Causation in Physics: Causal Processes and Mathematical Derivations

Nancy Cartwright

Stanford University

1. Introduction

Physics is, above all else, mathematical. But its causal stories can be told in words. How then does mathematics bear on causal claims in physics? I want to make some very simple philosophical points about derivations, mathematical dependencies, and causal relations in physics. I am going to focus on the question, "How does mathematics provide theoretical support for physics' causal claims?" Theoretical support is only one among a variety of kinds of support that a causal claim may have, but I am not going to discuss how to balance these various kinds of support. Instead I will concentrate entirely on theoretical support and lay out some necessary conditions for the ideal case. Although the points I will make are not very involved philosophically, I will flesh them in with a very detailed example concerning gas lasers.

My first point is that derivations do not provide maps of causal processes. A derivation may start with the basic equations that govern a phenomenon. It may be highly accurate, and extremely realistic. Yet it may not pass through the causes. This is exactly what happens in our laser example. Even a highly accurate derivation of a phenomenon may not support the causal story we want to tell about it.

Nevertheless, I shall argue, a derivation is necessary if the theory is to support the causal story. It is not the derivation itself that provides the support; but rather a kind of back tracing through the derivation to follow to their origins the features that are mathematically responsible for the effect. My second claim, then is that mathematical backtracking of dependencies of the kind I shall describe in section 4 is necessary for complete theoretical support of a causal story.

The third point is that this special kind of backtracking is not sufficient for causal support. This is not surprising if the derivation through which we are tracing did not go via the causes in the first place. In our example we shall see two structurally similar tracings, one which leads back to the causes, and one which does not.

PSA 1984, Volume 2, pp. 391-404

What picks the features targeted by one tracing as correct, and the others not, is the way they can be fitted into the causal process that we already know to be going on. This brings me to the last of my rather simple points: you can't get causes just from mathematics.

2. The Lamb Dip

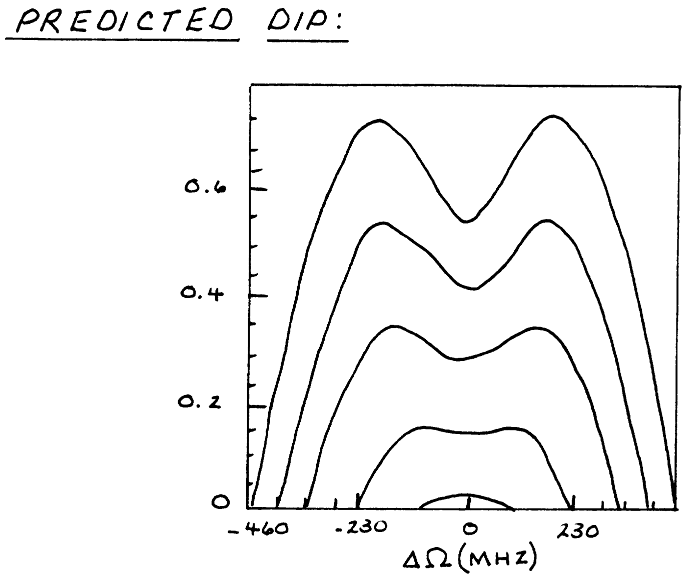
The effect we shall study occurs in gas lasers. It is called the Lamb dip, after Willis Lamb, who first predicted its occurrence. The dip occurs in the graph of laser intensity versus cavity frequency, as we see in Figure 2.1. The atoms in the cavity have a natural transition frequency, ω; the cavity also has a natural frequency, ν, depending on its physical structure. It is natural to expect that the intensity will be greatest when the cavity frequency matches the atomic transition. Indeed, Lamb reports, "I naively expected that the laser intensity would reach a maximum value when the cavity resonance was tuned to the atomic transition frequency. To my surprise, it seemed that there were conditions under which this would not be the case There could be a local minimum, or dip, when the cavity was tuned to resonance [i.e., cavity frequency = transition frequency]. I spent a lot of time checking and rechecking the algebra, and finally had to believe the results." (Lamb 1984, p. 553). Lamb didn't know it at the time, but the Lamb dip is caused by a combination of saturation, with its consequent hole-burning, and Doppler-shifting, which occurs for the moving atoms in a gas laser such as helium-neon.

I will explain in a moment what hole-burning and Doppler-shifting are, but first I want to give a short history of the dip. The concept of hole-burning is due to W.R. Bennett, and it was Bennett who first put together hole-burning and the Lamb dip in print (1962), though both a footnote in the Bennett paper and remarks of Lamb (conversation, Oct. 1, 1984) suggest that the connection was first seen by Gordon Gould. Bennett had been using hole-burning to explain unexpected beat frequencies he had been seeing in helium-neon lasers at Bell Lab in 1961. But, Bennett explains, "Ironically, a much more direct proof of the hole-burning process" is provided by the Lamb dip (1962, p. 58).

Bennett's paper is in Applied Optics in 1962. Lamb's paper was circulating at the time, but was not finally published until 1964. In fact, Lamb had been working on the calculations from the spring of 1961, and he says that he had already seen the dip (which Lamb calls "the double peak" after the humps rather than the trough) by the fall of 1961 (Lamb, conversation Oct. 1, 1984). Lamb wrote both to Bennett and to A. Javan about the prediction. Bennett, who had been measuring intensity versus tuning in the helium-neon laser sent back a tracing of a single peak. Javan answered more favorably, for he had been seeing frequency-pushing effects that could be easily reconciled with Lamb's general treatment. Javan then did a direct experiment to show the dip, which he published later with A. Szöke (Szöke and Javan 1963). The first published report of the dip was by R.A. McFarlane (McFarlane et al. 1963), who attributed earlier failures to see the dip to the use of natural neon, whose two isotopes confound the effect. McFarlane used a single isotope instead, and got the results shown in Figure 2.2 (Lamb, conversation Oct. 1, 1984; also Lamb 1984).

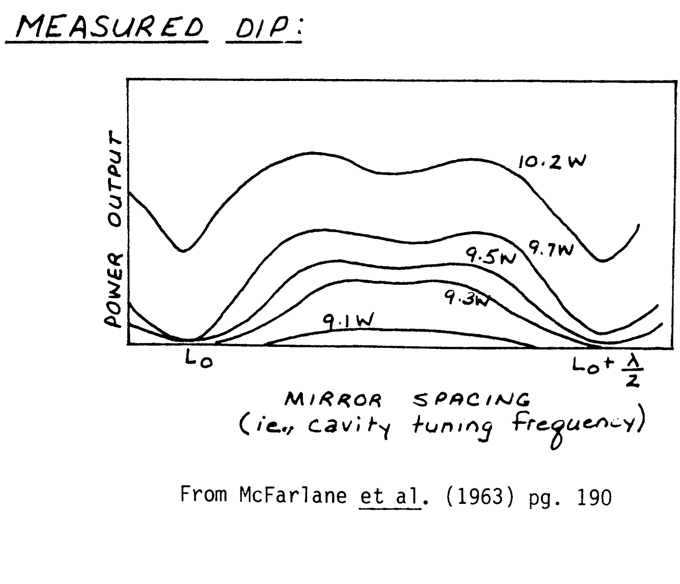
We see that Lamb worked on the paper for three years before it was

Figure 2.1



From Sargent et al. (1974) pg. 153

Figure 2.2



From McFarlane et al. (1963) pg. 190

published. He is in general very methodical, and slow to publish. But there was special motivation in this case for holding back: he did not know what caused the dip. He could predict it, and he knew it existed; but he did not know what physical process produced it. This raises the first philosophical point I want to make: the mathematical derivation of an effect may completely side-step the causal process which produces the effect; and this may be so even when the derivation is both a) faultless and b) realistic.

a) Lamb's mathematical treatment was highly accurate and very careful. Bennett described it as "an extremely detailed and rigorous development of the theory of optical maser oscillation" (1962, p. 58), and that is still the opinion today. In fact Lamb's study of gas lasers was the first full theoretical treatment of any kind of laser, despite the fact that Javan had produced a gas laser at the end of December, 1960, and ruby lasers had been operating since July of 1960. The work of Schawlow and Townes, which was so important for the development of these lasers, was by contrast piecemeal and unsystematic in its presentation.

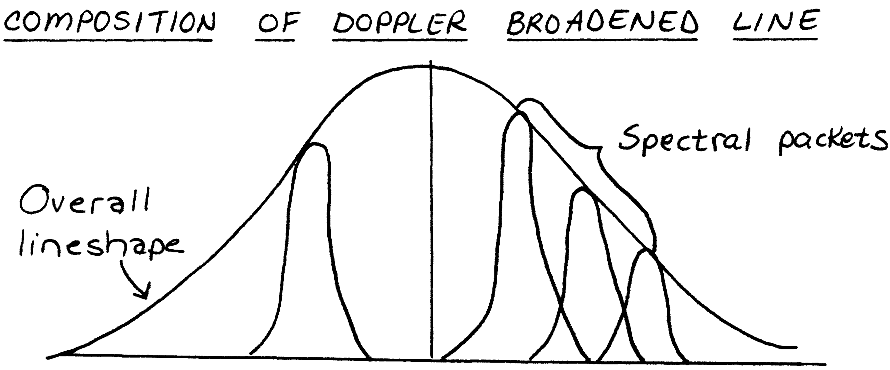
b) The calculations are based on a concrete, realistic model of the gas laser. This contrasts, for example, with an almost simultaneous theoretical treatment by Hermann Haken, which is highly formal and abstract (Haken and Sauermann 1963). Lamb's calculations refer to the real entities of the laser -- the moving gas molecules and the electromagnetic field that they generate; and the equations govern their real physical characteristics, such as population differences in the atoms and the polarization from the field. Nevertheless, the derivation fails to pass through the causal process. We will see exactly how this happened in section 5, but for now I want to stress this single point: a derivation of an effect may be both sound and realistic and yet provide no theoretical insight as to its causes.

3. Causes of the Lamb Dip

For simplicity, we will consider two level atoms with a transition frequency, ω. At thermal equilibrium, most atoms are in the ground state. For laser action we need to pump atoms up, creating a population inversion where the majority of atoms are in the excited state. When that occurs, a signal near the transition frequency will stimulate transitions in the atoms. The size of the atomic response will be proportional both to the applied signal and to the population difference, $\rho_{aa} - \rho_{bb}$. The stimulated emission in turn increases the signal, thereby stimulating an even stronger response. The response does not increase indefinitely because the signal depopulates the upper level, driving the population difference down, till a balance is achieved between the effects of the pumping and the signal. This is called saturation of the population difference.

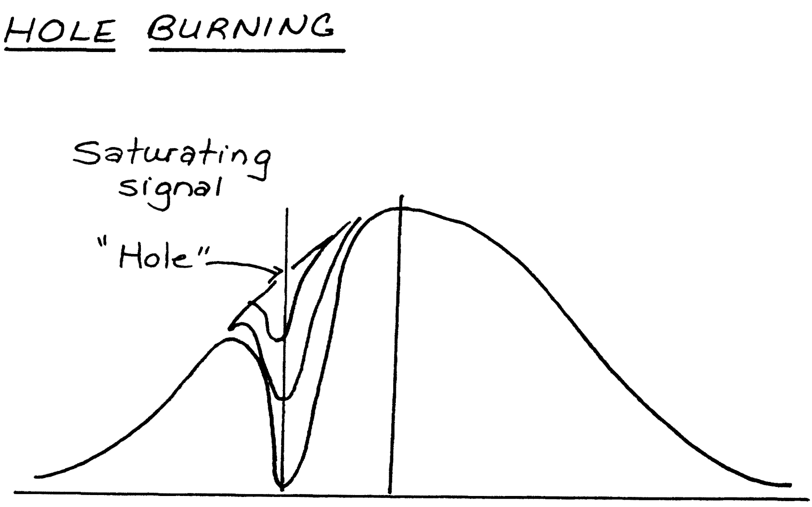
In a laser, oscillation begins when the gain of the beam in the cavity is enough to balance the losses due to things like leakage from the cavity. The intensity of the oscillations builds up till the oscillation saturates the gain, and brings it down. Steady state oscillation occurs at the point where the saturation brings the gain exactly equal to the losses.

Figure 3.1



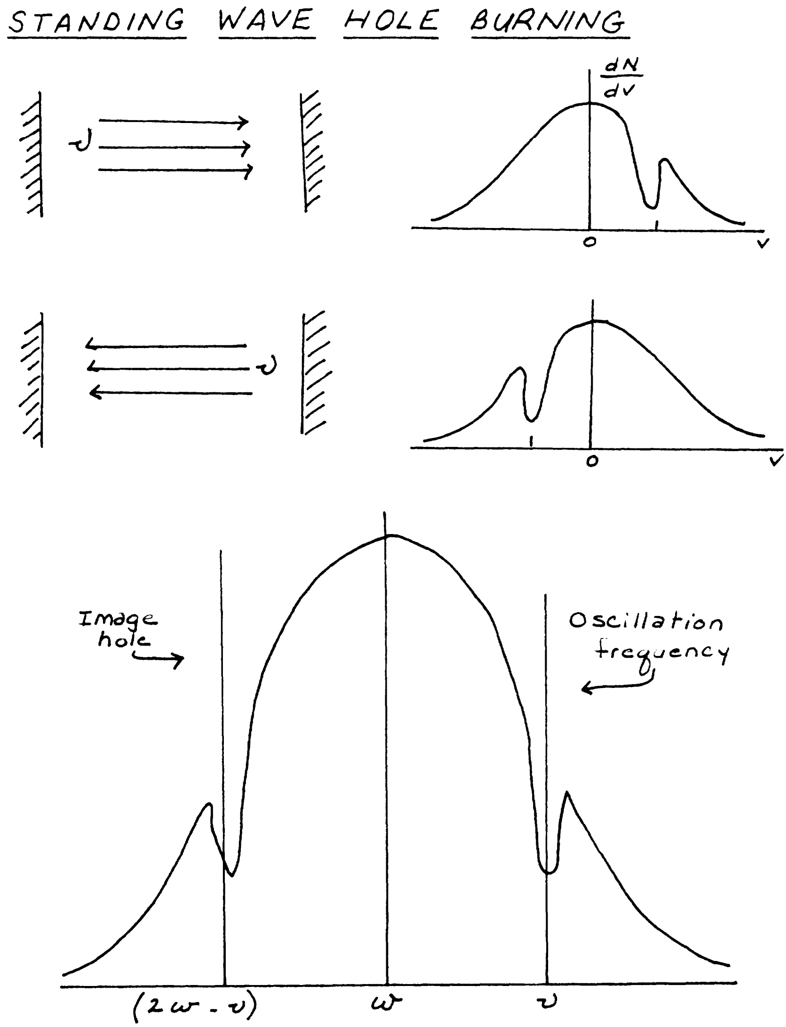
From Siegman (1971)

Figure 3.2



From Siegman (1971)

Figure 3.3



From Siegman (1971)

Saturation produces unexpected effects in gas lasers. In a gas laser the atoms are moving, and the observed spectral line will be much broader than the natural line for the atoms. That is because of Doppler-shifting. The moving atom sees a signal as having a different frequency than does the stationary one. The broadened line is actually made up of separate spectral packets with atoms of different velocities, where each packet itself has the natural line width, as in Figure 3.1. This gives rise to the possibility of hole-burning: an applied signal will saturate uniformly across the natural spectral packet that contains its frequency; but it will have almost no effect on other packets. So if we chart the population difference versus frequency across the Doppler broadened line, we will see a "hole" in the population difference near the saturating frequency. (Figure 3.2)

We now have all the concepts necessary to explain the Lamb dip. We have been discussing the Doppler shift due to the interaction of a moving atom with a travelling wave. In a laser cavity there are two travelling waves, oppositely directed, which superpose to form a single standing wave. So the standing wave interacts with two groups of molecules -- those whose velocities produce the appropriate Doppler-shifted frequency to interact with the forward wave; and those which interact with the backward one. These atoms have equal and opposite velocities. The holes pictured in Figure 3.3 result. As Murray Sargent and Marlan Scully explain, following Lamb's treatment in their Laser Handbook article, ". . . the holes, i.e., lack of population difference, represent atoms which have made induced transitions to the lower state. Hence the area of the hole gives a measure of the power in the laser . . . . For central tuning ($\omega = \nu$), the laser intensity is driven by a single hole because the two holes for v and -v coincide. The area of this single hole [i.e., the power] can be less than that for the two contributing to detuned oscillation provided the Doppler width and excitation are sufficiently large." (1972, p. 80). Hence we get a power or intensity dip at central tuning.

4. Mathematical Justification

We have now seen the qualitative account of what causes the Lamb dip. How does this account relate to Lamb's mathematical derivation? Here again our example is a particularly felicitous one, for this question is explicitly taken up in the advanced-level textbook by Sargent, Scully and Lamb (1974). Here I will summarize their defense (Chapter 10.1) of the hole-burning/Doppler-shift hypothesis.

Steady state oscillation occurs in a laser when the saturated gain, α_g, equals the losses due to structural features of the cavity. Thus, the amount of saturated gain at steady state is fixed. It is given by

$$(4.1)\quad \alpha_g = \int_{-\infty}^{+\infty} dv \, \frac{e^{-v^2/u^2}\left[\mathcal{L}(\omega-\nu+Kv)+\mathcal{L}(\omega-\nu-Kv)\right]}{1+\tfrac{1}{2}I\left[\mathcal{L}(\omega-\nu+Kv)+\mathcal{L}(\omega-\nu-Kv)\right]}$$

where the Lorentzian, $\mathcal{L}_\gamma(\omega) = \gamma^2/\gamma^2 + \omega^2$. For Doppler spreads large compared to the decay constant, γ, the Lorentzian in the numerator disappears on integration, and we have roughly

$$(4.2) \quad \alpha_g^{off} \approx 1/(1 + \tfrac{1}{2} I^{off})$$

$$(4.3) \quad \alpha_g^{on} \approx 1/(1 + I^{on})$$

So in order for $\alpha_g^{off} \approx \alpha_g^{on}$, we must have $I^{on} \approx 1/2\ I^{off}$. (This approximation ignores a factor $\exp[-(\nu-\omega)^2/(Ku)^2]$ in α_g^{off}. But from Sargent, Scully, and Lamb, figure 10-3, we expect to see a relative hump by $\nu - \omega = 230$ MHz for $Ku \approx 2\pi \times 1010$ MHz. Hence this factor is in the order of magnitude of e^{-100}. We are also ignoring a small contribution from the second Lorentzian in α_g^{off}.) The important question: why is there a factor of 1/2 in α_g^{off} and not in α_g^{on}? Because in the denominator of (4.2) only one of the Lorentzians from (4.1) contributes non-negligibly. But in (4.3) both have contributed non-negligibly. It is because both Lorentzians are there that I^{on} must be less than I^{off} to balance the same fixed cavity losses.

So, how do the two different Lorentzians get into the denominator of (4.3)? Sargent, Scully, and Lamb explain, "We recall from the discussion (of the equation that expands the standing wave as the sum of two oppositely directed running waves) that each Lorentzian . . . results from saturation by one running wave of the standing wave field."(1974, p. 154).

In more detail, the Lorentzians get into (4.3) through the population difference, and they enter the expression for the population difference through the saturation factor; that is, the factor that reduces the population difference from what there would be without a field oscillation. Looking at the equation for the population difference, "The Lorentzians . . . show that holes are burned in the plot of (the population difference) $\rho_{aa} - \rho_{bb}$ versus v. Off resonance ($\nu \neq \omega$), one of the Lorentzians is peaked at the detuning value $\omega - \nu = +Kv$, and one at $\omega - \nu = -Kv$, thereby burning two holes . . . . On resonance ($\nu = \omega$), the peaks coincide and a single hole is burned." (Sargent et al. 1974, p. 149). Populations are depleted at two velocities because there are two running waves for the atoms to interact with. "For $\nu > \omega$, an atom moving along the z axis sees the first of the running waves . . . 'stretched out' or Doppler downshifted . . . . Comparison of the equation (which writes the standing wave as a sum of the two running waves) with (the equation used to generate the population difference) reveals that the Lorentzian in (the saturation factor) result(s) from this running wave . . . . Similarly, an atom moving with velocity $-v$ sees the second standing wave downshifted, interacts strongly if the atom with velocity v did, and produces the (second) Lorentzian." (Sargent et al. 1974, p. 150). Thus our story is complete.

The Lamb dip has found its source in the combination of saturation and Doppler broadening, as promised.

The important philosophical point to notice is how Sargent, Scully, and Lamb trace the mathematics and the causal story in exact parallel. To make this point absolutely perspicuous, I present in Chart 1 a summary outline of the mathematical path, matching step-by-step with the causal story. In order to keep the exposition as simple as possible, I leave out a number of parallel stages in both-- for example, the critical role of the polarization.

Notice exactly what Sargent, Scully, and Lamb do here. They do not lay out a derivation. That has been done earlier. Rather, they take an intelligent back-look through the derivation to pick out the cause: First, they isolate the mathematical feature that is responsible for the exact characteristic of the effect in question -- in this case the fact that two Lorentzians contribute off resonance, and only one on resonance; second, they trace the genealogy of this feature back through the derivation to its origin -- here, to the two terms for the oppositely directed running waves; third, they note, as they go, the precise mathematical consequences that the source terms force at each stage. The mathematics supports the causal story when these mathematical consequences traced back through the derivation match stage-by-stage the hypothesized steps of the causal story. This is the kind of support that causal stories need, and until it has been accomplished, their theoretical grounding is inadequate.

We have looked at just one case, but it is not an untypical case in physics. Just staying within this particular example, for instance, we could easily lay out a similar retrospective mathematical tracing for the causal claim that an applied beam saturates the population difference. For contrast, we might look at Bennett's own "hole-burning model" (1962, 1967) which does not (so far as I can reconstruct) allow the kind of backwards causal matching that Lamb's does. I am not going to do that here, but simply summarize my conclusion: no matter how useful or insightful might be the more qualitative and piecemeal considerations of Bennett, they do not provide genuine theoretical support for the causal story connecting saturation and Doppler broadening with the Lamb dip. Only matching back-tracing through a rigorous derivation in a realistic model can provide true theoretical support for a causal hypothesis.

5. How Mathematics Falls Short

I have argued that a mathematical backtrack is necessary for a theory to support a causal story. In this section I want to illustrate that it is not sufficient. To do so, we need to look with some care at Lamb's derivation, and that involves one extra complication. So far, for simplicity, I have overlooked the role of the polarization. That will no longer be possible.

Lamb calls his theory a "self-consistency" treatment. Here is his own description, and diagram: "We understand the mechanism of maser oscillation as follows: an assumed electromagnetic field E (r, t) polarizes the atoms of the medium creating electric dipole moments P_i (r, t) which add up to produce a macroscopic polarization density

Chart 1

Off resonance — $a \Rightarrow b$ = a causes b	Off resonance — $a \rightarrow b$ = b is derived from
read down ↓	read up ↑
Existence of waves running in 2 directions $\Rightarrow$ saturation at the 2 velocities for which the 2 oscillating frequencies are appropriately up- and down-shifted	Each of the two running wave terms $\rightarrow$ one, each, of the two Lorentzians in the saturation factor
Saturation at each velocity $\Rightarrow$ reduction in the population difference at that velocity	each of the Lorentzians in the saturation factor $\rightarrow$ a Lorentzian (in the denominator) that reduces the population difference at the appropriate Doppler-shifting velocity
reduction in the population difference at 2 velocities $\Rightarrow$ greater intensity in the cavity	2 Lorentzians in the denominator of the population difference $\rightarrow$ 2 Lorentzians in the denominator of the gain formula $\rightarrow$ greater intensity in the cavity

P (r, t). This polarization acts further as the source of a reaction field E' (r, t) according to Maxwell's equations. The condition for maser oscillation is th[e]n that the assumed field be just equal to the reaction field." (Lamb 1963, p. 78).

* * * *

Schematic basis for calculation of the properties of the maser oscillator

$E(r,t) \xrightarrow[\text{mechanics}]{\text{quantum}} \langle P_i \rangle \xrightarrow[\text{summation}]{\text{statistical}} P(r,t) \xrightarrow[\text{equations}]{\text{Maxwell's}} E'(r,t)$

Figure 5.1. From Lamb (1963), pg. 78

* * * *

The self consistency assumption gives rise to equations of motion for the population matrix, that is, the matrix which describes the "state" of the population. I will write these equations here in order to point out just one important feature of them.

(5.1) $\dot{\rho}_{ab} = -i\omega\rho_{ab} - \gamma_{ab}\rho_{ab} + i\mathcal{V}(t)(\rho_{aa} - \rho_{bb})$

(5.2) $\dot{\rho}_{aa} = -\gamma_{a}\rho_{aa} + i\mathcal{V}(t)(\rho_{ab} - \rho_{ba})$

(5.3) $\dot{\rho}_{bb} = -\gamma_{b}\rho_{bb} - i\mathcal{V}(t)(\rho_{ab} - \rho_{ba})$

Here $\mathcal{V}$ is the dipole interaction potential between the atoms and the field. We have seen $\rho_{aa} - \rho_{bb}$ earlier. It is the population difference. The elements ρ_{ab} and ρ_{ba} represent the atomic contribution to the polarization. Notice that the changing values of the population difference and of the atomic polarization are yoked together. The rate of change of the population difference at _t_ depends on the polarization at _t_, which has itself been changing in a way dependent on the population difference at earlier times, which depends on the polarization at those times, and so on.

Lamb uses a perturbation analysis to solve these equations. Roughly, it goes like this. To get the 1st order approximation for ρ_{ab}, you insert the _initial_ (t = 0) value for $\rho_{aa} - \rho_{bb}$ in equation (5.1) for $\partial\rho_{ab}/\partial t$, and integrate over time. Similarly, for the 1st order approximation in $\rho_{aa} - \rho_{bb}$. For the second order approximations, in the time integrals you use not the initial values, but the first-order values; and so on.

It turns out that the contributions alternate as you go through the orders, in much the way we think about the problem. First, in 0th order, we have the original inverted population difference interacting with the field, which produces a polarization contribution in 1st order; this in turn makes for a new population contribution in 2nd order; which gives rise to a new polarization in 3rd order. The first order polarization, which has not yet taken into account any feedback from the stimulated atoms, is accurate enough to calculate the threshold of oscillation but not to study steady state oscillation. Hence Lamb concentrated on the 3rd order polarization.

Calculating the intensity from the 3rd order polarization, one discovers the Lamb dip. But Lamb did not see the cause for the dip. Why? Because of an unfortunate shift in the order of integration (see Lamb 1984, p. 553; Lamb 1964 pp. A1448, A1449). In calculating the 3rd order polarization one must, as I said, integrate through time over the 2nd order population difference. Recall, from our discussion of Doppler-shifting, that the population difference varies significantly with the velocity of the atoms. But the macroscopic polarization depends on the total contribution from all the atoms at all velocities, so the calculation must integrate over velocity as well. For mathematical convenience, Lamb did the velocity integral first, then the time integral. He wiped out the velocity information before solving for the population difference. He never saw the two holes at $+\underline{v}$ and $-\underline{v}$ that would account for the dip.

By 1963 Gould and Bennett had suggested the hole-burning explanation, and Lamb had inverted the integrals and derived the velocity dependence of the population difference. The calculation of the intensity, and the backtracking we studied in section 4, is routine from that point on.

But in 1961 and 1962 Lamb had not seen the true causal story, and he was very puzzled. What did he do? He did exactly the kind of mathematical backtracking that I have been advocating. He himself says, "I tried very hard to find [the] origin [of the dip] <u>in my equations</u>." (1984, p. 553, underlining added). Lamb's back tracing <u>is laid out</u> in section 17 of the 1964 paper. Notice how the language is just what one would expect: "The dependence of a typical term . . . is . . ."; "the physical consequence of the appearance of terms involving $\pm$Kv . . . is . . ."; "only the first two possibilities are able to lead to non vanishing interference . . ."

Finally, Lamb concludes, "Physically, one may say that a dominant type of process involves three interactions: first, one with a right (left) running wave at $\underline{t}'''$, then one with a left (right) running wave at $\underline{t}''$, and finally one with a left (right) running wave at t', with the time integrals obeying t - t' = t" - t'" so that the accumulated Doppler phase angle . . . cancels out at time $\underline{t}$." (1964, p. A1448).

Lamb spent a good deal of time trying to figure out the physical significance of these time interval terms but he could not find a causal role for them (conversation, Oct. 13, 1984). As he says in his published historical account, "I never was able to get much insight from this kind of thing. The correct interpretation would have been obvious if I had held back the $\underline{v}$ integration... ." (1984, p. 553).

So, here is a clear case. The mathematical backtracking that I claim is necessary to support a causal story is not sufficient to pick out the causal story. The velocity dependence of the population difference plays a significant role in the physical production of the dip, whereas facts about the time intervals are merely side effects. Yet the mathematical dependencies are completely analogous. The first is singled out rather than the second, not by mathematical backtracking, but by our antecedent causal knowledge, which in this case is highly detailed. Lamb starts with a sophisticated causal picture, whose outline is in Figure 5.1 -- a picture of an applied field which polarizes the atoms and produces dipole moments. The dipole moments add up to a macroscopic polarization that produces a field which polarizes the atoms, and so on. The velocity dependence fits in a clear and precise way into this picture. But no role can be found for the time difference equalities. These time factors find no place in the causal process that we already know to be taking place. Generalizing from this single - though typical - case we arrive at a not very surprising conclusion: mathematics is necessary to support the story, but you can't get new causes out without putting old ones in.

References

→ Bennett, W.R., Jr. (1962). "Gaseous Optical Masers." Applied Optics Supplement 1: 24-61.

-----------------. (1967). The Physics of Gas Lasers. New York: Gordon and Breach.

Haken, H. and Sauermann, H. (1963). "Theory of Laser Action in Solid-State, Gaseous, and Semiconductor Systems." In Quantum Electronics and Coherent Light. (Proceedings of the International School of Physics "Enrico Fermi." Course 31.) Edited by C.H. Townes and P.A. Miles. New York and London: Academic Press. Pages 111-155.

Lamb, W.E., Jr. (1963). "Theory of Optical Maser Oscillations." In Quantum Electronics and Coherent Light. (Proceedings of the International School of Physics "Enrico Fermi." Course 31.) Edited by C.H. Townes and P.A. Miles. New York and London: Academic Press. Pages 78-110.

→ --------------. (1964). "Theory of an Optical Maser." Physical Review 134: A1429-A1450.

→ --------------. (1984). "Laser Theory and Doppler Effects." IEEE Journal of Quantum Electronics QE-20: 551-555.

→ McFarlane, R.A.; Bennett, W.R., Jr.; and Lamb, W.E., Jr. (1963). "Single Mode Tuning Dip in the Power Output of an He-Ne Optical Maser." Applied Physics Letters 2: 189-190.

Sargent, Murray III and Scully, Marlan O. (1972). "Theory of Laser Operation." In Laser Handbook, Volume 1. Edited by F.T. Arecchi and E.O. Schulz-DuBois. Amsterdam: North Holland Publishing Co. Pages 45-114.

--; and Lamb, Willis E., Jr. (1974). Laser Physics. London: Addison-Wesley Publishing Co.

Siegman, A.E. (1971). An Introduction to Masers and Lasers. New York: McGraw-Hill Publishing Co.

→ Szöke, A. and Javan, A. (1963). "Isotope shift and saturation behavior of the 1.15 micron transition of Ne." Physical Review Letters 10: 521-526.