

Topics in optics and quantum electronics

Jinyuan Wu

April 15, 2024

1 Mode description of electromagnetic fields

In this section we mainly focus on the mode description of classical electrodynamics. Modes are seen in resonators, waveguides, photonic crystals, and more. Loss can then be included perturbatively and we get leaky modes.

In the case of electrodynamics in vacuum, in the Coulomb gauge (i.e. $\varphi = 0$), we have

$$\nabla \times \nabla \times \mathbf{A} = \left(\frac{\omega}{c}\right)^2 \mathbf{A}, \quad (1)$$

and under the inner product definition

$$\langle \mathbf{A} | \mathbf{B} \rangle := \int d\mathbf{r} \mathbf{A}^* \cdot \mathbf{B}, \quad (2)$$

the LHS of the equation on \mathbf{A} is Hermitian, guaranteeing that ω is real.

From the well-known relation $[a, a^\dagger] = 1$, it follows that the classical counterparts of creation and annihilation operators are complex variables which satisfies the relation

$$\{a, a^*\} = -i. \quad (3)$$

This means if the Poisson brackets are to be defined in terms of a and a^* , it should be defined as

$$\{A, B\} = \frac{1}{i} \left(\frac{\partial A}{\partial a} \frac{\partial B}{\partial a^*} - \frac{\partial A}{\partial a^*} \frac{\partial B}{\partial a} \right), \quad (4)$$

where a and a^* are to be regarded as independent variables. The general relation between a, a^* and x, p is

$$a = \alpha q + i\beta p, \quad q = \frac{a + a^*}{2\alpha}, \quad p = \frac{a - a^*}{2i\beta}, \quad 2\alpha\beta = 1. \quad (5)$$

If q and p follow the canonical commutation relation, then a and a^* follow the commutation relation between a creation operator and an annihilation operator, and vice versa.

We can also justify (3) by Hamiltonian's equations. From (5) and the chain rule, we find

$$\begin{aligned} \dot{a} &= \frac{\partial a}{\partial q} \dot{q} + \frac{\partial a}{\partial p} \dot{p} = \frac{\partial a}{\partial q} \frac{\partial H}{\partial p} - \frac{\partial a}{\partial p} \frac{\partial H}{\partial q} \\ &= \frac{\partial a}{\partial q} \left(\frac{\partial H}{\partial a} \frac{\partial a}{\partial p} + \frac{\partial H}{\partial a^*} \frac{\partial a^*}{\partial p} \right) - \frac{\partial a}{\partial p} \left(\frac{\partial H}{\partial a} \frac{\partial a}{\partial q} + \frac{\partial H}{\partial a^*} \frac{\partial a^*}{\partial q} \right) \\ &= -2i\alpha\beta \frac{\partial H}{\partial a^*}, \end{aligned} \quad (6)$$

and similarly

$$\dot{a}^* = 2i\alpha\beta \frac{\partial H}{\partial a}. \quad (7)$$

Now that means we have to define the Poisson bracket as (4) to keep the form

$$\frac{dA}{dt} = \{A, H\}. \quad (8)$$

The $2\alpha\beta = 1$ constraint can be explained by conservation of area in phase space: roughly speaking the area inside the trajectory of $a(t)$ or $(q(t), p(t))$ is the “strength of oscillation”, like the number of photons.

Now going back to the problem of electromagnetic fields. Consider the one dimensional periodic boundary condition problem. We have (in phaser form)

$$\tilde{E}_\pm(x, t) = \tilde{E}_\pm^0 e^{ikx - i\omega t}, \quad (9)$$

and we consider A to be the coordinate and hence $-E = \partial A / \partial t$ is the momentum. To decompose A into modes, we write A as static, spatially varying factors weighted by time-dependent real values, and thus $A = q(t)A_0 e^{ikz}$,

2 Perturbation theory

Stationary electrodynamics in terms of \mathbf{E} is not a Hermitian problem, but a generalized Hermitian one: the problem is

$$\nabla \times \nabla \times \mathbf{E} = \left(\frac{\omega}{c}\right)^2 \epsilon_r(\mathbf{r}) \mathbf{E}(\mathbf{r}), \quad (10)$$

and although the operators on both LHS and RHS are Hermitian, the LHS multiplied with the inverse of the RHS is not. The time-independent perturbation theory for this problem is therefore slightly more complicated, and the “energy” being solved is

$$\alpha := \frac{\omega^2}{c^2}, \quad (11)$$

instead of the frequency.

$$\Delta\omega_n^{(1)} = -\frac{\omega_n^{(0)}}{2} \langle \tilde{\mathbf{E}}_n | \epsilon_0 \Delta\epsilon_r | \tilde{\mathbf{E}}_n \rangle. \quad (12)$$

The minus sign can be understood using the following line of argumentation: if a piece of dielectric is inserted into an isolated system, the total energy is conserved but some energy is stored within the internal degree of freedom dielectric, so $\omega(a^\dagger a + 1/2)$, the vacuum part of the Hamiltonian has to decrease. One important caveat is that the leading order of energy correction *can't* be evaluated by $\langle \mathbf{E}_n | \Delta\epsilon_r | \mathbf{E}_n \rangle$: since the mapping from the electric field to the magnetic field involves at least one time derivative, the electromagnetic Hamiltonian in the form of $\mathbf{E}^2 + \mathbf{B}^2$ can't be seen as an ordinary Hamiltonian whose corresponding wave functions are electric fields, and therefore ordinary perturbation techniques fail. One way to work around this is doing everything in terms of modes. This is done in homework 2.

3 Loss as perturbation

Now we study the influence of ϵ_2 , the imaginary part of ϵ_r , as a perturbation. We can just inject $\Delta\epsilon_r = i\epsilon_2$ into (12), and we get

$$\Delta\omega_n^{(1)} = -i\frac{1}{\tau_n}, \quad \frac{1}{\tau_n} = \frac{\omega_n}{2} \langle \tilde{\mathbf{E}}_n | \epsilon_0 \epsilon_2 | \tilde{\mathbf{E}}_n \rangle. \quad (13)$$

This means the time evolution of the corresponding mode a_n contains a e^{-t/τ_n} factor, and the EOM seems to be

$$\dot{a} = -i\omega a - \frac{1}{\tau} a. \quad (14)$$

We can add a driving field to the RHS of this equation of motion; without considering any nonlinear effects,¹ the lifetime remains the same regardless of the possible external driving. The question is whether (14) is truly the correct equation for a damped system without external driving. It is of course possible to have a different kind of damping: if we choose to not start from the imaginary part of ϵ_r and instead insert a $-kv$ type of damping into the EOM of p and q , then the damping term for a should assume the form of

$$\dot{a} + i\omega_0 a = -\frac{a}{\tau} + \frac{a^*}{\tau}. \quad (15)$$

Interestingly, from an equally intuitive starting point, now we get a set of coupled EOMs for a and a^* , while in (14), the EOMs for a, a^* are not coupled.

The “real” form of the damping term depends on how the system couples to the bath; if we are studying a harmonic oscillator then (15) is exact, while (14) is not; the most natural way to include loss in Maxwell's equations also leads to something similar to (15). So indeed, although (14) seems quite natural, it's not exact, and the question is how it's related to the exact equations.

¹Damping always comes from coupling of the system with a much larger, external bath. The first order correction from the bath adds a finite imaginary part to the self energy of the system; the interaction with the bath of course also modifies the coupling between the bath and the system, but this is not first-ordered, and can be ignored when damping is small.

The answer is rotating wave approximation. From the usual procedure we will find the coupled terms in the loss are related to fast oscillation terms, which can be ignored for the low frequency part of a, a^* . This is valid as long as there is separation between the time scales of damping and of oscillation. This is the condition of the validity of RWA of an external driving field, which, in the latter case, means the frequency of driving should be close to ω_0 while the driving *amplitude* should be small. Since the damping term is not oscillating, the validity condition of RWA for damping is simpler.

TODO: does rotating wave approximation change the way the system couple to external fields? (c.f. general relativity and the $f(R)$ theory)

4 Mode coupling

Consider a ring resonator. We have a right mover mode and a left mover mode, which we refer to as

$$\tilde{E}_k = a_k E_k^0 e^{ikz} \quad (16)$$

and

$$\tilde{E}_{-k} = a_{-k} E_{-k}^0 e^{-ikz}, \quad (17)$$

respectively.

Now suppose we have such a field configuration:

$$E = a_k \tilde{E}_k + \text{c.c.} + a_{-k} \tilde{E}_{-k}, \quad (18)$$

and there is a $\Delta\epsilon_r$ perturbation in the system. The change of the frequency is therefore given by

$$\Delta\omega^{(1)} = -\frac{\omega^{(0)}}{2} \langle \Delta\epsilon_r \rangle \propto (a_k^* + a_{-k})(a_k^* + a_{-k}) + \text{c.c.}, \quad (19)$$

and therefore the contribution of $\Delta\epsilon_r$ to the energy of the system looks like

$$\Delta E = -\tilde{\gamma}(a_k^* + a_{-k})(a_k^* + a_{-k}) + \text{c.c.} \quad (20)$$

The EOM of a_k therefore is (TODO: formalism)

$$\dot{a}_k = -i\omega_k a_k + 2i\gamma(a_k^* + a_{-k}). \quad (21)$$

Again we can apply RWA and remove the “fast” a_k^* term, so eventually the EOMs of the two modes are (TODO: RWA without external driving)

$$\dot{a}_k = -i\omega_0 a_k + i\tilde{g}a_{-k}, \quad \dot{a}_{-k} = -i\omega_0 a_{-k} + i\tilde{g}^* a_k, \quad \tilde{g} = 2\gamma. \quad (22)$$

The EOMs are still within the Hamiltonian dynamics framework, so we can diagonalize them and find that when

$$\Delta\epsilon_r = \Delta\epsilon \cos(2kz), \quad (23)$$

the new mode $a_k + a_{-k}$ has a lower frequency $\omega_0 - |\tilde{g}|$ while the new mode $a_k - a_{-k}$ has a higher frequency $\omega_0 + |\tilde{g}|$. But the above formalism is not restricted to this perturbation: whatever $\Delta\epsilon_r$ is, the way it affects the two modes we consider here is by influencing \tilde{g} . Specifically, $|\tilde{g}|$ can be used to estimate the roughness of a ring resonator.

Of course, we can also perturb the system by modulation in *time*. Consider a classic LC -circuit, for example: it can be verified that an oscillating C results in

$$H = \omega_0 a^* a + \frac{\gamma(t)}{2} (a + a^*)^2, \quad (24)$$

because there is no L^2 term that cancels the aa^* terms. The resulting EOMs are

$$\dot{a} = -i\omega_0 a - i\gamma(t)(a + a^*), \quad (25)$$

and the perturbation $\gamma(t)$ usually takes the form like

$$\gamma(t) = \frac{\tilde{\gamma}_0}{2} e^{-i\Omega t} + \text{c.c.} \quad (26)$$

where $\Omega = 2\omega_0$. After RWA the only term that survives is

$$\dot{a} = -i\omega_0 a - i\frac{\tilde{\gamma}_0}{2} e^{-i\Omega t} a^* \quad (27)$$

and hence we find

$$\frac{d}{dt} \begin{bmatrix} a e^{-i\omega_0 t} \\ a^* e^{i\omega t} \end{bmatrix} = \begin{bmatrix} 0 & -i\frac{\tilde{\gamma}_0}{2} \\ i\frac{\tilde{\gamma}_0^*}{2} & 0 \end{bmatrix} \begin{bmatrix} a e^{-i\omega_0 t} \\ a^* e^{i\omega t} \end{bmatrix} \quad (28)$$

We can easily see that the “source” of the $e^{-i\Omega t}$ term is a^* , which in turn comes from the $e^{i\Omega t}$ term and a . This eventually gives us a feedback (usually *positive* but not always – see below) to a , leading to the increase of the total energy. The derivative of the “free” energy is

$$\dot{H}_{\text{SHO}} = \text{Re}(-i|\tilde{\gamma}_0| e^{i\phi} a_0), \quad (29)$$

where ϕ is the phase of $\tilde{\gamma}_0$, and a_0 is the slow oscillating part of a . This expression tells us two things. First, the phase of the pumping matters: it can be chosen in a specific way to *reduce* the total energy. Second, the growth or damping of the total energy is exponential, and therefore if at first you don’t have a finite amplitude, there is no growth at all. To start the growth, noise is important; classically this comes from thermal fluctuation; in the quantum case, the above EOM involves some terms from non-trivial commutation relations and we may say the growth is started by the zero-point energy.

5 Nonlinearity

We know the free $\epsilon_0 \epsilon_r \mathbf{E}^2/2 + \mathbf{B}^2/2\mu_0$ Hamiltonian can be written as $\sum \omega a^* a$; similarly a nonlinear Hamiltonian can be written as $\sum a a^* + \text{c.c.}$; in principle all modes can be mixed together by nonlinearity; usually however only transitions satisfy the phase matching condition are strong enough; the definition of “phase matching” is somehow unclear: for plain waves, it means $\sum \mathbf{k} = 0$ and has to be satisfied *strictly*; but in other cases, like in a photonic crystal, if \mathbf{k} means Bloch wave vectors, then the TODO: again, if a transition channel doesn’t satisfy the phase matching condition, does it mean this channel vanish in reciprocal space?

Phase matching in the time domain has to be exactly satisfied only when we are dealing with a scattering problem, in which we integrate from $t = -\infty$ to $t = \infty$ and get a $\sum \omega = 0$ condition; for a finite time problem, this condition doesn’t need to be exactly satisfied, although a transition channel satisfying the $\sum \omega = 0$ condition still likely to be the most influential one. In Feynman diagrams we do have strict frequency conservation conditions at each vertex but then we don’t have the on-shell condition for internal lines; by the correspondence between Feynman diagrams and Goldstone diagrams, we can let the internal lines to satisfy the on-shell conditions but let the total energy of the system to be off-shell, therefore arriving at results obtained from the EOM approach.

The EOM of one mode looks like (here NL means “nonlinear”)

$$\dot{a}_1 = -i\omega_0 a_1 - i\frac{\partial H_{\text{NL}}}{\partial a_1^*}. \quad (30)$$

For SHG, for example, we have

$$\begin{aligned} \dot{a}_1 &= -i\omega_1 a_1 + i\tilde{\gamma} a_1^* a_2, \\ \dot{a}_2 &= -i\omega_2 a_2 + i\tilde{\gamma}^* a_1 a_1. \end{aligned} \quad (31)$$

We can also study a nonlinear system using nonlinear Maxwell’s equations. This can be seen as a cross-verification of the correctness of the mode-based, usually somehow simplified formalism; when the dielectric function has a strong “dispersion” – i.e. frequency dependence – the Hamiltonian approach is also hard to use.

6 Temporal module coupling

Now we consider the coupling of modes with the external world. The formalism is almost identical to the so-called input-output formalism in quantum optics. As a motivation, consider

the following model: imagine we have a one-dimensional photonic crystal, with one defect (for example a missing cavity in a string of cavities). Now if a pulse is injected from one end of the photonic crystal we expect to see reflectivity and radiation to the 3D space around the defect. Whatever the defect is, it can be modeled as a mode (or several modes) coupled with “ports” around it. In the absence of the ports, the EOM of the system always takes the form $\dot{a} = -i\omega_0 a$, or possibly with a decay. Now we consider the structure of a port. We represent the waves going “into” the defect as S_i and the waves going “out” of the defect as S_o . Both variables may be linear combination of modes in the ports, or even contain some nonlinear effects. The EOM of the defect then is (TODO: Peter says it only holds under RWA)

$$\dot{a} = -i\omega_0 a - \frac{a}{\tau} + K S_i. \quad (32)$$

Here we do the Markovian approximation and assume that the a/τ term faithfully captures the coupling of the system with the output port. Note that this assumption is completely *orthogonal* to the assumption that the bath is thermalized, and therefore the equation above is *not* a Langevin equation, and still we can further loosen this approximation and include frequency dependence in the energy loss term. Indeed, although from the perspective of the mode a , the interaction with modes in the port causes an exponential decay, from the perspective of modes in the ports there is no real dissipation and coupling with the disorder can be captured as a scattering process – see the end of this section. Of course it’s possible to have “dissipative loss”, which means we have modes that are not considered as a part of the port but are nevertheless coupled to a and damp it;² here we ignore these dissipation channels and assume that dissipation mainly comes from the output modes of the port. Now if we somehow manage to integrate out a and consider the theory of the port. Ignoring all nonlinear effects and retardation effects, the relation between S_o and S_i, a can only take the following form:

$$S_o = \beta S_i + \gamma a, \quad (33)$$

where β and γ have to be related to τ .

When S_i, S_o are appropriately normalized we expect to see that $|S_i|^2$ is the power incident while $|S_o|^2$ is the output power. For example, suppose we have a ring resonator, and we place a defect in it. Since a ring resonator is usually thin, it can be modeled as a pipe with periodic boundary condition, and the power going through an arbitrary point in the ring resonator takes the form of

$$P = v_g \frac{\omega_0 a^* a}{L}, \quad (34)$$

where a is a mode in the ring resonator (not the mode around the defect), and therefore S should be defined such that

$$P = S S^* = v_g \frac{\omega_0 a^* a}{L}. \quad (35)$$

Moreover we normalize a so that $|a|^2$ is the mode energy.

Now we seriously consider how to derive the relation between τ and γ, K . The most famous example for coupling with outgoing modes leading to dissipation is arguably the fact that an infinite network of LC circuits is equivalent to a resistor. Similar systems include a mass point connected to an infinite spring or a leaky cavity. When there is no input to the defect mode, the loss of energy in the mode equals the outgoing flux S_o , and therefore

$$\frac{dU_{\text{defect}}}{dt} = -\frac{2}{\tau} U_{\text{defect}} = -|S_o|^2, \quad (36)$$

and therefore under the normalization convention

$$U_{\text{defect}} = |a|^2, \quad (37)$$

we have

$$S_o = \underbrace{\sqrt{\frac{2}{\tau}} e^{i\alpha}}_{\gamma} a. \quad (38)$$

²Sometimes this kind of loss is known as **internal loss**; note that here the term *internal* means “not a part of the system we are interested in”.

The phase factor is a variable that depends on the exact definition of the output of the system; consider for example a defect embedded into a wave guide: we need to choose a point at which we decide “we are no longer near the defect”, and the definition – hence the phase – of the outgoing mode depends on this point. TODO: what happens when the input is turned on? We will get something like $2 \operatorname{Re} K S_i^* a = |S_i|^2$. Whether this is always true can be verified using a harmonic oscillator; essentially that’s about what is the definition of “outgoing”. Using MBPT this can be easily justified: when everything is linear, couplings between the S_i and S_o modes do not influence each other.

We still have K and β . Since a, S_i, S_o are assumed to form a closed system, we expect they form a system with time reversal symmetry. After time reversal operation, hence the EOM of a becomes

$$-\frac{d}{dt}a^*(-t) = -i\omega_0 a^*(-t) - \frac{a^*(-t)}{\tau} + \mathcal{T} K S_i, \quad (39)$$

where we don’t immediately write what’s the time reverse of S_i : time reversal operation also changes the mode label, and the correct transform is $\mathcal{T} S_i = S_o^*(-t)$ (in principle we can put an additional phase factor between $\mathcal{T} S_i$ and S_o^* ; for convenience this is not done here), so eventually

$$-\frac{d}{dt}a^*(-t) = -i\omega_0 a^*(-t) - \frac{a^*(-t)}{\tau} + K S_o^*(-t), \quad (40)$$

and therefore

$$\frac{d}{dt}a(-t) = -i\omega_0 a(-t) + \frac{a(-t)}{\tau} - K^* S_o(-t). \quad (41)$$

Similarly

$$S_i(-t) = \beta^* S_o(-t) + \gamma^* a(-t). \quad (42)$$

Note that \mathcal{T} doesn’t change the coefficients but does turn a to a^* because the latter are complex fields. Now the EOMs should be equivalent to the EOMs before time reversal operation, and when the input modes are turned off, from

$$S_o(-t) = -\frac{\gamma^*}{\beta^*} a(-t) = -\frac{e^{-i\alpha}}{\beta^*} \sqrt{\frac{2}{\tau}} a(-t)$$

we find

$$\beta = -e^{2i\alpha}. \quad (43)$$

Similarly by comparing the EOMs of a , when the input is turned off, we have

$$\frac{d}{dt}a(-t) = -i\omega_0 a(-t) + \frac{a(-t)}{\tau} - K^* e^{i\alpha} \sqrt{\frac{2}{\tau}} a(-t),$$

and since we also expect to see

$$\frac{d}{dt}a(-t) = \frac{1}{\tau} a(-t),$$

we have

$$K = e^{i\alpha} \sqrt{\frac{2}{\tau}}, \quad (44)$$

and similarly

$$\gamma = e^{i\alpha} \sqrt{\frac{2}{\tau}}. \quad (45)$$

The most general EOMs therefore are

$$\dot{a} = -i\omega_0 a - \frac{a}{\tau} + e^{i\alpha} \sqrt{\frac{2}{\tau}} S_i, \quad S_o = -e^{2i\alpha} S_i + e^{i\alpha} \sqrt{\frac{2}{\tau}} a. \quad (46)$$

The positions of the $e^{i\alpha}$ factors are quite clear: it means the time it takes for the input current to arrive at a and goes back. The minus sign reminds us of the phase shift in electromagnetic wave reflection. For the single-mode-single-port case, since α can be shifted by redefining the modes, usually either we choose $e^{i\alpha} = 1$ or $e^{i\alpha} = -1$. In the former convention we have

$$\dot{a} = -i\omega_0 a - \frac{a}{\tau} + \sqrt{\frac{2}{\tau}} S_i, \quad S_o = -S_i + \sqrt{\frac{2}{\tau}} a. \quad (47)$$

We can easily extend the above EOMs to the case with multiple ports, and the only thing needed is to sum over all $1/\tau$ terms and S_i terms, while keeping the $S_o = -S_i + \sqrt{2/\tau}a$ equation for each port, assuming that we only have hidden degrees of freedom that make the ports interact with each other or otherwise we should include a non-trivial scattering matrix to link the input and output modes of different ports.

From the equations concepts like impedance can then be defined: we can switch to the frequency domain, and we will find

$$\frac{S_o(\omega)}{S_i(\omega)} = -\frac{\omega - \omega_0 - i/\tau}{\omega - \omega_0 + i/\tau}. \quad (48)$$

Note that the relation between $S_i(\omega)$ and $S_o(\omega)$ – which, when there is no energy degeneracy, are just the value of *one* mode at $t = -\infty$ and $t = \infty$ – is basically a relatively trivial scattering matrix and is a pure-state, unitary relation, even through there is “dissipation” from the perspective of mode a . It’s clear that Markov approximation doesn’t make the system “noisy”; that’s to say, Markov approximation doesn’t necessitate mixed-state formalisms, like the quantum Langevin equation: we need to explicitly declare that the bath is indeed noisy to finally get the quantum Langevin equation. The scattering matrix element (48) can actually be derived using Feynman diagrams: we are interested in $b(\omega)$ -to- $b(\omega)$ scattering, and according to LSZ reduction formula, the corresponding S -matrix equals to one plus the sum of all amputated diagrams. In the present case we only have one such diagram, which contains one b -to- a vertex, one a propagator, and one a -to- b vertex, and the a propagator has its own self-energy correction due to the presence of b modes, which has been approximated by the Markov approximation. Thus the S matrix is

$$\sim 1 + \left(-i\sqrt{\frac{2}{\tau}}\right) \frac{i}{\omega - \omega_0 + \frac{i}{\tau}} \left(-i\sqrt{\frac{2}{\tau}}\right),$$

which evaluates to exactly (48) (except one minus sign which comes from the definition of S_i and S_o).

We don’t have any “real” mixed state

7 Supermodes

Now consider two modes a and b , each of which has a port linked to it. Again assuming linear coupling between them only and turning off the ports for the moment, we have

$$\dot{a} = -i\omega_a a + K_{ab}b, \quad \dot{b} = -i\omega_b b + K_{ba}a, \quad (49)$$

and the energy conservation condition (or in other words, unitarity) is equivalent to

$$K_{ba} = -K_{ab}^*. \quad (50)$$

We can then diagonalize the dynamic matrix and obtain two resulting modes, often known as **supermodes**. In one of them a, b move in the same direction at the same time, while in the other a, b move in opposite directions at the same time. What makes the supermodes different from ordinary modes is that the ports are attached to a and b , not them, and therefore the time evolution of the supermodes may be rather exotic. If we have two ports, then the phases of the two S_i need to match in a specific way to excite one supermode. If a and b are attached to the same port, then their (spontaneous) radiations to the port cancel each other in one supermode while they enhances each other in another, so the first supermode is incredibly long lived while the second damps rather fast; similarly the first supermode will be incredibly hard to excite if we are only allowed to use the ports.

Consider one case: mode a_2 is connected to a port (S_i, S_o), while mode a_1 is coupled to mode a_2 . When there is no input energy, the equation about the port is

$$S_o = -i\sqrt{\frac{2}{\tau}}a_2. \quad (51)$$

Further, assuming that we have no excitation in the “minus” supermode, we have

$$a_1 = a_2 = \frac{1}{\sqrt{2}}a_+, \quad (52)$$

and therefore the output power is

$$P_{\text{out}} = |S_o|^2 = \frac{2}{\tau} |a_2|^2 = \frac{2}{\tau} \frac{|a_+|^2}{2}, \quad (53)$$

and putting this to the EOM of the energy of a_+ , we find the lifetime of a_+ is 2τ , not τ .

Then let's consider the case where two modes are coupled to one port. The output flux now is

$$S_o = S_i - i\sqrt{\frac{2}{\tau_1}} a_1 - i\sqrt{\frac{2}{\tau_2}} a_2, \quad (54)$$

and again we ignore the input flux. Note that we haven't derived any equation about the coupling between two modes and one port; in most experimental settings, we can divide a large port (say, a waveguide) into several small ports and apply the one-mode-one-port theory to each of them. This leads to (54), but leaves the possibility that S_{o1} and S_{o2} have a phase difference, which is indeed the case for a waveguide if the positions where a_1 and a_2 are connected to the waveguide are far from each other. Also if the two modes are in the same device, nothing prohibits us to have a non-diagonal matrix between S_o and a 's. Here for the sake of simplicity we ignore this possibility. Suppose we only have excitations in the $+$ supermode, and this means

$$a_1 = a_2 = \frac{a_+}{\sqrt{2}}. \quad (55)$$

The output flux is

$$S_o = -i \left(\sqrt{\frac{1}{\tau_1}} + \sqrt{\frac{1}{\tau_2}} \right) |a_+|^2, \quad (56)$$

and therefore the loss of the $+$ supermode is enhanced (just consider the case $\tau_{1,2} = \tau$).

On the other hand, for the $-$ supermode we have

$$a_1 = \frac{a_+}{\sqrt{2}}, \quad a_2 = -\frac{a_+}{\sqrt{2}}, \quad (57)$$

and S_o now is

$$S_o = -i \left(\sqrt{\frac{1}{\tau_1}} - \sqrt{\frac{1}{\tau_2}} \right), \quad (58)$$

and clearly when $\tau_1 = \tau_2$ we don't have damping of a_- at all!

Modes that never damp and correspondingly can't be excited by S_i are known as **dark modes**.

8 Scattering matrices

Suppose we have a two-port system; a_1 and a_2 are input modes of the two ports, while b_1 and b_2 are output modes of the two ports. The relation between $a_{1,2}$ and $b_{1,2}$ is called the scattering matrix of the system. What it means by "the relation between ..." needs some explanation.

This scattering matrix is not the same as the two-particle scattering matrix: if the two-port system is linear we still need a 2×2 matrix to describe the relation between the input and output fields, because the input field consists of at least two modes ($a_{1,2}$) and so does the output field; the scattering matrix here however indeed is the same as the *one-particle* S -matrix, although under a different basis. The input modes are obtained by diagonalizing the free Hamiltonian (e.g. a Hamiltonian whose eigenstates are plane waves) *in the "input" subspace*, which means the Hilbert space form by collecting all waves going towards the scattering center, and the output modes are defined accordingly. The normalization of the input and output wave functions should be chosen wisely such that

$$|\text{scattering stationary state}\rangle = |\text{one input state}\rangle + \sum c_i |\text{output states}\rangle \quad (59)$$

is already normalized under the inner product defined in the *whole* Hilbert space when

$$\sum |c_i|^2 = 1. \quad (60)$$

Of course, that means the amplitudes of the input and output wave functions are smaller than one. For a one dimensional scattering problem where the scattering center is at $x = 0$, this means a input state look like

$$\psi(x) = \frac{1}{\sqrt{L}} \begin{cases} e^{ikx}, & x < 0, \\ 0, & x > 0. \end{cases} \quad (61)$$

where L is the “sample size” and $-L/2 \leq x \leq L/2$. Note that the normalization coefficient is not the one used in the half-space but the one used in the whole space. Based on the settings above, it can be proved that the S matrix defined in ordinary scattering matrix,

$$S = U(\infty, -\infty) = \sum_i e^{-i\omega_i t} |\text{stationary state } i\rangle \langle \text{stationary state } i|$$

is exactly the same as what you get by arranging coefficients of output states in scattering stationary states that contain only one input state component.

Until now we are working in the Schrodinger picture. Turning to the Heisenberg picture, we know the time evolution of annihilation operators (and hence field operators) in the Heisenberg picture is the same as the time evolution of the wave function in the Schrodinger picture (consider how $\langle 0|a|\psi\rangle$ involves, which is independent to the picture used), and hence S can also be seen as the time evolution operator of mode annihilation operators, and hence sometimes people write $b_i = S_{ij}a_j$, and thus the scattering matrix becomes “the relation between the output modes and the input modes”. This understanding of equations with the form of $b_i = S_{ij}a_j$ no longer works when two modes are connected together to form a close loop, but things like $b_i = S_{ij}a_j$ can still be understood as shorthands for analyses of *wave functions* in a stationary state.

We also note that the scattering matrix formalism is itself a *pure state* formalism: in high-energy physics the scattering matrix is related to scattering experiments which involve measurement, but the scattering matrix itself is evaluated thoroughly in the pure state formalism. This is also the case here.

Scattering matrices often have **reciprocity**. The reciprocal condition for the Green’s function $G_\omega(\mathbf{r}, \mathbf{r}')$ is quite simple:

$$G_\omega(\mathbf{r}, \mathbf{r}') = G_\omega(\mathbf{r}', \mathbf{r}). \quad (62)$$

This condition means the *transpose* of the scattering matrix is just the origin scattering matrix. We note that this is *not* time reversal symmetry. In terms of the scattering matrix, reciprocity means

$$S^\top = S. \quad (63)$$

The energy conservation condition is equivalent to unitarity. This can be easily proven: the energy conservation condition is equivalent to

$$b^\dagger b = a^\dagger a, \quad (64)$$

where a is the column vector containing all input modes and b is the column vector containing all output modes. Since $b = Sa$, we have

$$a^\dagger (S^\dagger S - I) a = 0, \quad (65)$$

and hence (TODO: corner case)

$$S^\dagger S = I. \quad (66)$$

Time reversal symmetry is almost always true; of course when an external magnetic field is present we don’t have time reversal symmetry, but this merely suggests at the necessity to change the direction of the magnetic field when doing the time reversal operation. Here we define (TODO: equivalent definitions: the direction of p is changed; or $i \rightarrow -i$)

$$\text{TR}[Q(t)] = Q^*(-t). \quad (67)$$

Time reversal symmetry means

$$S^* = S^{-1}. \quad (68)$$

There is a lot of abuse (or confusing usage) of the term *breaking the time reversal symmetry*. As is mentioned, when we have a strong magnetic field it’s often said that the time reversal

symmetry is broken; when the magnetic field is included as a part of the system (not an arbitrary external field), time reversal symmetry isn't truly broken, but reciprocity is indeed broken. Consider a circulator of electrons: if an electron is injected into it through one port and we record the port from which the electron goes out, an electron going into the second port doesn't go out of the first port. Of course we can still say time reversal symmetry is broken here if we keep the magnetic field not flipped when we apply the time reversal operation.

9 Noise

Often the modes going in and out in the ports are indeed noisy, and in this case the temporal coupled theory becomes Langevin-like; the Langevin-like theory can of course be derived from the intuition that energy loss shouldn't make a theory about *a* non-unitary and some sort of noise has to be there to make the theory unitary, or the intuition that coupling with the external environment. From equipartition theorem we can see that the power spectral density going into the system, i.e. the power per unit of frequency, is

$$\frac{dP}{d\omega} \simeq \frac{k_B T}{2\pi}. \quad (69)$$

Since now the modes in the ports are indeed noisy, we can study them using theories about stochastic signals. To define the average of $s(t)$ or $s(t_1)s(t_2)$ or \dots , we can either take the ensemble average (a probabilistic average according to the probabilistic distribution of trajectories) or time average (average over a large period of time). In physics the most frequent types of noises are stationary, meaning that the the occasion we take the average doesn't matter, and ergodic, meaning that time average is equivalent to ensemble average.

The (two-point) correlation function, $\langle s^*(t_1)s(t_2) \rangle$, can actually be measured directly experimentally. We can adapt a interferometer to do it: instead of measuring the *sum* of the two signals from two paths, we measure the *product* of the two signals. Because of the stationary nature of noises, we have

$$\langle s^*(t)s(t+\tau) \rangle = C_s(\tau). \quad (70)$$

By Fourier transform, if we choose the following normalization:

$$s(t) = \int_{-T/2}^{T/2} d\omega s(\omega) e^{-i\omega t}, \quad (71)$$

we have

$$C_s(\tau) = \lim_{T \rightarrow \infty} \frac{2\pi}{T} \int_{-\infty}^{\infty} d\omega \langle |s(\omega)|^2 \rangle e^{-i\omega\tau} =: \int_{-\infty}^{\infty} d\omega C_s(\omega) e^{-i\omega\tau}, \quad (72)$$

and therefore

$$C_s(\omega) = \lim_{T \rightarrow \infty} \frac{2\pi}{T} \langle |s(\omega)|^2 \rangle = \frac{1}{\delta(0)} \langle |s(\omega)|^2 \rangle. \quad (73)$$

This is the Wiener–Khinchin theorem in the specific case of time averaging. The theorem can be proven as well with ensemble average; this is more easily done by starting from $\langle s^*(\omega)s(\omega) \rangle$. As $s(\omega)$ is just a mode in $s(t)$, this means the two-point correlation function and the power spectral density are related by Fourier transform.

This means from the power spectral density of a noise we can know a lot about its time correlation. For example, a truly white noise has $\delta(\tau)$ time correlation. Or conversely if we need a wide power spectral density we then need an extremely narrow pulse. Note that $C_s(\omega)$ equals to $\langle s^*(\omega)s(\omega') \rangle$ divided by a common $\delta(\omega - \omega')$ factor: as long as the system has time translational symmetry, in the frequency domain correlation function we always have a $\delta(\omega - \omega')$ factor. It's what *else* there is besides this δ -function factor that $C_s(\omega)$ describes.

Now we turn to the temporal mode coupling theory. We still have

$$\dot{a} = -i\omega_0 a + \frac{a}{\tau} + \sqrt{\frac{2}{\tau}} b_i, \quad (74)$$

where b_i is just S_i . The correlation function of b_i can be derived from the equipartition theorem; for a one-dimensional transmission line it's

$$\langle b_i^*(\omega)b_i(\omega') \rangle = \frac{k_B T}{2\pi} \delta(\omega - \omega'), \quad (75)$$

because by definition we know

$$P = \langle b_i^*(t)b_i(t) \rangle = C_{b_i}(t) = \int C_{b_i}(\omega) d\omega, \quad (76)$$

and therefore

$$\frac{dP}{d\omega} = \frac{k_B T}{2\pi} \Rightarrow C_{b_i}(\omega) = \frac{k_B T}{2\pi}, \quad (77)$$

and from Wiener–Khinchin theorem we have

$$\langle b_i^*(\omega)b_i(\omega') \rangle = \frac{k_B T}{2\pi} \delta(\omega - \omega'). \quad (78)$$

The correlation function of a can be calculated from

$$a(\omega) = \sqrt{\frac{2}{\tau}} \frac{b_i(\omega)}{i(\omega_0 - \omega) + \frac{1}{\tau}}. \quad (79)$$

We find

$$\langle a^*(\omega)a(\omega') \rangle = \frac{1}{\tau} \frac{k_B T}{\pi} \frac{1}{(\omega - \omega_0)^2 + \frac{1}{\tau^2}} \delta(\omega - \omega'). \quad (80)$$

The equipartition theorem for a can then be directly verified: we have

$$\int \int d\omega d\omega' \langle a^*(\omega')a(\omega) \rangle = k_B T. \quad (81)$$

The spectrum of the reflected wave can be calculated as well. From

$$b_o = -b_i + \sqrt{\frac{2}{\tau}} a, \quad (82)$$

we get

$$b_o(\omega) = \frac{\frac{1}{\tau} - i(\omega_0 - \omega)}{\frac{1}{\tau} + i(\omega_0 - \omega)} b_i(\omega), \quad (83)$$

and

$$\langle b_o^*(\omega')b_o(\omega) \rangle = \frac{k_B T}{2\pi} \delta(\omega - \omega'). \quad (84)$$

So the energy distribution of b_o is exactly the same as the energy distribution of b_i . The only difference between b_i and b_o is a phase factor.

Now we take a slightly different approach. If we know nothing about the port stuff, can we still get the correlation function of the noise term coming together with the dissipation term? From

$$\dot{a} = -i\omega_0 a - \frac{a}{\tau} + \eta(t) \quad (85)$$

we get

$$\frac{d}{dt} |a|^2 = -\frac{2}{\tau} |a|^2 + (a^* \eta + \eta^* a). \quad (86)$$

After taking the time average on both sides, since $\frac{d}{dt} \langle |a|^2 \rangle = 0$, we find

$$\frac{2}{\tau} \langle a^* a \rangle = \langle a^* \eta + \eta^* a \rangle. \quad (87)$$

The relation between a and η is given by

$$a(t) = \int_{-\infty}^t e^{-i\tilde{\omega}_0(t-t')} \sqrt{\frac{2}{\tau}} \eta(t') dt', \quad (88)$$

where

$$\tilde{\omega}_0 = \omega_0 - \frac{i}{\tau}. \quad (89)$$

So now the RHS of (87) becomes an expression of $\langle \eta(t)\eta(t') \rangle$. Now we have to introduce the approximation that

$$\langle \eta(t)\eta(t') \rangle = A\delta(t - t') \quad (90)$$

to simply the following procedure. Using

$$\int dt \Theta(t - t') \delta(t - t') = \frac{1}{2},$$

we find

$$\frac{2}{\tau} \langle a^* a \rangle = \frac{2}{\tau} k_B T = A,$$

and eventually

$$\langle \eta(t) \eta(t') \rangle = \frac{2}{\tau} k_B T \delta(t - t'). \quad (91)$$

10 Input-output formalism

Now we quantize the temporal coupling theory. Since we were mostly working on physical quantity EOMs before, the forms of the eventual effective EOMs should be exactly the same in classical and quantum systems. The main difference is the interpretation of the variables. For example, in the quantum case, $\langle a^\dagger a \rangle$ has a clear meaning: the number of photons in the mode a .

The Hamiltonian takes the form

$$H = \hbar \omega_0 a^\dagger a + \sum_k \hbar \omega_k b_k^\dagger b_k - \hbar \sum_k \tilde{g}_k a^\dagger b_k + \text{h.c.}, \quad (92)$$

and the EOMs are

$$\dot{b}_k = -i\omega_k b_k + i g_k^* a, \quad \dot{a} = -i\omega_0 a + i \sum_k g_k b_k. \quad (93)$$

The following derivation follows exactly what is done in the classical case; now however we want to make sense of what exactly are S_i and S_o . Assume that the port is indeed a transmission line, and k used to label b_k then is the wave vector. Suppose the length of the port is L , and we have

$$k = \frac{2\pi}{L} m, \quad (94)$$

where m is an integer. The real space b is

$$b(x) = \frac{1}{\sqrt{L}} \sum_k e^{ikx} b_k, \quad (95)$$

and its “value”, like the expectation value, gives the shape of wave packets in a quantum state. Here the normalization is equivalent to assuming that

$$U_{\text{tot}} = \int dx |b(x, t)|^2 = \sum_k b_k^\dagger b_k. \quad (96)$$

Of course we can split an arbitrary wave packet into a right-mover and a left-mover, and therefore we define

$$b_+(x) = \frac{1}{\sqrt{L}} \sum_{k>0} e^{ikx} b_k, \quad (97)$$

and

$$b_-(x) = \frac{1}{\sqrt{L}} \sum_{k<0} e^{ikx} b_k. \quad (98)$$

Going to the continuous limit, we make the following substitution:

$$b(k)^{\text{continuous } k} = \frac{\sqrt{L}}{2\pi} b_k^{\text{discrete } k}. \quad (99)$$

This leads to

$$b(x) = \int dk e^{ikx} b(k). \quad (100)$$

Next we evaluate the noise spectrum. We have (here we work in the classical case)

$$\begin{aligned}
\langle b^*(k)b(k') \rangle &= \lim_{L \rightarrow \infty} \left\langle \frac{\sqrt{L}}{2\pi} b_k^* \frac{\sqrt{L}}{2\pi} b_{k'} \right\rangle = \frac{1}{2\pi} \lim_{L \rightarrow \infty} \underbrace{\frac{L}{2\pi}}_{1/\Delta k} \langle b_k^* b_{k'} \rangle \\
&= \frac{1}{2\pi} \lim_{L \rightarrow \infty} k_B T \delta_{kk'} / \Delta k \\
&= \frac{k_B T}{2\pi} \delta(k - k').
\end{aligned} \tag{101}$$

11 Quantum noise

Replacing a and b by their quantum operator counterparts leads us into the regime of quantum noise and fluctuation, in the Heisenberg picture. This also relates the most natural way to express the rotating wave approximation to the standard definition of rotating wave approximation in quantum mechanics, i.e. by applying a picture transform.

The key thing that needs to be highlighted is now dynamic variables don't always commute. This means we don't always have $C(\tau) = C(-\tau)^\dagger$ (TODO). One way is to always work with $(AB + BA)/2$.

The time evolution operator of a time dependent Hamiltonian is

$$U = \mathcal{T} \exp \left(-\frac{i}{\hbar} \int dt' H(t') \right). \tag{102}$$

When the Hamiltonian is time dependent but we have

$$[H(t), H(t')] = 0, \tag{103}$$

this can be simplified into

$$U = \exp \left(-\frac{i}{\hbar} \int dt' H(t') \right). \tag{104}$$

An example of time dependent Hamiltonian commuting with itself at different times is

$$H = (\omega + \Delta\omega(t)) a^\dagger a, \tag{105}$$

while a Hamiltonian that doesn't commute with itself at a different time is

$$H = \omega a^\dagger a + f(t)(a + a^\dagger). \tag{106}$$

From $[a, a^\dagger] = 1$ it follows the trivial result $\frac{d}{dt}[a(t), a(t)^\dagger] = 0$. This result helps us to determine some basic behaviors of the Langevin term in a system coupled to a Markovian external bath. From the equation without the Langevin term

$$\dot{a} = -i\omega_0 a - \frac{1}{\tau} a$$

we get the troublesome commutation relation

$$\frac{d}{dt}[a, a^\dagger] = -\frac{2}{\tau}[a, a^\dagger],$$

which is expected since unitarity is broken. Now starting from

$$\dot{a} = -i\omega_0 a - \frac{1}{\tau} a + \eta \tag{107}$$

we get

$$\frac{d}{dt}[a, a^\dagger] = -\frac{2}{\tau}[a, a^\dagger] + [\eta, a^\dagger] + [a, \eta^\dagger], \tag{108}$$

and since the LHS is zero, and therefore $[a, a^\dagger]$ is always unity, we find

$$[\eta, a^\dagger] + [a, \eta^\dagger] = \frac{2}{\tau}. \tag{109}$$

This is a condition that *all* Langevin “noises” (or non-thermalized port operators) have to follow.

Actually we can even derive the Langevin equation under some assumptions. We know there is a direct relation between a and η , namely

$$a(t) = \int \Theta(t-t') dt' e^{-i(t-t')(\omega+i/\tau)} \eta(t'), \quad (110)$$

and therefore

$$[a, \eta] = \int_0^t e^{-i(\omega+i/\tau)(t-t')} [\eta(t'), \eta(t)^\dagger] dt'. \quad (111)$$

If we assume (here is the assumption!) that

$$[\eta(t'), \eta^\dagger(t)] = D\delta(t-t'), \quad (112)$$

we find (note that $\int_{-\infty}^0 \delta(t) dt = 1/2$)

$$[a, \eta] = \int_0^t e^{-i(\omega+i/\tau)(t-t')} D\delta(t-t') dt' = \frac{1}{2}D,$$

and we eventually find

$$D = \frac{2}{\tau}, \quad [\eta(t), \eta^\dagger(t')] = \frac{2}{\tau}\delta(t-t'). \quad (113)$$

12 Quantum bath and quantum noise

We eventually move to the fully quantum treatment of noise. We normalize $b(\omega)$ so that the correlation function reads

$$\langle b^\dagger(\omega)b(\omega') \rangle = n_{\text{th}}(\omega) \frac{\delta(\omega-\omega')}{2\pi}, \quad \langle b(\omega)b^\dagger(\omega') \rangle = (n_{\text{th}}(\omega) + 1) \frac{\delta(\omega-\omega')}{2\pi}. \quad (114)$$

The first equation, the equation about the normal-ordered correlation function, gives the rate of dissipation, because it shows the expectation value of photons emitted to the port. It vanishes when $T \rightarrow 0$. The second equation replicates the bosonic canonical commutation relation and shows the quantum nature of the noise. The bosonic distribution $n_{\text{th}}(\omega)$ depends on ω , and it also depends on the temperature and therefore thermodynamics still enters the correlation function. The commutation relations are still

$$[b(\omega), b^\dagger(\omega')] = \frac{\delta(\omega-\omega')}{2\pi}, \quad (115)$$

which comes from the canonical commutation relation of oscillators. We should now be very careful about the order of the operators: the total photon flux is

$$\Phi = \int d\omega \Phi(\omega) = \int d\omega \int d\omega' \langle b^\dagger(\omega)b(\omega') \rangle. \quad (116)$$

Here $\langle \dots \rangle$ is the average operation in a thermal state; when $T \rightarrow 0$, we should get

$$\langle [b(\omega), b(\omega')] \rangle = \frac{\delta(\omega-\omega')}{2\pi}. \quad (117)$$

Now we can move forward to calculating the fluctuation of a . Defining $\eta = \sqrt{\frac{2}{\tau}}b$, and noticing that (here a is normalized in the same way as b does)

$$a(\omega) = \frac{\eta}{i(\omega_0 - \omega) + \frac{1}{\tau}}, \quad (118)$$

we find

$$\langle n_a \rangle = \frac{1}{2\pi} \frac{2}{\tau} \int \frac{n(\omega) d\omega}{(\omega - \omega_0)^2 + \frac{1}{\tau^2}} \approx n(\omega_0), \quad (119)$$

where at the last step we have assumed that τ is large enough so that the integrand is an approximate δ -function. This result also follows from the classical equipartition theorem; the difference is $n(\omega)$ deviates from the classical case. We can also verify using exactly the same routine in the classical case that

$$b_o(\omega) = e^{i\theta(\omega)} b_i(\omega), \quad (120)$$

so that we have detailed balance of energy and the system is indeed thermally equilibrium.

13 Amplification

The (quantum) Langevin formalism describes not only loss but also gain, the latter appearing when there is *population inversion* in the medium as in, say, laser. In the so-called *small σ limit* (i.e. the limit in which the scattering cross section between light field and the atoms is small) we can ignore what happens in the medium and assume the medium's energy never gets depleted. In the more general case we also need to consider the competition between repopulation rates and various decay rates.

We consider a prototypical laser device. The reason we have gain is we are letting the pulse going around three mirrors over and over again and between two of the mirrors there is a gain medium. Of course, the round trip time is small, so the details of the round trip can be ignored in the EOM of the photon mode a . In each round trip, suppose there are N two-level systems in total, the rate equation of the photon flux is (here Φ_{vac} is a term that we are going to relate with fluctuation)

$$\frac{d\Phi}{dz} = \Phi\sigma N(p_e - p_g) + \Phi_{\text{vac}}\sigma p_e. \quad (121)$$

The change of the photon flux after a thin sample with a width of Δz is

$$\Delta\Phi_{\text{tot}} = \Delta\Phi_{\text{stim}} + \Delta\Phi_{\text{sp}} = (G - 1)\Phi_{\text{input}} + \Phi_{\text{vac}}(G - 1)n_{\text{sp}}, \quad (122)$$

where

$$G - 1 = \sigma N(p_e - p_g)\Delta z, \quad n_{\text{sp}} = \frac{p_e}{p_e - p_g}. \quad (123)$$

This is a rate equation formalism of the gain medium, which ignores non-trivial time evolution of ρ_{eg} of the medium; usually the rate equation works well enough for the basic features of, say, the laser line width; it doesn't work well enough for more coherent phenomena which on the other hand are not touched here anyway.

Now we compare the above with the photon fluxes we get from (note the sign before the a/τ term; here τ_g means the atom can't stay on the ground state for too long)

$$\dot{a} = -i\omega_0 a + \frac{1}{\tau_g} a + \eta. \quad (124)$$

When the sample is not thick, approximately we can relate the photon flux to the photon number n :

$$\langle\Phi\rangle \approx \frac{\langle n\rangle}{V} v_g, \quad (125)$$

and therefore from this proportional relation between Φ and n and (122), we find

$$\frac{d\langle n\rangle}{dt} = \frac{G - 1}{\tau_{\text{rt}}} + \underbrace{\frac{G - 1}{\tau_{\text{rt}}} n_{\text{sp}} n_{\text{vac}}}_{R_{\text{sp}}}, \quad (126)$$

and from the fact that n follows an EOM like $\dot{n} = \frac{2}{\tau_g} n + \dots$,

$$\frac{2}{\tau_g} = \frac{G - 1}{\tau_{\text{rt}}}, \quad (127)$$

where “rt” means “round trip”. As we mentioned above, this equation ignores the details of how $|a|^2$ grows: here we essentially assume it grows in a simple, exponential way, although since the gain medium is not necessarily everywhere in the laser device, if we check the details in each round trip, we will find $|a|^2$ doesn't really grow in such a simple way. But since τ_{rt} is expected to be very small, we can “smoothen” $|a|^2$'s time evolution and work in a longer time scale, where $|a|(t)^2$ looks quite close to an exponentially growing function.

Now we work to refine the picture and find what η does and how it relates to Φ_{vac} . The complete EOM of n is

$$\frac{d\langle a^\dagger a\rangle}{dt} = \frac{2}{\tau_g} \langle a^\dagger a\rangle + \langle \eta^\dagger a + a^\dagger \eta\rangle. \quad (128)$$

To keep the system unitary, we need $[a(t), a^\dagger(t)]$ to always be unity, and we find

$$[\eta(t), \eta(t')] = -\frac{2}{\tau_g} \delta(t - t') \quad (129)$$

works (of course it's still not the only possible solution). Strangely, η now looks like a *creation* operator – in the theory of loss it behaves like a *annihilation* operator. η should be normalized so that

$$R_{\text{sp}} = \langle \eta^\dagger a + a^\dagger \eta \rangle, \quad (130)$$

the LHS agreeing with (126). a can be solved and expressed in terms of η , and by energy conservation we find

$$\langle \eta^\dagger(t) \eta(t') \rangle = \frac{2}{\tau_g} n_{\text{sp}} \delta(t - t'), \quad \langle \eta(t) \eta^\dagger(t') \rangle = \frac{2}{\tau_g} (n_{\text{sp}} - 1) \delta(t - t'), \quad (131)$$

and hence

$$R_{\text{sp}} = \frac{2}{\tau_g} n_{\text{sp}}. \quad (132)$$

Our derivation for a Langevin formalism of amplification has been completed here.

We can have a Langevin formalism for a system with both gain and loss; this happens when the laser has leaky mirrors and the atoms have spontaneous decays that contribute nothing to a . And there is no necessity that the ports related to loss are thermalized: when there is such a port, we can calculate the response function of the laser cavity which, in this case, is an optical amplifier. For example, the following system

$$\dot{a} = -i\omega_0 a + \left(\frac{1}{\tau_g} - \frac{1}{\tau_a} - \frac{1}{\tau_e} \right) a + \eta_g + \eta_a + \sqrt{\frac{2}{\tau_e}} b_i, \quad b_o = -b_i + \sqrt{\frac{2}{\tau_e}} a. \quad (133)$$

The system has one gain channel, one absorption channel and one channel to direct energy out of the cavity to a port we have full control of. The frequency-resolved photon flux in the port, when there is no input, is

$$\Phi(\omega) = \frac{1}{2\pi} \frac{2}{\tau_e \tau_g} \frac{n_{\text{sp}}}{(\omega - \omega_0)^2 + \left(\frac{1}{\tau_g} - \frac{1}{\tau_e} - \frac{1}{\tau_a} \right)^2}. \quad (134)$$

This means the gain tunes the line width and changes the line width. The output photon flux in the port, when there is no input, is

$$\Phi = \frac{2}{\tau_e \tau_g} \frac{n_{\text{sp}}}{\tau_e^{-1} + \tau_a^{-1} - \tau_g^{-1}}. \quad (135)$$

It can be seen that when

$$\frac{1}{\tau_g} = \frac{1}{\tau_e} + \frac{1}{\tau_a}, \quad (136)$$

the output is infinite: this means right at the threshold of laser – the threshold that stimulated emission dominates spontaneous emission – if we assume the medium always keeps its population inversion state, infinite energy flows into the medium and then the output laser beam! This of course is not physical and usually laser saturation effects appear before this happen.

Now we consider the line width. For convenience we ignore port b at the moment. The EOM of the laser intensity always takes the form of

$$\frac{d}{dt} \langle |a|^2 \rangle = - \left(\frac{2}{\tau_e} - \frac{2}{\tau_g} \right) \langle |a|^2 \rangle + R_{\text{sp}}. \quad (137)$$

Since the spontaneous emission rate R_{sp} is always larger than zero, when the system reaches equilibrium, we always have to have

$$\frac{1}{\tau_e} > \frac{1}{\tau_g}. \quad (138)$$

This means in a physical circumstance – in which $|a^2|$ can't grow unboundedly and has to eventually arrive at equilibrium – the gain is always weaker than the dissipation. It also means

(trivially) that to have a usable laser device that gives you steady output, you have to expect the output to be kind of noisy. We can view the inverse of

$$\Gamma = \frac{2}{\tau_e} - \frac{2}{\tau_g} \quad (139)$$

as the decoherence time: at each time step, a (and the output b , if the b_i, b_o port is turned on) look to be in a coherent state, but $a(0)$ and $a(1/\Gamma)$ almost have no correlation and are no longer coherent to each other. We also note that R_{sp} comes from nothing but η_g , i.e. the noise that comes together with the gain, because η_e has zero expectation.

The line width of the spectrum of a is

$$\langle n(\omega) \rangle = \int d\omega' \langle a^\dagger(\omega) a(\omega') \rangle = \frac{1}{\pi} \frac{n_{sp}/\tau_g}{(\omega - \omega_0)^2 + \Gamma^2/4}, \quad (140)$$

and we find the line width is

$$\Delta\omega_{FWHM} = \Gamma = \frac{2}{\tau_g} \frac{n_{sp}}{\langle n \rangle} \quad (141)$$

which is known as Schawlow-Townes line width. The saturation number is defined when saturation effects reduce the line width to

$$\Delta\omega_{FWHM} = \frac{\Gamma}{2}. \quad (142)$$

We have already mentioned the concept of laser threshold. The reason it is called a threshold can be investigated by calculating the output of the laser. At the threshold we can observe two effects. First, the output power suddenly increases drastically, and second, the output state looks much closer to a coherent state (with some noise) than to a (amplified) thermal state. Question: could these effects be derived from the equations above?

As we have mentioned above, the laser output, when the laser device arrives at equilibrium, always contains noise, and the output state is not a perfect coherent state. If we observe the output state as a whole using one quasi-distribution function, we can see two effects: first, thermal photons from spontaneous emission of course change the output power, so we can see amplitude noise; but more importantly, the phases of thermal photons are totally random, so the predominant noise is actually the phase noise: we can see the peak of the quasi-distribution function does a random walk on a circle.