# On loss compensation, amplification and lasing in metallic metamaterials

Sotiris Droulias<sup>1\*</sup>, Thomas Koschny<sup>2</sup>, Maria Kafesaki<sup>1,3</sup> and Costas M. Soukoulis<sup>1,2</sup>

<sup>1</sup>Institute of Electronic Structure and Laser, FORTH, 71110 Heraklion, Crete, Greece <sup>2</sup>Ames Laboratory and Department of Physics and Astronomy, Iowa State University, Ames, Iowa 50011, USA <sup>3</sup>Department of Materials Science and Technology, University of Crete, 71003 Heraklion, Crete, Greece \*e-mail: sdroulias@iesl.forth.gr

# **Supplementary Material**

### 1. 4-level gain system



**Figure S1.** Schematic of the four-level gain medium. The pump is assumed to be homogeneous and emission is assumed to be an electric dipole transition with frequency  $\omega_a$ . The nonradiative decay processes between the *i*<sup>th</sup> and *j*<sup>th</sup> energy levels are described by the  $1/\tau_{ij}$  decay rates.  $N_i$  is the population at level *i*.

The gain material is homogeneously embedded in the dielectric host medium, which has relative permittivity  $\varepsilon_{r,host} = 9$ , and is modeled as a four-level quantum system<sup>14,15,18,19,21,S1</sup>, as shown in Fig. S1. The pumping takes place between the ground state  $(N_0)$  and the third level  $(N_3)$  via the pumping rate  $R_p$ , which is homogeneous, corresponding to electrical pumping as in typical semiconductor lasers<sup>S2</sup>. Emission takes place between the second level  $(N_2)$  and the first level  $(N_1)$ , which are called the *upper* and *lower lasing levels*, respectively. The rate equations that describe our model are:

$$\frac{dN_3}{dt} = -\frac{1}{\tau_{32}}N_3 - \frac{1}{\tau_{30}}N_3 + R_p N_0 \qquad (S1a)$$
$$\frac{dN_2}{dt} = +\frac{1}{\tau_{32}}N_3 - \frac{1}{\tau_{21}}N_2 + \frac{1}{\hbar\omega_a}\mathbf{E} \cdot \frac{\partial\mathbf{P}_a}{\partial t} \quad (S1b)$$
$$\frac{dN_1}{dt} = +\frac{1}{\tau_{21}}N_2 - \frac{1}{\tau_{10}}N_1 - \frac{1}{\hbar\omega_a}\mathbf{E} \cdot \frac{\partial\mathbf{P}_a}{\partial t} \quad (S1c)$$
$$\frac{dN_0}{dt} = +\frac{1}{\tau_{10}}N_1 + \frac{1}{\tau_{30}}N_3 - R_p N_0 \qquad (S1d)$$

where the gain-induced polarization density is given by:

$$\frac{\partial^2 \mathbf{P}_a}{\partial t^2} + \Gamma_a \frac{\partial \mathbf{P}_a}{\partial t} + \omega_a^2 \mathbf{P}_a = -\sigma_a \left( N_2 - N_1 \right) \mathbf{E}$$
(S2)

and  $\sigma_a$  is the coupling strength of  $P_a$  to the electric field and  $\Delta N = N_2 - N_1$  is the population difference that drives the polarization. Depending on the sign of  $\Delta N$ , energy can be transferred from the fields to the medium (i.e. absorption for  $\Delta N < 0$ ) or from the medium to the fields (i.e. amplification for  $\Delta N > 0$ ). The latter case refers to what is widely known as 'population inversion' and is a prerequisite for a material to provide gain.

Depending on the experiment to be simulated, the pump can be either pulsed or constant and this is reflected in the evolution of populations over time; for pulsed pump ( $R_p = R_p(t)$  with  $\max(R_p(t)) = R_{p0}$ ) the populations are initially redistributed, but then relax to their initial condition (Fig. S2a), while for constant pump ( $R_p = \text{const.}$ ) they reach a steady state (Fig. S2b). The pump conditions in Fig. 2 have been chosen in accordance with the discussion in Section 4, i.e. for the constant pump  $R_p = 3 \times 10^8 \text{s}^{-1}$  and for the pulsed pump  $R_{p0} = 3 \times 10^{11} \text{s}^{-1}$  and  $\tau_{pump} = 0.15$  ps. The constant pump reaches a steady state after approximately 200 ps, which is the gain used in the examples of Fig.4-6. This amount of gain is achieved with the pulsed pump, approximately 9 ps after excitation, as is shown in Fig. S2a, as the populations  $N_2$  and  $N_1$  become equal to their CW equivalents in (b).



**Figure S2.** Evolution of populations for our gain system when the pump is (a) pulsed with 0.15 ps duration,  $R_{p0} = 1 \times 10^{11} \text{ s}^{-1}$  and (b) constant with  $R_p = 3 \times 10^8 \text{ s}^{-1}$ . The red dotted lines denote the asymptotic limit of the populations, which have been normalized to the total population  $N_{total} = 5 \times 10^{23}$ . For the considered pump conditions in (a), 9 ps after excitation, the populations  $N_2$  and  $N_1$  become equal to their CW equivalents in (b).

For a certain constant pump rate  $R_p$ , the population difference  $\Delta N = N_2 \cdot N_1$  provided by a certain gain material can be directly calculated from Eq. (S1) with setting the fields and the derivatives to zero, to account for t $\rightarrow \infty$ . Then we obtain the result<sup>23</sup>:

$$\Delta N = \frac{\tau_{30} \left(\tau_{21} - \tau_{10}\right) R_p}{\tau_{32} + \tau_{30} \left[1 + R_p \left(\tau_{10} + \tau_{21} + \tau_{32}\right)\right]} N_{total}$$
(S3)

For another gain material with the exact same parameters, but with different total population  $N_{total}^* = C \times N_{total}$ , as discussed in the main paper (section: Observable regimes for different gain materials), to achieve the same population inversion (gain), a different pump rate  $R_p^*$  has to be applied, as already shown in Fig. 10. To relate the two pump rates, Eq. (S3) can be applied for both systems and equating  $\Delta N$  for both cases leads to:

$$R_{p}^{*} = \left(C\frac{1}{R_{p}} + (C-1)\frac{\tau_{30}}{\tau_{30} + \tau_{32}}(\tau_{10} + \tau_{21} + \tau_{32})\right)^{-1}$$
(S4)

For C<1, the denominator of the right side of Eq. (S4) can become negative, which is an unphysical solution. This indicates that the same population inversion  $\Delta N$  cannot be achieved with the new gain material. In our examples, this occurs for the cases with C = 0.15 and C = 0.06, which are the two cases shown to have inadequate gain to reach lasing.

### 2. FDTD self-consistent calculations

The gain material is characterized by the lifetimes  $\tau_{32} = 0.05$  ps,  $\tau_{21} = 80$  ps and  $\tau_{10} = 0.05$  ps and the coupling constant is  $\sigma_{\alpha} = 10^{-4} \text{ C}^2/\text{kg}$  ( $\tau_{30}$  is assumed for simplicity to be very large, i.e. the nonradiative  $N_3 \rightarrow N_0$  transition to be absent). The gain medium is assumed to have a Lorentzian response which is homogeneously broadened with linewidth  $\Gamma_{\alpha} = 2\pi \times 20 \times 10^{12}$  rad/s and emission frequency  $\omega_{\alpha} = 2\pi \times 200 \times 10^{12}$  rad/s. To avoid effects from frequency mismatch between the gain emission and the SRR resonance, we tune  $\omega_{\alpha}$  to coincide with the resonant frequency of the SRR-gain composite system. For example, for the strongly coupled case ( $\delta z = 0$  nm) we set  $\omega_{\alpha} = 2\pi \times 184 \times 10^{12}$  rad/s and for the uncoupled case ( $\delta z = 80$  nm) we set  $\omega_{\alpha} = 2\pi \times 199 \times 10^{12}$  rad/s. Similarly, we repeat this tuning for each  $\delta z$  separation.

In our simulations the total electron density is considered to be  $N_0(t=0) = N_0(t) + N_1(t) + N_2(t) + N_3(t) = C \times 5 \times 10^{23} \text{ m}^{-3}$  and the initial condition is that all electrons are in the ground state and all electric, magnetic and polarization fields are zero. For the major part of the paper we have used C = 1, while for the simulations presented in section "*Observable regimes for different gain materials*", we have set C to 10, 0.5, 0.15 and 0.06 to account for gain materials that provide different amounts of maximum gain. For the pump-probe simulations we first pump the system with a pulsed pump and then probe it with a weak Gaussian pulse and repeat for different pump-probe delays (for the CW pump simulations the pump-probe delay is irrelevant, because the gain is constant). For the lasing simulations (Fig. 9) we insert noise in the system, then pump with a CW pump and do not send a probe pulse, but monitor the outgoing waves instead and this procedure is repeated for several pump rates. In all cases, the system of the Maxwell equations coupled with the atomic rate equations is self-consistently solved in a Finite-Difference Time-Domain (FDTD) scheme<sup>S3</sup> using an approach similar to the one outlined in<sup>15</sup>.

### 3. Retrieval of effective parameters

The retrieval of effective metamaterial parameters is a very popular technique<sup>30,31</sup>, according to which the effective refractive index *n* and impedance *z* are calculated from the reflection and transmission coefficients, *r* and *t* respectively. To do so, it is assumed that the metamaterial is subwavelength enough to be homogenizable and it is then replaced by a homogeneous slab of thickness *d* with the respective properties *n* and *z* that need to be calculated. Then, the effective permittivity and permeability of the metamaterial are simply expressed as  $\varepsilon_r = n/z$  and  $\mu_r = nz$ . The inversion of *r* and *t*, however, introduces ambiguity both in the sign of *z* and the value of *n*. The sign of *z* can be determined by causality arguments, but lifting the ambiguity in *n* is a more demanding task, as multi-valued functions are involved.

The expression for *n* can be written as  $n(\omega)\omega = f(\omega)$  where  $f(\omega) = \frac{c}{d}(-j\ln|g(\omega)| + Arg[g(\omega)] + 2\pi q)$ ,  $q \in \mathbb{Z}$ ,

 $g(\omega) = \frac{1}{2t} \left(1 - r^2 + t^2\right) - \frac{2r}{t} \left(z - \frac{1}{z}\right)^{-1}$  and Arg[g] is the principal argument of g defined to lie in the interval  $(-\pi,\pi]$ . It is evident that while  $g(\omega)$  is a single-valued function of  $\omega$  that is directly calculated from r and t,  $f(\omega)$  (and hence n) consists of a multi-

valued real part and a single-valued function of  $\omega$  that is directly calculated from *i* and *i*,  $f(\omega)$  (and hence *n*) consists of a multi-valued real part and a single-valued imaginary part and is therefore a multi-valued function of  $\omega$ . The ambiguity in  $\text{Im}[f(\omega)]$  can be lifted again by causality arguments (e.g. if the material is passive, then  $\text{Im}[f(\omega)] \ge 0$ ) and hence, the only ambiguity left is with the multiple branches of  $\text{Re}[f(\omega)]$ , i.e. the multiple branches of  $\text{Re}[n(\omega)]$ .

It is crucial to note that  $Arg[g(\omega)]$  is wrapped in the interval  $(-\pi,\pi]$  by definition. This renders  $Re[n(\omega)]$  a piecewise continuous function of  $\omega$  for each q, i.e. a family of piecewise continuous functions of  $\omega$ . From this point on, the correct solution of  $Re[n(\omega)]$  can be obtained via two approaches:

1) The solution can be built piece by piece by choosing the correct part of each branch $^{30,33}$ .

2)  $Arg[g(\omega)]$  can be unwrapped, thus rendering  $Re[n(\omega)]$  a family of continuous curves. Then, the solution is obtained by identifying the appropriate continuous curve, i.e. the appropriate  $q^{31}$ .

With our modified retrieval technique, we manage to lift this ambiguity. Differentiation of  $f(\omega)$  gives:

$$n(\omega)\omega = f(\omega) \Rightarrow \frac{\partial n(\omega)}{\partial \omega}\omega + n(\omega) = \frac{\partial f(\omega)}{\partial \omega}$$
(S5),

where

$$\frac{\partial f(\omega)}{\partial \omega} = \frac{c}{d} \left( -j \frac{1}{\ln |g(\omega)|} \frac{\partial \ln |g(\omega)|}{\partial \omega} + \frac{\partial Arg[g(\omega)]}{\partial \omega} \right) \quad (S6)$$

It is obvious that not only is  $2\pi q$  now removed, but also the unknown function  $f(\omega)$  has given its place to the derivative  $\frac{\partial f(\omega)}{\partial \omega}$ , which can be calculated directly from  $g(\omega)$  (see Eq. (S6)). Consequently, the solution of Eq. (S5) gives the exact value of  $n(\omega)$ .

To cast the solution in numerical form, let us assume a set of N+1 frequency points  $\{\omega_i\}$  (i = 0, 1, 2, ..., N), i.e.  $\omega = \{\omega_0, \omega_1, \omega_2, ..., \omega_N\}$ , at which the reflection and transmission coefficients are measured. At these frequency points  $r = \{r_0, r_1, r_2, ..., r_N\}$  and  $t = \{t_0, t_1, t_2, ..., t_N\}$ , where  $r_i = r(\omega_i)$  and  $t_i = t(\omega_i)$ . All quantities involved in the calculations are also discretized and consequently  $g = \{g_0, g_1, g_2, ..., g_N\}$ ,  $f = \{f_0, f_1, f_2, ..., f_N\}$  and  $n = \{n_0, n_1, n_2, ..., n_N\}$ , where  $g_i = g(\omega_i)$ ,  $f_i = f(\omega_i)$  and  $n_i = n(\omega_i)$ . At a frequency point i = m (0 < m < N), the discretized impedance is expressed as:

$$z_m = \pm \sqrt{\frac{\left(1 + r_m\right)^2 - t_m^2}{\left(1 - r_m\right)^2 - t_m^2}}$$
(S7)

To calculate the discretized n, equation Eq. (S5) is expressed as:

$$\frac{n_{i+1} - n_i}{\omega_{i+1} - \omega_i} \frac{\omega_{i+1} + \omega_i}{2} + \frac{n_{i+1} + n_i}{2} = \frac{f_{i+1} - f_i}{\omega_{i+1} - \omega_i} \Longrightarrow$$
(S8)
$$\Rightarrow \begin{cases} n_{i+1} = (f_{i+1} - f_i + n_i \omega_i) / \omega_{i+1} \\ or \\ n_i = (f_i - f_{i+1} + n_{i+1} \omega_{i+1}) / \omega_i \end{cases}$$
(S9b)

The recursive relations Eq. (S9a) and Eq. (S9b) are equivalent and express the new *n* in terms of the old, either with increasing frequency Eq. (S9a) or decreasing Eq. (S9b). All  $f_i$  and  $\omega_i$  are known, hence Eq. (S9) is easily solved if recursion starts at a frequency  $\omega_{i,start}$  where *n* is known. The problem of calculating *n* is hence transferred to the correct determination of a starting value  $n_{i,start}$ .



Figure S3. Schematic of numerical implementation of the retrieval algorithm

With a simple inspection of Eq. (S5), it is obvious that if the derivative on the left hand side vanishes, i.e. *if the refractive index is constant around a certain frequency range*, this value can be directly calculated as  $n(\omega) = \frac{\partial f(\omega)}{\partial \omega}$ . In this case we can use Eq. (S9a) or Eq. (S9b) with

$$n_{i,start} = \frac{f_{i,start+1} - f_{i,start}}{\omega_{i,start+1} - \omega_{i,start}}$$
(S10)

From Eq. (S5) it is evident that if  $n(\omega) = C$ , then  $\frac{f(\omega)}{\omega} = C$  and consequently  $\frac{\partial \left(\frac{f(\omega)}{\omega}\right)}{\partial \omega} = 0$ , where C is a constant. It is straightforward to show that if  $\frac{f(\omega)}{\omega} = C$  then also  $\frac{\partial f(\omega)}{\partial \omega} = C$ , which is a weaker condition than the former (necessary but not

sufficient).

To conclude,  $n_{i,start}$  is unambiguously calculated directly from the data in hand in frequency areas that satisfy

 $\left(\frac{f(\omega)}{\omega}\right) = 0$ . One could claim that these conditions are rather strict. Fortunately, most materials of interest exhibit a number

of resonances, which die out at distant frequencies. There, these conditions are satisfied.

## References

S1. Siegman AE. Lasers, chapters 2, 3, 6, and 13. Sausalito: University Science, 1986.

S2. Yariv A and Yeh P. Optical electronics in modern communications. New York: Oxford University Press, 2007.

S3. Taflove A. Computational electrodynamics: the finite difference time domain method, chapters 3, 6, and 7. London: Artech House, 1995.