gain was reduced by 50% from the standard limit, resulting in an overall reduction of 25% in the linewidth. Presumably higher-order nonlinear optical processes could more closely approach the ultimate limit. However, the difficulty in producing such nonlinear optical processes, and the fact that even a third-order nonlinearity fails to reach the ultimate limit, suggests that the adiabatic transfer is a better experimental option for probing the ultimate quantum limit to the laser linewidth.

The ultimate limit for the rate of phase diffusion attained by eliminating gain noise can also be obtained, for short times, by instead eliminating loss noise. This can be achieved by coupling the laser output into a squeezed vacuum rather than a normal vacuum. This only works for short times because it requires a specific phase relation between the squeezed vacuum and the coherent field in the laser, which will not remain valid since the laser phase continues to diffuse. It was suggested in Ref. [24] that it might be possible to produce the squeezed vacuum by driving the squeezing device with the laser itself. In this case the whole squeezing device should really be considered as part (an internal absorber, in fact) of the laser, so that $\bar{n}$ in the original laser cavity should no longer be used as a good measure of the total stored excitation. Similar comments could be made about the proposal of Ghosh and Agarwal [29], who also misquote the expression for the standard quantum limit given in Ref. [1] by a factor of two as their Eq. (18). I believe that a rigorous analysis of these proposals would reveal no reduction below the standard quantum limit.

\section{VII. Conclusion}

In the introduction I reproduced a simple argument purporting to use the time-energy uncertainty principle to derive the standard laser linewidth as a consequence of phase diffusion due to the gain process. The results of this paper show that any such simple argument is untenable since the gain process contributes only half of the standard phase diffusion rate. To compensate for disposing of this simple argument, I will conclude this paper with a (not quite so simple) argument deriving the ultimate Markovian quantum limit $\ell_{\text{ult}}$ from another uncertainty principle argument.

Instead of the energy-time relation, which is of doubtful content, I will use the quadrature uncertainty relation

\[ V(X)V(Y) \geq 1, \]  

(7.1)

where $X/2$ and $Y/2$ are the real and imaginary components of the laser mode amplitude $a$. Clearly the vacuum state is rotationally symmetric with

\[ V(X) = V(Y) = 1, \]  

(7.2)

and this holds also for a coherent state (which is the state the laser mode can be assumed to be in).

Let the mean amplitude of the coherent state be real and positive so that $X/2 = \sqrt{\bar{n}}$ and $Y = 0$. The phase variance is

\[ V(\phi) = V \left( \arctan \frac{Y}{X} \right) \geq V(Y/X^2) = \frac{1}{4\bar{n}} \]  

(7.3)

for $\bar{n} \gg 1$. Now the effect of linear damping for an infinitesimal time $dt$ is to reduce the mean photon number of the coherent state from $\bar{n}$ to $\bar{n}(1 - \kappa dt)$. Thus the change in the phase variance is

\[ dV(\phi) = \frac{\kappa dt}{4\bar{n}}. \]  

(7.4)

A noiseless gain process will return the mean photon number to $\bar{n}$ without increasing the phase noise. Therefore the phase variance increases at least as

\[ V(\phi) \sim \frac{\kappa t}{4\bar{n}}. \]  

(7.5)

The linewidth is defined from the two-time correlation function

\[ \langle a^{\dagger}(t)a(0) \rangle \sim \bar{n}e^{-V(\phi)/2} \sim \bar{n}e^{-\kappa t/8\bar{n}}. \]  

(7.6)
The Fourier transform of this expression is a Lorentzian with a FWHM of
\[ \ell = \frac{\kappa}{4n}, \] (7.7)
which is the ultimate quantum limit to the laser linewidth, as claimed.

ACKNOWLEDGMENTS

I would like to thank D. Pope for a critical reading of this paper. The work was undertaken with the support of the Australian Research Council.

APPENDIX: REFINING THE SCHAWLOW-TOWNES LIMIT.

The Schawlow-Townes expression
\[ \ell_{ST} = \frac{\hbar \omega}{P_{out}} \gamma^2. \] (A1)
was derived in the days before good optical cavities, and hence implicitly assumes that the atomic linewidth \( \gamma \) is much smaller than the (FWHM) cavity linewidth \( \kappa \). With \( \kappa \lesssim \gamma \), it is necessary to replace \( \gamma \) by the bare linewidth of the laser \( \ell_{bare} \). This is the frequency spread the output would have if the pump were suddenly turned off and all of the energy allowed to escape. For a large class of lineshapes, it can be shown that a reasonable approximation to the bare linewidth including contributions from the atomic (or other gain) medium and the cavity is
\[ \ell_{bare}^{-1} = \gamma^{-1} + \kappa^{-1}. \] (A2)

For instance, this expression agrees with that given by Haken (p. 103 of Ref. citeHak84) for the case where \( \kappa \gtrsim \gamma \). In the other cases, where \( \kappa \ll \gamma \), the \( \gamma \) in the Schawlow-Townes expression is simply replaced by \( \kappa \) which also agrees with Eq. (A2). The corrected Schawlow-Townes expression is thus
\[ \ell_{ST} = \frac{\hbar \omega}{P_{out}} \ell_{bare} = \frac{\hbar \omega}{P_{out}} \gamma^2 \kappa^2. \] (A3)

The second correction which must be made to the Schawlow-Townes linewidth relates to its use of the output power. Say, for argument’s sake, that one one has a laser with linewidth given by the Schawlow-Townes limit, with all of the power coming out of one mirror. Then say that the mirror is replaced by one of the same reflectance, but with larger internal absorption. Then the power loss per round trip is identical, so the laser dynamics remain the same and the linewidth would remain the same. But the power out would be reduced because the transmittance is reduced. Therefore the Schawlow-Townes formula would now predict an increased linewidth, which does not occur. In other words, the actual new linewidth would be less than the quantum limit set by the Schawlow-Townes formula. It is obviously inappropriate that a quantum limit can be surpassed by building a worse device.

The resolution to this problem with the Schawlow-Townes linewidth is to eliminate \( P_{out} \) from the expression by recognizing that
\[ \frac{P_{out}}{\ell_{bare}} \] (A4)
is an upper bound on the mean energy \( \bar{E} \) stored as coherent excitations in the laser system. If all of the stored coherent excitation eventually makes it into the output beam of the laser then the bare linewidth \( \ell_{bare} \) is due wholly to the output coupling and \( P_{out} = \ell_{bare} \bar{E} \). In general \( P_{out} \) is less than this. Reducing the output coupling efficiency (as discussed in the preceding paragraph) will not affect \( \bar{E} \) so it is the correct parameter to use, rather than \( P_{out} \). The doubly corrected Schawlow-Townes limit is thus
\[ \ell_{ST}'' = \frac{\ell_{bare} \hbar \omega}{\bar{E}} \leq \frac{\hbar \omega}{P_{out}} \frac{\gamma^2 \kappa^2}{(\gamma + \kappa)^2}, \] (A5)
where the inequality becomes an equality only for perfectly efficient output coupling.

It is convenient to define the number of quanta of coherent excitation, \( \bar{N} = \bar{E} / \hbar \omega \). For the case \( \kappa \gg \gamma \) the excitation stored in the gain medium is negligible and \( \bar{N} = \bar{n} \), where the latter represents the mean photon number in the cavity. If the gain medium cannot be adiabatically eliminated then \( \bar{N} \) must include the excitations stored coherently in the gain medium as well. If \( \gamma \ll \kappa \), as in the original Schawlow-Townes expression, these excitations in the gain medium will be the dominant ones.

The final correction which needs to be made to the Schawlow-Townes linewidth is to insert a factor of \( \frac{1}{2} \). The Schawlow-Townes limit without this factor is appropriate to a laser below threshold in which the complex amplitude of the field undergoes large slow fluctuations (for \( N \gg 1 \) which is the limit in which the Schawlow-Townes equation is valid). Above threshold, the laser intensity fluctuations are almost eliminated, leaving only phase fluctuations. This increases the coherence time by a factor of two, so that the final corrected expression for the laser linewidth is
\[ \ell_{ST} = \frac{\ell_{bare}}{2\bar{N}} \leq \frac{\hbar \omega}{2P_{out}} \frac{\gamma^2 \kappa^2}{(\gamma + \kappa)^2}. \] (A6)
Here st stands for standard (quantum limit) as opposed to ST which stands for Schawlow-Townes.

In the limit \( \kappa \ll \gamma \), which is the usual limit for most lasers, and which allows the gain medium to be adiabatically eliminated from the field equations, one obtains
\[ \ell_0 = \frac{\kappa}{2\bar{n}} \leq \frac{\hbar \omega}{2P_{\text{out}}} \kappa^2. \]  

(A7)

This result has often (including by myself [2]) been quoted as the Schawlow-Townes limit, despite the obvious differences from Eq. (A1). Here I will call it instead the standard Markovian quantum limit to the laser linewidth. “Markovian” refers to the fact that the equations of motion for the laser mode, including gain and loss, are well-approximated by Markovian equations. For the gain process this is a consequence of adiabatically eliminating the gain medium. For the loss process, it is a consequence of assuming a high-\(Q\) cavity. Corrections (upwards) for non-Markovian loss (low-\(Q\) cavities) are discussed for example in Ref. [27], but here I will always assume a high-\(Q\) cavity.

Obviously for \(\gamma \lesssim \kappa\), the linewidth of Eq. (A6) will be less than the standard Markovian quantum limit of Eq. (A7). That is a reflection of the fact that in this case the bare linewidth \(\ell_{\text{bare}}\) is less than \(\kappa\), and also that the gain medium is an extra reservoir of energy (coherent with the laser mode) so that \(\bar{N}\) is greater than \(\bar{n}\). A linewidth which, in the ideal limit, reduces to \(\ell_{\text{st}}\) was recently derived in Ref. [23], for a laser with \(\gamma \lesssim \kappa\). These authors claimed that this was “reduced compared to the Schawlow-Townes limit” because they followed the common (but, in my opinion, erroneous) practice of identifying \(\ell_0\) as the Schawlow-Townes limit. To me this seems to be an example of imprecise terminology obscuring an otherwise valuable contribution to fundamental laser physics. In this paper I always work with models in which the gain medium can be adiabatically eliminated, so that \(\ell_0 = \ell_{\text{st}}\).

References:

[1] A.L. Schawlow and C.H. Townes, Phys. Rev. 112, 1940 (1958).
[2] H. Haken, Laser Theory, (Springer, Berlin, 1984).
[3] W.H. Louisell, Quantum Statistical Properties of Radiation (John Wiley & Sons, New York, 1973).
[4] M. Sargent, M.O. Scully, and W.E. Lamb, Laser Physics (Addison-Wesley, Reading Mass., 1974).
[5] R. Loudon, The Quantum Theory of Light (Oxford University Press, Oxford, 1983).
[6] H. Weichel, Am. J. Phys. 44, 839 (1976).
[7] R.J. Glauber, Phys. Rev. 130, 2529 (1963); ibid. 131, 2766 (1963).
[8] E.C.G. Sudarshan, Phys. Rev. Lett. 10, 277 (1963).
[9] D.F. Walls and G.J. Milburn, Quantum Optics (Springer, Berlin, 1994).
[10] S.M. Barnett, S. Stenholm, and D.T. Pegg, Opt. Comm. 73, 314 (1989).
[11] M. Hillery, M. Freyburger, and W. Schleich, Phys. Rev. A 51, 1792 (1995).
[12] H.M. Wiseman, Phys. Rev. A 56, 2068 (1997).
[13] H.M. Wiseman, Phys. Rev. A 47, 5180 (1993).
[14] For the use of this superoperator inverse in a different context, see S.L. Braunstein and C.M. Caves, Phys. Rev. Lett. 72, 343 (1977).
[15] G. Lindblad, “On the Generators of Quantum Dynamical Semigroups” Commun. math. Phys. 48, 199 (1976).
[16] C.W. Gardiner, Quantum Noise (Springer-Verlag, Berlin, 1991).
[17] L. Susskind and J. Glogower, “Quantum Mechanical Phase and Time Operator” J. Physics A 51, 49 (1964).
[18] A.S. Parkins et al., Phys. Rev. A 51, 1578 (1995).
[19] It might be thought the presence of an external classical field would disqualify this device from the class of lasers, as defined in Ref. [23]. However, it is in fact not necessary for the second field to have a well-defined coherent amplitude, or even a very large photon number, but only for it to have a negligible probability of being in the vacuum state.
[20] P. Filipowicz, J. Javanainen, and P. Meystre, Phys. Rev. A 34, 3077 (1986).
[21] G. Raithel et al., in Cavity Quantum Electrodynamics, edited by Paul R. Berman, Adv. At. Mol. Opt. Phys. Suppl. 2 (1994).
[22] M.O. Scully et al., Phys. Rev. A 44, 5992 (1991).
[23] R.R. McGowen and W.C. Schieve, Phys. Rev. A 55, 3818 (1997) and references therein.
[24] J. Gea-Banacloche Phys. Rev. Lett. 59, 543 (1987).
[25] M.A.M. Marte and D.F. Walls, Phys. Rev. A 37, 1235 (1988).
[26] R. Ghosh and G.S. Agarwal Phys. Rev. A 39, 1582 (1989).
[27] S. Prasad and B.S. Abbott, Phys. Rev. A 38, 3551 (1988).
[28] M.P. van Exter, S.J.M. Kuppens, and J.P. Woerdman, Phys. Rev. A 51, 809 (1995).
FIG. 1. Schematic of an ideal laser. An atom in the upper state passes through the cavity and its state is then detected. If the atom remains in the upper state, the process is repeated until it is detected in the lower state. The time for this process (including repetitions) is assumed to be very short compared to the cavity damping time $\kappa^{-1}$. Once the atom is detected in the lower state, a new upper-state atom is injected after a random waiting time $\tau$ having an exponential distribution $w(\tau) = \exp(-\kappa \mu \tau)$. Here $\mu$ is the desired mean number of photons in the cavity.