The spaser as a nanoscale quantum generator and ultrafast amplifier

Mark I Stockman[1,2,3]

[1] Department of Physics and Astronomy, Georgia State University, Atlanta, GA 30303, USA
[2] Max Planck Institute for Quantum Optics, Hans-Kopfermann-Strasse 1, 85748 Garching, Germany
[3] Ludwig Maximilian University Munich, Am Coulombwall 1, 85748 Garching, Germany[4]

E-mail: mstockman@gsu.edu

Abstract
Nanoplasmonics has recently experienced explosive development with many novel ideas and dramatic achievements in both fundamentals and applications. The spaser has been predicted and observed experimentally as an active element—a generator of coherent local fields. Even greater progress will be achieved if the spaser can function as an ultrafast nanoamplifier—an optical counterpart of the MOSFET (metal–oxide–semiconductor field effect transistor). A formidable problem with this is that the spaser has inherent feedback, causing quantum generation of nanolocalized surface plasmons and saturation and consequent elimination of the net gain, making it unsuitable for amplification. We have overcome this inherent problem and shown that the spaser can perform functions of an ultrafast nanoamplifier in two modes: transient and bistable. On the basis of quantum density matrix (optical Bloch) equations we have shown that the spaser amplifies with gain $\gtrsim 50$ with a switching time $\lesssim 100$ fs (potentially, ~ 10 fs). This prospective spaser technology will further broaden both fundamental and applied horizons of nanoscience, in particular enabling ultrafast microprocessors working at 10–100 THz clock speed. Other prospective applications are in ultrasensing, ultradense and ultrafast information storage, and biomedicine. The spasers are based on metals and, in contrast to semiconductors, are highly resistive to ionizing radiation, high temperatures, microwave radiation, and other adverse environments.

1. Introduction

Nanoplasmonic phenomena (see, e.g., [1]) unfold on the spatial scale between the skin depth $l_s \approx 25$ nm (in noble metals) and the nonlocality radius $l_{nl} \sim v_F/\omega \sim 1$ nm, where v_F is the electron speed at the Fermi surface, and ω is the optical frequency. Nanoplasmonics is ultrafast: the temporal scale of the nanoplasmonic phenomena is between the coherence time of a hundred attoseconds defined by the inverse bandwidth of the plasmonic frequency range (between UV and mid-IR for plasmonic metals and doped semiconductors) [2, 3] and the surface plasmon (SP) relaxation time $\gamma_p^{-1} \sim 10$–100 fs (for noble metals in the visible to near-IR frequency range) [4].

Not just a promise any longer [5], nanoplasmonics has delivered a number of important applications: ultrasensing [6], scanning near-field optical microscopy [1, 7], SP-enhanced photodetectors [8], thermally assisted magnetic recording [9], generation of extreme UV [10], biomedical tests [6, 11], SP-assisted thermal cancer treatment [12], and many others.

To continue its vigorous development, nanoplasmonics needs an active device—a near-field generator and amplifier of nanolocalized optical fields, which has until recently been absent. A nanoscale amplifier in microelectronics is the metal–oxide–semiconductor field effect transistor (MOSFET) [13, 14], which has enabled all contemporary digital electronics, including computers and communications, and formed the present day technology as we know it.

---

[4] http://www.phy-astr.gsu.edu/stockman.

However, the MOSFET is limited by frequency and bandwidth to ≲100 GHz, which is already a limiting factor in further technological development. Another limitation of the MOSFET is its high sensitivity to temperature, electric fields, and ionizing radiation, which limits its use in extreme environmental conditions.

An active element of nanoplasmonics is the spaser (surface plasmon amplification by stimulated emission of radiation) that was proposed [4] (see also [15]) as a nanoscale quantum generator of nanolocalized coherent and intense optical fields. It has recently been observed experimentally [16]. However, there is a formidable problem for setting the spaser as a nanoscale quantum amplifier. A principal difference between a laser (quantum optical oscillator) and a quantum optical amplifier is that the amplifier does not possess feedback: it does not have mirrors and any parasitic feedback due to light scattering should be carefully minimized to prevent spontaneous generation. In sharp contrast, in the spaser, as we discuss below in this introduction, the feedback is always inherently present because the metal plasmonic nanoparticle (spaser's core) supports SP modes whose fields exert periodic perturbation on the gain medium causing the feedback, which *fundamentally* cannot be removed. This feedback will cause the SP generation and the subsequent saturation of the gain. For any spaser (or any laser, for that matter) in a stationary (CW) regime, the net amplification must be zero (which is a condition of the stable CW operation). Therefore, one might assume that the spaser cannot serve as a nanoamplifier.

In this article we have solved this problem by suggesting two approaches to setting the spaser as a nanoamplifier: (i) the dynamic or transient approach, which is based on the fact that during a femtosecond transient process after the population inversion is created but before the CW regime is established, the spaser possesses a net amplification, and (ii) the bistability approach that is based on the inclusion of a saturable absorber in the gain medium, which prevents spontaneous generation and sets the spaser as a bistable (logical) ultrafast nanoamplifier. This is done by employing a quantum kinetic theory that is based on density matrix equations, which are the adaptation of the optical Bloch equations of the laser theory to the case of the spaser. From these equations we have also obtained new and potentially useful results describing not only the ultrafast dynamics of the spaser but also its CW mode: the 'spasing curve' (that is a relation between the SP population of the spasing mode and the pumping rate) and also the linewidth of the spaser. These results provide an excellent qualitative explanation of the recent experimental data [16].

On the basis of the present theory, one may envisage femtosecond-cycle nanoplasmonic chips with a high degree of integration where spasers communicate and control each other through their local optical fields or are connected with nanoplasmonic wires. These can perform ultrafast microprocessor functions. The spasers can also be integrated with nanophotodetectors and nanosensors to perform complex functions of intelligent ultrafast detection and sensing. In contrast to semiconductor technology, the spasers are based on metals and, therefore, are highly resistive to ionizing radiation, high temperatures and other adverse environments, with possible applications in space, nuclear industries and defense. These functions of spasers will further widen both the fundamental and applied horizons of nanoplasmonics and, generally, science and technology.

1.1. Spaser fundamentals

We have recently discussed the physical principles of the spasing [15] but will briefly reiterate them here for the sake of completeness and self-containment. The spaser is a nanoplasmonic counterpart of the laser [4, 15]. The laser has two principal elements: the resonator (or cavity) that supports photonic mode(s) and the gain (or active) medium that is population inverted and supplies energy to the lasing mode(s). An inherent limitation of the laser is that the size of the laser cavity in the propagation direction is at least half a wavelength and practically more than that even for the smallest lasers developed [17–19]. In the spaser [4] this limitation is overcome. The spasing modes are surface plasmons (SPs) whose localization length is on the nanoscale [20] and is only limited by the minimum inhomogeneity scale of the plasmonic metal and the nonlocality radius [21] $l_{nl} \sim 1$ nm. So, the spaser is truly nanoscopic; its minimum total size can be just a few nm.

The resonator of a spaser can be any plasmonic metal nanoparticle whose total size R is much less than the wavelength λ and whose metal thickness is between l_{nl} and l_s, which supports a SP mode with the required frequency ω_n. This metal nanoparticle should be surrounded by the gain medium that overlaps with the spasing SP eigenmode spatially and whose emission line overlaps with this eigenmode spectrally [4]. As an example, we consider a model of a nanoshell spaser [15, 22], which is illustrated in figure 1. Panel (a) shows a silver nanoshell carrying a single SP (plasmon population number $N_n = 1$) in the dipole eigenmode. It is characterized by a uniform field inside the core and hot spots at the poles outside the shell with the maximum field reaching ∼10^6 V cm^{-1}. Similarly, figure 1(b) shows the quadrupole mode in the same nanoshell. In this case, the mode electric field is non-uniform, exhibiting hot spots of ∼1.5×10^6 V cm^{-1} of the modal electric field at the poles. These high values of the modal fields are the underlying physical reason for the very strong feedback in the spaser. Under our conditions, the electromagnetic retardation within the spaser volume can be safely neglected. Also, the radiation of such a spaser is a weak effect: the decay rate of plasmonic eigenmodes is dominated by the internal loss in the metal. Therefore, it is sufficient to consider only quasistatic eigenmodes [20, 23] and not their full electrodynamic counterparts [24], which also is necessary for the ultrafast operation of the spaser as a nanoamplifier.

Note that for high aspect ratios (say, $\eta = 0.95$, considered in some of the examples) and relatively small radii ($R = 12$ nm) the thickness of the metal shell may be rather small (<1 nm), which may lead to nonlocal effects including Landau damping. In this case, the radius of the
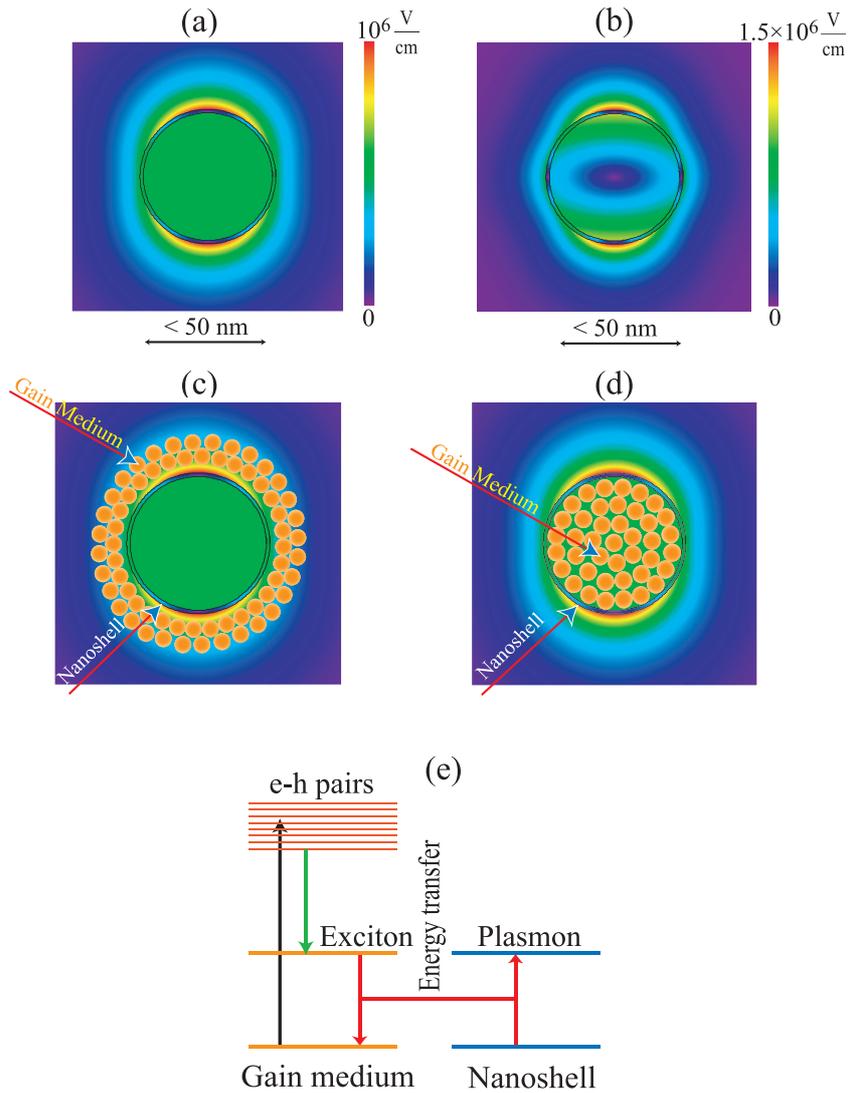
Figure 1. Schematic of the spaser geometry, local fields, and fundamental processes leading to spasing. (a) Nanoshell geometry and the local optical field distribution for one SP in an axially symmetric dipole mode. The nanoshell has aspect ratio $\eta = 0.95$. The local field magnitude is color coded by the scale bar in the right-hand side of the panel. (b) The same as (a) but for a quadrupole mode. (c) Schematic of a nanoshell spaser where the gain medium is outside of the shell, on the background of the dipole mode field. (d) The same as (c) but for the gain medium inside the shell. (e) Schematic of the spasing process. The gain medium is excited and population inverted by an external source, as depicted by the black arrow, which produces electron–hole pairs in it. These pairs relax, as shown by the green arrow, to form the excitons. The excitons recombine, emitting SPs in the nanoshell. The plasmonic oscillations of the nanoshell stimulate this emission, supplying the feedback for the spaser action.

spaser can be safely increased by a factor of 2–3 (this will not adversely effect the spasing as the spaser is fully scalable, as we show below in section 2.2). Also one should keep in mind that the nonlocal effects in thin nanoshells are small for special symmetry reasons [25]. We also consider a number of examples with significantly higher aspect ratios where the metal thickness is a few nanometers or larger.

For the sake of numerical illustration of our theory, we will use the dipole eigenmode (figure 1(a)). There are two basic ways to place the gain medium: (i) outside the nanoshell, as shown in panel (c), and (ii) in the core, as in panel (d), which was originally proposed in [22]. As we have verified, these two designs lead to comparable characteristics of the spaser. However, the placement of the gain medium inside the core

illustrated in figure 1(d) has a significant advantage, because the hot spots of the local field are not covered by the gain medium and are sterically available for applications.

Note that any l-multipole mode of a spherical particle is, indeed, $2l + 1$-times degenerate. This may make the spasing mode polarization unstable, like in lasers without polarizing elements. In reality, the polarization may be pinned and become stable due to deviations from the perfect spherical symmetry, which exist naturally or can be introduced deliberately. A more practical shape for a spaser may be a nanorod, which has a mode with the stable polarization along the major axis. However, a nanorod is a more complicated geometry for theoretical treatment, and we will consider it elsewhere.

The level diagram of the spaser gain medium and the plasmonic metal nanoparticle is displayed in figure 1(e) along with a schematic of the relevant energy transitions in the system. The gain medium chromophores may be semiconductor nanocrystals [4, 26], dye molecules [27, 28], rare-earth ions [22], or electron–hole excitations of a bulk semiconductor [17, 18]. For certainty, we will use the semiconductor science language of electrons and holes.

The pump excites electron–hole pairs in the chromophores (figure 1(e)), as indicated by the vertical black arrow, which relax to form excitons. The excitons constitute the two-level systems that are the donors of energy for the SP emission into the spasing mode. In vacuum, the excitons would recombine emitting photons. However, in the spaser geometry, the photoemission is strongly quenched due to the resonance energy transfer to the SP modes, as indicated by the red arrows in the panel. The plasmons in the spaser mode create the high local fields that excite the gain medium and stimulate more emission to this mode, which is the feedback mechanism. If this feedback is strong enough and the lifetime of the spaser SP mode is long enough, then an instability develops leading to the avalanche of the SP emission in the spasing mode and spontaneous symmetry breaking, establishing the phase coherence of the spasing state. Thus the establishing of spasing can be called a nonequilibrium phase transition, as in the physics of lasing.

1.2. Brief overview of the latest progress in spasers and nanolasers

Since the original theoretical proposal and prediction of the spaser [4], there has been active development in this field, both theoretical and experimental. We comment below only on some representative publications. Among theoretical developments, a nanolens spaser has been proposed [29], which possesses a nanofocus ('the hottest spot') of the local fields. In [4, 29], only the necessary condition of spasing has been established on the basis of the perturbation theory.

There have been theories published describing the spaser (or, 'nanolaser' as it is sometimes called) phenomenologically, on the basis of classic linear electrodynamics by considering the gain medium as a dielectric with a negative imaginary part of the permittivity [22, 30–32]. Such electrodynamic approaches do not take into account the nature of the spasing as a spontaneous symmetry breaking. This leads to principal differences of their results from the present microscopic quantum mechanical theory in the region of spasing, as we discuss below in section 3.1 in conjunction with figure 2. Note that before the spasing threshold, the above-mentioned phenomenological theories are, in principle, applicable. There has also been a theoretical publication on a bowtie spaser (nanolaser) with electrical pumping [33]. It is based on balance equations and only the CW spasing generation intensity is described. Yet another theoretical development has been a proposal of the lasing spaser [34], which is made of a planar array of spasers. The theoretical publications mentioned above in this paragraph deal with the CW regime. In contrast, in this article we are most interested in the ultrafast behavior of the spaser.

There has also been a theoretical proposal of a spaser ('nanolaser') consisting of a metal nanoparticle coupled to a single chromophore [35]. In this paper, a dipole–dipole interaction is illegitimately used at very small distances r where it has a singularity (diverging for $r \rightarrow 0$), leading to a dramatically overestimated coupling with the SP mode. As a result, a completely unphysical prediction of CW spasing due to a single chromophore has been obtained [35]. In contrast, our theory is based on the full (exact) field of the spasing SP mode. As our results of section 3 below show, hundreds of chromophores per metal nanoparticle are realistically required for the spasing even under the most favorable conditions.

There has been a vigorous experimental quest toward the spaser. Stimulated emission of SPPs has been observed in a proof-of-principle experiment using pumped dye molecules as an active (gain) medium [27]. There have also been later experiments that demonstrated strong stimulated emission compensating a significant part of the SPP loss [28, 36–38]. As a step toward the lasing spaser, the first experimental demonstration of compensating Joule losses in a metallic photonic metamaterial using optically pumped PbS semiconductor quantum dots has been reported [26]. In another development, a tunable free-electron light nanosource ('light well') has been demonstrated [39], which is based on spontaneous emission in a periodically layered metal–dielectric structure. There have also been experimental investigations reporting the stimulated emission effects of SPs in plasmonic metal nanoparticles surrounded by gain media with dye molecules [40, 41]. An electrically pumped nanolaser with a semiconductor gain medium has been demonstrated [18] where the lasing modes are SPPs. A nanolaser with an optically pumped semiconductor gain medium and a hybrid semiconductor silver SPP waveguide has been demonstrated with an extremely tight transverse mode confinement [19].

Finally, an observation has been reported of a true spaser, by Noginov *et al* [16]. This spaser is a chemically synthesized gold nanosphere of radius 7 nm surrounded by a dielectric shell of 21 nm outer radius containing immobilized dye molecules, which is sometimes referred to as a 'Cornell nanodot'. Under nanosecond optical pumping in the absorption band of the dye, this spaser develops a relatively narrow spectrum and intense visible emission that exhibits a pronounced threshold in pumping intensity. The observed characteristics of this spaser are in excellent qualitative agreement and can be fully understood on the basis of the corresponding theoretical results obtained below in section 3.1.

This article is organized as follows. In section 2.1 for the case of the spaser, we formulate quantum density matrix equations (also called in the literature the optical Bloch equations). In section 2.2 we consider these equations for a stationary (CW) case where we find the conditions of spasing, the population of the SPs in the spasing mode, and the spasing line shape. In section 3.1 we illustrate the spasing curve and spectral composition of the spaser generation. In section 3.1 we describe the CW spaser as a bistable nanodevice. The major results of this article regarding the ultrafast kinetics of the spaser as a generator and quantum nanoamplifier are presented in section 4. We briefly discuss the results obtained and conclude in section 5.

2. Equations for the spaser

2.1. Quantum density matrix (optical Bloch) equations for the spaser

The SP eigenmodes $\varphi_n(\mathbf{r})$ are described by a wave equation (with homogeneous boundary conditions) [4, 20]

$$\nabla \Theta(\mathbf{r}) \nabla \varphi_n(\mathbf{r}) = s_n \nabla^2 \varphi_n(\mathbf{r}), \tag{1}$$

where n is the mode number, s_n is corresponding eigenvalue, and $\Theta(\mathbf{r})$ is the characteristic function equal to 1 for $\mathbf{r}$ in the metal component and 0 for $\mathbf{r}$ in the dielectric. Note that the eigenvalues s_n are all real and contained in the range $1 \geqslant s_n \geqslant 0$. The eigenmodes are normalized by an integral over the volume V of the system, $\int_V |\nabla \varphi_n(\mathbf{r})|^2 \, \mathrm{d}^3 r = 1$. The physical frequency ω_n of the SPs is defined by an equation $\mathrm{Re}[s(\omega_n)] = s_n$, where $s(\omega) = \varepsilon_d/[\varepsilon_d - \varepsilon_m(\omega)]$ is Bergman's spectral parameter, ε_d is the permittivity of the ambient dielectric, and $\varepsilon_m(\omega)$ is the metal permittivity.

The electric field operator[5] of the quantized SPs is [4]

$$\mathbf{E}(\mathbf{r}) = -\sum_n A_n \nabla \varphi_n(\mathbf{r})(\hat{a}_n + \hat{a}_n^\dagger), \qquad A_n = \left(\frac{4\pi \hbar s_n}{\varepsilon_d s_n'} \right)^{1/2}, \tag{2}$$

where $\hat{a}_n^\dagger$ and $\hat{a}_n$ are the SP creation and annihilation operators, and $s_n' = \mathrm{Re}[\mathrm{d}s(\omega_n)/\mathrm{d}\omega_n]$.

The spaser Hamiltonian has the form

$$H = H_g + \hbar \sum_n \omega_n \hat{a}_n^\dagger \hat{a}_n - \sum_p \mathbf{E}(\mathbf{r}_p) \mathbf{d}^{(p)}, \tag{3}$$

where H_g is the Hamiltonian of the gain medium, p is an index (label) of a gain medium chromophore, $\mathbf{r}_p$ is its coordinate vector, and $\mathbf{d}^{(p)}$ is its dipole moment operator. In this paper, we will treat the active medium quantum mechanically but the SPs quasiclassically, considering $\hat{a}_n$ as a classical quantity (c-number) a_n with time dependence given by $a_n = a_{0n} \exp(-\mathrm{i}\omega t)$, where a_{0n} is a slowly varying amplitude. The number of coherent SPs per spasing mode is then given by $N_p = |a_{0n}|^2$. This approximation neglects the quantum fluctuations of the SP amplitudes. However, when necessary, we will take into account these quantum fluctuations, in particular, to describe the spectrum of the spaser.

Introducing $\rho^{(p)}$ as the density matrix of a pth chromophore, we can find its equation of motion in a conventional way by commutating it with the Hamiltonian (3) as $\mathrm{i}\hbar \dot{\rho}^{(p)} = [\rho^{(p)}, H]$, where the dot denotes the temporal derivative. We will use the standard rotating wave approximation (RWA), which only takes into account the resonant interaction between the optical field and chromophores. We denote as $|1\rangle$ and $|2\rangle$ the ground and excited states of a chromophore, with the transition $|2\rangle \rightleftharpoons |1\rangle$ resonant to the spasing plasmon mode n. In this approximation, the time dependence of the non-diagonal elements of the density matrix is $(\rho^{(p)})_{12} = \bar{\rho}_{12}^{(p)} \exp(\mathrm{i}\omega t)$, and $(\rho^{(p)})_{21} = \bar{\rho}_{12}^{(p)*} \exp(-\mathrm{i}\omega t)$, where $\bar{\rho}_{12}^{(p)}$ is a time-independent amplitude

---

[5] Note that we have corrected a misprint in [4], replacing the coefficient 2π by 4π.

defining the coherence (polarization) for the $|2\rangle \rightleftharpoons |1\rangle$ spasing transition in a pth chromophore of the gain medium.

Introducing a rate constant Γ_{12} to describe the polarization relaxation and a difference $n_{21}^{(p)} = \rho_{22}^{(p)} - \rho_{11}^{(p)}$ as the population inversion on this spasing transition, we derive an equation of motion for the non-diagonal element of the density matrix as

$$\dot{\bar{\rho}}_{12}^{(p)} = -[\mathrm{i}(\omega - \omega_{12}) + \Gamma_{12}]\bar{\rho}_{12}^{(p)} + \mathrm{i}n_{21}^{(p)}\Omega_{12}^{(p)*}, \tag{4}$$

where $\Omega_{12}^{(p)} = -A_n \mathbf{d}_{12}^{(p)} \nabla \varphi_n(\mathbf{r}_p) a_{0n}/\hbar$ is the Rabi frequency for the spasing transition in a pth chromophore, and $\mathbf{d}_{12}^{(p)}$ is the corresponding transitional dipole element. Note that always $\mathbf{d}_{12}^{(p)}$ is either real or can be made real by a proper choice of the quantum state phases, making the Rabi frequency $\Omega_{12}^{(p)}$ also a real quantity.

An equation of motion for n_{21}^p can be found in a standard way by commutation with H. To provide conditions for the population inversion ($n_{21}^p > 0$), we imply existence of a third level. For simplicity, we assume that it very rapidly decays into the excited state $|2\rangle$ of the chromophore, so its own population is negligible. It is pumped by an external source from the ground state (optically or electrically) with some rate that we will denote as g. In this way, we obtain the following equation of motion:

$$\dot{n}_{21}^{(p)} = -4\,\mathrm{Im}[\bar{\rho}_{12}^{(p)}\Omega_{21}^{(p)}] - \gamma_2(1 + n_{21}^{(p)}) + g(1 - n_{21}^{(p)}), \tag{5}$$

where γ_2 is the decay rate $|2\rangle \rightarrow |1\rangle$.

The stimulated emission of the SPs is described as their excitation by the coherent polarization of the gain medium. The corresponding equation of motion can be obtained using Hamiltonian (3) and adding the SP relaxation with a rate of γ_n as

$$\dot{a}_{0n} = [\mathrm{i}(\omega - \omega_n) - \gamma_n]a_{0n} + \mathrm{i}\sum_p \rho_{12}^{(p)*}\Omega_{12}^{(p)}. \tag{6}$$

Another relevant process is spontaneous emission of SPs by a chromophore into a spasing SP mode. The corresponding rate $\gamma_2^{(p)}$ for a chromophore at a point $\mathbf{r}_p$ can be found in a standard way using the quantized field (2) as

$$\gamma_2^{(p)} = 2\frac{A_n^2}{\hbar \gamma_n}|\mathbf{d}_{12}\nabla \varphi_n(\mathbf{r}_p)|^2 \frac{(\Gamma_{12} + \gamma_n)^2}{(\omega_{12} - \omega_n)^2 + (\Gamma_{12} + \gamma_n)^2}. \tag{7}$$

As in the Schawlow–Townes theory of laser linewidth, this spontaneous emission of SPs leads to the diffusion of the phase of the spasing state. This defines the width γ_s of the spasing line as

$$\gamma_s = \frac{\sum_p (1 + n_{21}^{(p)})\gamma_2^{(p)}}{2(2N_p + 1)}. \tag{8}$$

This width is small for a case of developed spasing when $N_p \gg 1$. However, for $N_p \sim 1$, the predicted width may be too high because the spectral diffusion theory assumes that $\gamma_s \lesssim \gamma_n$. To take into account this limitation in a simple way, we will interpolate to find the resulting spectral width Γ_s of the spasing line as $\Gamma_s = (\gamma_n^{-2} + \gamma_s^{-2})^{-1/2}$.

We will also examine the spaser as a bistable (logical) amplifier. One of the ways to set the spaser in such a mode is to add a saturable absorber. This is described by the same

equations (4)–(6) where the chromophores belonging to the absorber are not pumped by the external source directly, i.e., for them in equation (5) one has to set $g = 0$.

Numerical examples are given for a silver nanoshell where the core and the external dielectric have the same permittivity of $\varepsilon_d = 2$; the permittivity of silver is adopted from [42]. The following realistic parameters of the gain medium are used (unless indicated otherwise): $d_{12} = 1.5 \times 10^{-17}$ esu, $\hbar\Gamma_{12} = 10$ meV, $\gamma_2 = 4 \times 10^{12}$ s^{-1} (this value takes into account the spontaneous decay into SPs), and the density of the gain medium chromophores is $\rho = 2.4 \times 10^{20}$ cm^{-3}, which is realistic for dye molecules but may be somewhat high for the semiconductor quantum dots that were proposed as the chromophores [4] and used in experiments [26]. We will assume a dipole SP mode and chromophores situated in the core of the nanoshell as shown in figure 1(d). This configuration is of advantage both functionally (because the region of the high local fields outside the shell is accessible for various applications) and computationally (the uniformity of the modal fields makes the summation of the chromophores trivial, thus greatly facilitating numerical procedures).

2.2. Equations for the stationary (CW) regime

Physically, the spaser action is a result of spontaneous symmetry breaking when the phase of the coherent SP field is established from the spontaneous noise. Mathematically, the spaser is described by homogeneous differential equations (4)–(6). These equations become homogeneous algebraic equations for the stationary (or continuous wave (CW)) case. They always have a trivial, zero solution. However, when their determinant vanishes, they also possess a nontrivial solution describing spasing, whose condition is

$$(\omega_s - \omega_n + i\gamma_n)^{-1}(\omega_s - \omega_{21} + i\Gamma_{12})^{-1} \sum_p |\tilde{\Omega}_{12}^{(p)}|^2 n_{21}^{(p)} = -1,$$
(9)

where $\tilde{\Omega}_{12}^{(p)} = -A_n \mathbf{d}_{12}^{(p)} \nabla\varphi_n(\mathbf{r}_p)/\hbar$ is the single-plasmon Rabi frequency. The population inversion of a pth chromophore $n_{21}^{(p)}$ is explicitly expressed as

$$n_{21}^{(p)} = (g - \gamma_2)$$
$$\times \{g + \gamma_2 + 4N_n|\tilde{\Omega}_{12}^{(p)}|^2/[(\omega_s - \omega_{21})^2 + \Gamma_{12}^2]\}^{-1}. \quad (10)$$

From the imaginary part of equation (9) we immediately find the spasing frequency

$$\omega_s = (\gamma_n \omega_{21} + \Gamma_{12}\omega_n)/(\gamma_n + \Gamma_{12}), \quad (11)$$

which generally does not coincide with either the gain transition frequency ω_{21} or the SP frequency ω_n, but is between them (this is a frequency walk-off phenomenon similar to that of laser physics). Substituting equation (11) back into equations (9)–(10), we obtain a system of equations

$$\frac{(\gamma_n + \Gamma_{12})^2}{\gamma_n\Gamma_{12}[(\omega_{21} - \omega_n)^2 + (\Gamma_{12} + \gamma_n)^2]} \sum_p |\tilde{\Omega}_{12}^{(p)}|^2 n_{21}^{(p)} = 1, \quad (12)$$

$$n_{21}^{(p)} = (g - \gamma_2)\left[g + \gamma_2 + \frac{4N_n|\tilde{\Omega}_{12}^{(p)}|^2(\Gamma_{12} + \gamma_n)}{(\omega_{12} - \omega_n)^2 + (\Gamma_{12} + \gamma_n)^2}\right]^{-1}.$$
(13)

This system defines the stationary (CW) number of SPs per spasing mode N_n.

Since $n_{21}^{(p)} \leqslant 1$, from equations (12) and (13) we immediately obtain a necessary condition for the existence of spasing:

$$\frac{(\gamma_n + \Gamma_{12})^2}{\gamma_n\Gamma_{12}[(\omega_{21} - \omega_n)^2 + (\Gamma_{12} + \gamma_n)^2]} \sum_p |\tilde{\Omega}_{12}^{(p)}|^2 \geqslant 1. \quad (14)$$

This expression is fully consistent with [4]. The following order of magnitude estimate of this spasing condition has a transparent physical meaning and is of heuristic value:

$$\frac{d_{12}^2 Q N_c}{\hbar\Gamma_{12} V_n} \gtrsim 1, \quad (15)$$

where $Q = \omega/\gamma_n$ is the quality factor of SPs, V_n is the volume of the spasing SP mode, and N_c is the number of gain medium chromophores within this volume. Deriving this estimate, we have neglected the detuning, i.e., set $\omega_{21} - \omega_n = 0$. We also used the definitions of A_n of equation (2) and $\tilde{\Omega}_{12}^{(p)}$ given after equation (9), and the estimate $|\nabla\varphi_n(\mathbf{r})|^2 \sim 1/V$ following from the normalization of the SP eigenmodes $\int |\nabla\varphi_n(\mathbf{r})|^2 \, d^3r = 1$ of [20]. The result of equation (15) is, indeed, in agreement with [4] where it was obtained in slightly different notations.

It follows from equation (15) that for the existence of spasing it is beneficial to have a high quality factor Q, a high density of the chromophores, and a large transition dipole (oscillator strength) of the chromophore transition. The small modal volume V_n (at a given number of the chromophores N_c) is beneficial for this spasing condition: physically, it implies strong feedback in the spaser. Note that for the given density of the chromophores $\rho_c = N_c/V_n$, this spasing condition does not explicitly depend on the spaser size, which opens up the possibility of spasers of a very small size limited from the bottom by only the nonlocality radius $l_{nl} \sim 1$ nm. Another important property of equation (15) is that it implies a quantum mechanical nature of spasing and spaser amplification: this condition essentially contains the Planck constant $\hbar$ and, thus, does not have a classical counterpart. Note that in contrast to lasers, the spaser theory and equation (15) in particular do not contain the speed of light, i.e., they are quasistatic.

3. The spaser in the mode of a continuous wave nanoscale quantum generator

3.1. Kinetics of the CW spaser

The 'spasing curve', i.e., the dependence of the coherent SP population N_n on the excitation rate g, obtained by solving equations (12) and (13), is shown in figure 2(a) for four kinds of silver nanoshells with the frequencies of the spasing dipole modes as indicated, which are in the range from near-IR ($\hbar\omega_s = 1.2$ eV) to mid-visible ($\hbar\omega_s = 2.2$ eV). In all cases, there is a pronounced threshold of the spasing at an excitation rate $g_{th} \sim 10^{12}$ s^{-1}. Soon after the threshold, the dependence $N_n(g)$ becomes linear, which means that every quantum of excitation added to the active medium with a high probability
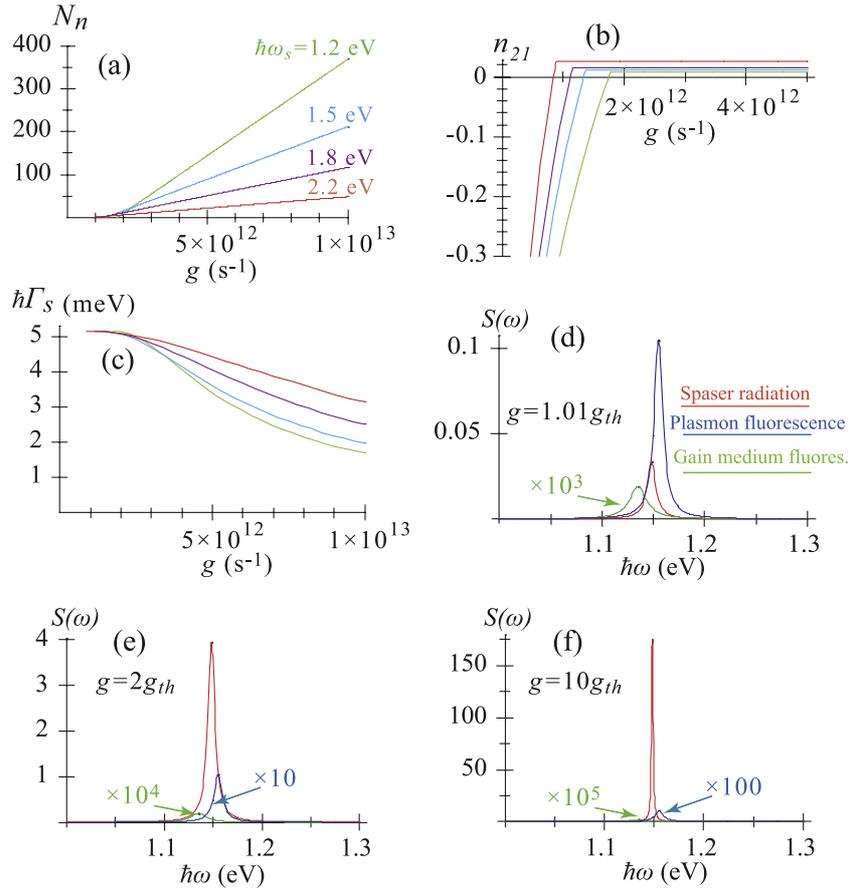
Figure 2. Spaser SP population and spectral characteristics in the stationary state. The computations are done for a silver nanoshell with the external radius $R_2 = 12$ nm; the detuning of the gain medium from the spasing SP mode is $\hbar(\omega_{21} - \omega_n) = -0.02$ eV. The other parameters are indicated in section 2. (a) Number N_n of plasmons per spasing mode as a function of the excitation rate g (per chromophore of the gain medium). Computations are done for the dipole eigenmode with the spasing frequencies ω_s as indicated, which were chosen by the corresponding adjustment of the nanoshell aspect ratio. (b) Population inversion n_{12} as a function of the pumping rate g. The color coding of the lines is the same as in panel (a). (c) The spectral width Γ_s of the spasing line (expressed as $\hbar\Gamma_s$ in meV) as a function of the pumping rate g. The color coding of the lines is the same as in panel (a). (d)–(f) Spectra of the spaser for the pumping rates g expressed in the units of the threshold rate g_{th}, as indicated in the panels. The curves are color coded and scaled as indicated.

is stimulated to be emitted as a SP, adding to the coherent SP population.

While this is similar to the case for conventional lasers, there is a dramatic difference for the spaser. For the lasers a similar relative rate of the stimulated emission is achieved at a photon population of $\sim 10^{18}$–10^{20}, while for the spaser the SP population is $N_n \lesssim 100$. This is due to the much stronger feedback in spasers because of the much smaller modal volume V_n—see the discussion of equation (15). The shape of the spasing curves of figure 2(a) (the well-pronounced threshold with the linear dependence almost immediately above the threshold) is in excellent qualitative agreement with the experiment [16]. Note that the recently demonstrated SPP nanolaser with a tight transverse confinement of the lasing modes does not exhibit a pronounced threshold [19], in contrast to the spaser.

The population inversion number n_{21} as a function of the excitation rate g is displayed in figure 2(b) for the same set of frequencies (and with the same color coding) as in panel (a). Before the spasing threshold, n_{21} increases with g to become positive with the onset of the population inversion just before

the spasing threshold. For higher g, after the spasing threshold is exceeded, the inversion n_{21} becomes constant (inversion pinning). The pinned levels of the inversion are very low, $n_{21} \sim 0.01$, which again is due to the very strong feedback in the spaser.

The spectral width Γ_s of the spaser generation is due to the phase diffusion of the quantum SP state caused by the noise of the spontaneous emission of the SPs into the spasing mode, as described by equation (8). This width is displayed in figure 2(c) as a function of the pumping rate g. At the threshold, Γ_s is that of the SP line γ_n, but for stronger pumping, as the SPs accumulate in the spasing mode, it decreases $\propto N_n^{-1}$, as given by equation (8). This decrease of Γ_s reflects the higher coherence of the spasing state with the increased number of SP quanta and, correspondingly, lower quantum fluctuations. This is similar to the lasers as described by the Schawlow–Townes theory [43].

The spasing developed in a dipole SP mode will show itself in the far field as an anomalously narrow and intense radiation line. The shape and intensity of this line in relation to the lines of the spontaneous fluorescence of the isolated
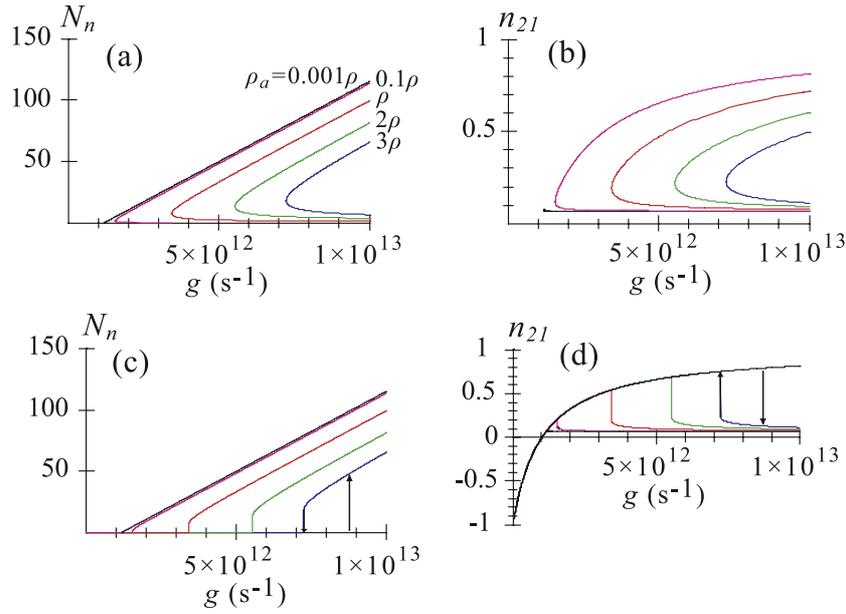
Figure 3. Bistability in a spaser with a saturable absorber in a stationary spasing state. (a) Dependence of the SP population number N_n in the spasing mode on the pumping rate g for different concentrations ρ_a of the saturable absorber. The curves are color coded corresponding to ρ_a shown in units of the concentration ρ of the active medium chromophores. The black line shows the threshold curve (separatrix) between the bistable and uniquely stable solutions. (b) The dependence of the population inversion n_{21} on the pumping rate g for the spasing states. The color coding is the same as in panel (a). Note that the lower branches of the curves in this panel correspond to the upper ones in panel (a) and vice versa. (c) The same as panel (a) but only the stable branches of the curves are displayed, illustrating the physical behavior of the spaser population N_n. Note that one of the stable branches coincides with the horizontal axis. The vertical arrows illustrate the hysteretic transitions between the two stable branches for $\rho_a = 3\rho$. (d) The same as panel (c) but for the population inversion n_{21}. The black curve starting with $n_{21} = -1$ shows a trivial stable solution corresponding to the absence of spasing ($N_n = 0$).

gain medium and its SP-enhanced fluorescence line in the spaser are illustrated in figures 2(d)–(f). Note that for the system under consideration, there is a 20 meV red shift of the gain medium fluorescence with respect to the SP line center. It is chosen so as to illustrate the spectral walk-off of the spaser line. For one per cent in the excitation rate above the threshold of the spasing (panel (d)), a broad spasing line (red color) appears, comparable in intensity to the SP-enhanced spontaneous fluorescence line (blue color). The width of this spasing line is approximately the same as that of the fluorescence, but its position is shifted appreciably (spectral walk-off) toward the isolated gain medium line (green color). For pumping twice more intense (panel (e)), the spaser line radiation dominates, but its width is still close to that of the SP line due to significant quantum fluctuations of the spasing state phase. Only when the pumping rate is an order of magnitude above the threshold does the spaser line strongly narrow (panel (f)), and it also completely dominates the spectrum of the radiation. This is a regime of small quantum fluctuations, which is desired in applications.

These results for the spasing region are different in the most dramatic way from those for previous phenomenological models, which are based on the consideration of a gain medium that has negative imaginary part of its permittivity plus a lossy metal nanosystem, described purely electrodynamically [22, 31]. For instance, in a 'toy model' [31], the width of the resonance line tends to zero at the threshold of spasing and then broadens again. This distinction of the present theory is due to the nature of the

spasing as a spontaneous symmetry breaking (nonequilibrium phase transition) leading to the establishing of the coherent SP state. The spasing state with its spontaneous symmetry breaking cannot be obtained in electrodynamics, which is a linear response theory, from the initial, non-spasing state of the system. The SP population in the spasing state is determined by nonlinear saturation of the gain. This state is influenced by the phase relaxation due to the spontaneous SP emission into the spasing mode leading to the finite spectral width of the spasing line $\Gamma_s \propto N_n^{-1}$. Therefore the present theory requires a quantum mechanical consideration of the gain medium. Note that below the spasing threshold, the phenomenological theories [22, 32] are applicable, describing such important effects as loss compensation in metamaterials [44].

3.2. The bistable CW spaser

Bistable devices in microelectronics are MOSFET-based devices that have two stable states describing logical 1 and 0. Their output 'logical' level changes when the input exceeds a certain threshold value. Below we show that the spaser can operate as a bistable device on the basis of the quantum amplification.

Generally, bistability is a result of nonlinearity in the system. We examine below the bistability in the spaser where the nonlinearity is due to the presence of a saturable absorber. Such an absorber is a chromophore whose absorption overlaps with the spasing line, but which does not directly absorb the radiation that pumps the spaser. Formally, it is described by equations (12) and (13) where the pumping rate $g = 0$ for

the index p corresponding to the saturable absorber. We will assume that this absorber is distributed in space in the same way as for the gain chromophores but with a different density ρ_a.

Results of a numerical solution of equations (12) and (13) for different values of ρ_a relative to the concentration ρ of the gain medium chromophores are shown in figure 3. Panel (a) displays the dependence of the SP population number N_n in the spasing mode on the pumping rate g, and panel (b) shows the corresponding dependence of the population inversion n_{21} of the gain medium. We note first that there always exists also a trivial, non-spasing solution $N_n = 0$. For values of g above critical (depending on ρ_a), there are also nontrivial solutions. For $\rho_a > 10^{-3}\rho$, these nontrivial solutions consist of two branches: high N_n and low N_n. Note that the high N_n branch in panel (a) corresponds to the low n_{21} branch in panel (b), and vice versa.

Thus it appears that there is a tristability. However, this impression is incorrect. The lower branches in figure 3(a) (and, correspondingly, the upper branches in panel (b)) describe unstable solutions that are not realizable physically. This can be understood already from the fact that along these branches the SP population decreases with increasing pumping, which is completely unphysical.

It is a bistability, not the tristability, which takes place in actuality. This is illustrated in figures 3(c) and (d) obtained by isolating the stable branches of the solutions. As one can see, with an increase of the pumping rate g, the solutions appear at critical pumping rates that increase with the saturable absorber concentration ρ_a. As the critical g for a given ρ_a is reached, the branch with nonzero N_n appears with a discontinuity. Both the $N_n = 0$ and $N_n > 0$ branches are stable and can retain their states indefinitely. The transition between these two stable states can be induced by either adding SP quanta to or removing them from the spasing mode, as illustrated by arrows for $\rho_a = 3\rho$. This shows that a spaser with a saturable absorber can function as both a nanoscopic memory cell and a bistable device on the basis of quantum amplification. The dynamics of such a device is femtosecond, as we show below in section 4.

In the bistable case there is no perfect pinning of the population inversion n_{21}, as figure 3(d) shows. When the pumping rate g is increased, the system moves along the stable no-spasing ($N_n = 0$) solution denoted by the black line, where the inversion n_{21} significantly overshoots its values for the spasing branches (colored lines). When the transition to spasing occurs, induced, e.g., by an injection of SP quanta, it is always discontinuous, as the vertical arrows indicate. Along the spasing ($N_n > 0$) branches (colored lines), the inversion counter intuitively decreases with increased pumping due to the stimulated emission.

It is instructive to compare the behavior of the spaser shown above in figure 3 (see also its discussion) with the corresponding behavior of lasers. As we have already mentioned in section 1, there are similarities and also substantial differences. A principal source of these differences is due to the fact that the spaser SP modes are much more strongly localized than the photonic modes in conventional lasers and even the SPP modes in the nanolasers of [18, 19].

The much smaller modal volume V_n of the spasing SP modes causes much higher local fields of these SPs (cf figure 1) and, consequently, much stronger feedback in the spaser— cf equation (15). Therefore the spaser operates—including the bistable operation—at a population of SPs many orders of magnitude smaller (typically, $N_n \sim 100$) per spasing mode as compared to the number of quanta in the resonators of the conventional lasers ($\sim 10^{18}$–10^{20}) and even the polaritonic nanolasers. Related to this fact are the very high values of the local fields in the spaser $\sim 10^6 \sqrt{N_n} \sim 10^7$ V cm^{-1}, as pointed out above in this paragraph. These strong local fields and the strong feedback caused by them are also responsible for the pronounced linear dependence of the SP population N_n on the pumping rate (see figure 3(c)) and the pronounced pinning of the population inversion n_{21} of the gain medium (see figures 2(b) and 3(d)) in the spasing state.

4. The spaser as an ultrafast quantum nanoamplifier

4.1. The problem of setting the spaser as an amplifier

Now we consider the central point of this article: setting the spaser as an ultrafast quantum nanoamplifier. As we have already mentioned in the introduction (section 1), the principal and formidable problem is that, in contrast to the conventional lasers and amplifiers in quantum electronics, the spaser has an inherent feedback that cannot be removed. Thus the spaser above the spasing threshold will always develop generation and accumulation of the macroscopic number of coherent SPs in the spasing mode. This leads to the population inversion pinning in the CW regime at a very low level—cf figure 2(b). This CW regime corresponds to the net amplification equal to zero, which means that the gain exactly compensates the loss, which condition is expressed by equation (12). This is a consequence of the nonlinear gain saturation. This holds for any stable CW generator (including any spaser or laser) and precludes using them as amplifiers.

We propose two regimes for setting the spaser as a quantum nanoamplifier. The first is a transient regime based on the fact that the establishing of the CW regime and the consequent inversion pinning and total gain vanishing require some time that is determined mainly by the rate of the quantum feedback, but depends also on the relaxation rates of the SPs and the gain medium. After the population inversion is created by the onset of pumping and before the spasing spontaneously develops, as we show below in this section, there is a time interval of approximately 250 fs, during which the spaser provides usable (and as predicted, quite high) amplification.

The second way to set the spaser as a quantum nanoamplifier is a bistable regime that is achieved by introducing a saturable absorber, which prevents the spontaneous spasing. Then injection of a certain above-threshold amount of SP quanta will saturate the absorber and initiate the spasing. Such a bistable quantum amplifier will be considered in the next subsection. The bistability is the most promising regime which allows not only the functioning of the spaser as an ultrafast threshold (logical) amplifier, but also a quasi-CW bistability where the spaser works as a memory element (cf section 3.2) with an ultrafast switching time.
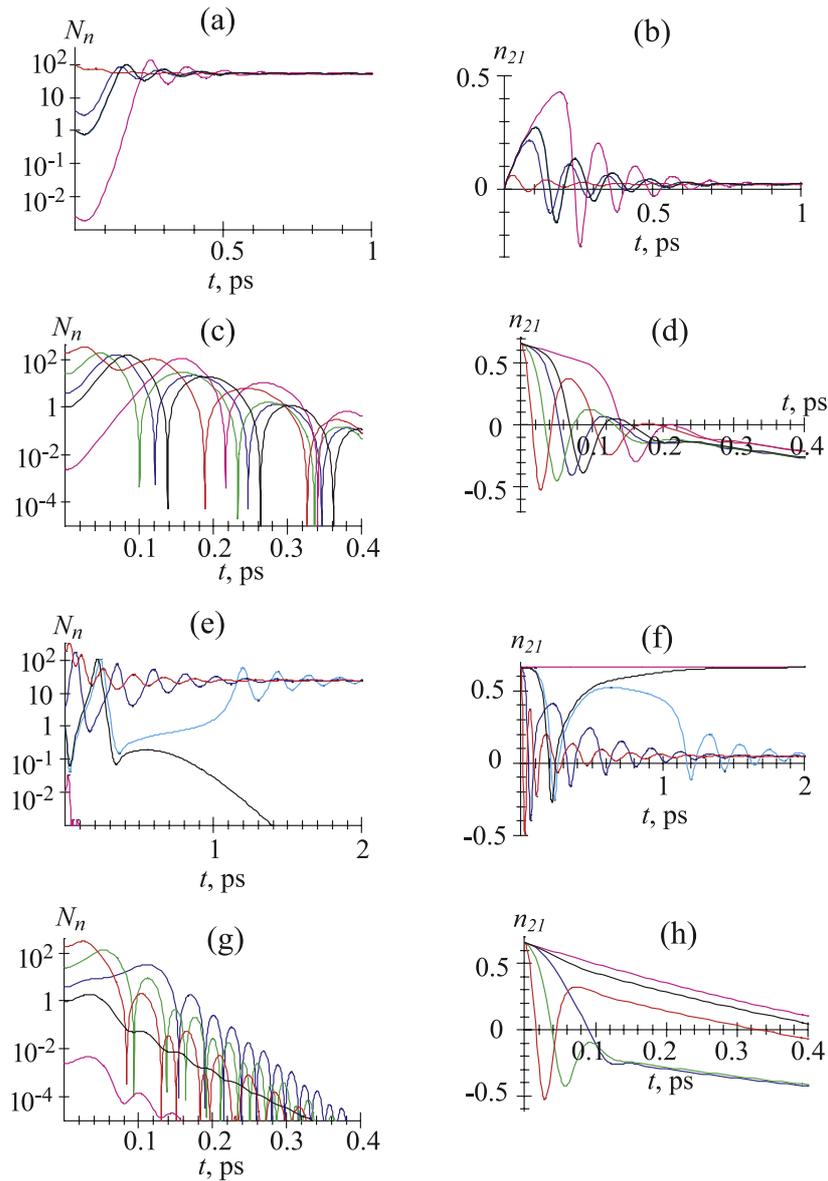
Figure 4. Ultrafast dynamics of the spaser. (a) For a monostable spaser (without a saturable absorber), the dependence of the SP population in the spasing mode N_n on time t. The spaser undergoes stationary pumping at a rate of $g = 5 \times 10^{12}$ s^{-1}. The color coded curves correspond to the initial conditions with the different initial SP populations, as shown in the graphs. (b) The same as (a) but for the temporal behavior of the population inversion n_{21}. (c) Dynamics of a monostable spaser (no saturable absorber) with the pulsed pumping described as the initial inversion $n_{21} = 0.65$. The coherent SP population N_n is displayed as a function of time t. Different initial populations are indicated by color coded curves. (d) The same as (c) but for the corresponding population inversion n_{21}. (e) The same as (a) but for a bistable spaser with the saturable absorber for the concentration $\rho_a = 0.66\rho$. (f) The same as (b) but for the bistable spaser. (g) The same as (e) but for the pulsed pumping with the initial inversion $n_{21} = 0.65$. (h) The same as (g) but for the corresponding population inversion n_{21}.

The temporal behavior of the spaser has been found by direct numerical solution of equations (4)–(6). This solution is facilitated by the fact that in the model under consideration all the chromophores experience the same local field inside the nanoshell, and there are only two kinds of such chromophores: belonging to the gain medium and to the saturable absorber, if it is present.

4.2. The monostable spaser as a nanoamplifier in the transient regime

In this subsection we consider a monostable spaser in a transient regime. This implies that no saturable absorber is

present. We will consider two pumping regimes: stationary and pulsed.

Starting with the stationary regime, we assume that the pumping at a rate (per chromophore) of $g = 5 \times 10^{12}$ s^{-1} starts at a moment of time $t = 0$ and stays constant after that. Immediately at $t = 0$, a certain number of SPs are injected into the spaser. We are interested in its temporal dynamics from this moment on.

The dynamical behavior of the spaser under this pumping regime is illustrated in figures 4(a) and (b). As we see, the spaser, which starts from an arbitrary initial population N_n, rather rapidly, within a few hundred femtoseconds approaches

the same stationary ('logical') level. At this level, an SP population of $N_n = 67$ is established, while the inversion is pinned at a low level of $n_{21} = 0.02$. On the way to this stationary state, the spaser experiences relaxation oscillations in both the SP numbers and inversion, which have a tendency to oscillate out of phase (compare panels (a) and (b)). This temporal dynamics of the spaser is quite complicated and highly nonlinear (anharmonic). It is controlled not by a single relaxation time but by a set of the relaxation rates. Clearly, among these are the rate of energy transfer from the gain medium to the SPs and the relaxation rates of the SPs and the chromophores.

In this mode, the main effect of the initial injection of the SPs (described theoretically as different initial values of N_n) is on the interval of time that is required for the spaser to reach the final (CW) state. For very small N_n, which in practice can be supplied by the noise of the spontaneous SP emission into the mode, this time is approximately 250 fs (cf the corresponding SP relaxation time being less then 50 fs). In contrast, for the initial values of $N_n = 1$–5, this time shortens to 150 fs. Whether this is a practically usable difference remains to be seen at more advanced stages of development.

Now consider the second regime: pulsed pumping. The gain medium population of the spaser is inverted at $t = 0$ to saturation with a short (much shorter than 100 fs) pump pulse. Simultaneously, at $t = 0$, some number of plasmons are injected (say, by an external nanoplasmonic circuitry). In response, the spaser should produce an amplified pulse of the SP excitation. Such a function of the spaser is illustrated in figures 4(c) and (d).

As we see from panel (c), independently of the initial number of SPs, the spaser always generates a series of SP pulses, of which only the first pulse is large (at or above the logical level of $N_n \sim 100$). (An exception is a case of little practical importance when the initial $N_n = 120$ exceeds this logical level, when two large pulses are produced.) The underlying mechanism of such a response is the rapid depletion of the inversion seen in panel (d), where energy is dissipated in the metal of the spaser. The characteristic duration of the SP pulse, ~100 fs, is defined by this depletion, controlled by the energy transfer and SP relaxation rates. This time is much shorter than the spontaneous decay time of the gain medium. This acceleration is due to the stimulated emission of the SPs into the spasing mode (which can be called a 'stimulated Purcell effect'). There is also a pronounced trend: the lower the initial SP population N_n, the later the spaser produces the amplified pulse. In a sense, this spaser functions as a converter of pulse amplitude to time delay.

4.3. The bistable spaser (with saturable absorber) as an ultrafast nanoamplifier

Now let us consider the most important point of this article: a bistable spaser as a quantum threshold ('logical') nanoamplifier. Such a spaser contains a saturable absorber mixed with the gain medium with parameters indicated at the end of section 2.1. In particular, the concentration of the saturable absorber $\rho_a = 0.66\rho$. This case of a bistable

spaser amplifier is of a particular interest because in this regime the spaser comes as close as possible in its functioning to the semiconductor-based (mostly, MOSFET-based) digital nanoamplifiers. As in the previous subsection, we will consider two cases: stationary pumping and pumping with short pulses.

We again start with the case of the stationary pumping at a rate of $g = 5 \times 10^{12}$ s^{-1}. We show in figures 4(e) and (f) the dynamics of such a spaser. For a small initial population $N_n = 5 \times 10^{-3}$ simulating the spontaneous noise, the spaser is rapidly (faster than in 50 fs) relaxing to zero population (panel (e)), while its gain medium population is equally rapidly approaching a high level (panel (f)) $n_{21} = 0.65$ that is defined by the competition of the pumping and the enhanced decay into the SP mode (the purple curves). This level is so high because the spasing SP mode population vanishes and the stimulated emission is absent. After reaching this stable state (which one could call, say, 'logical 0'), the spaser stays in it an indefinitely long time despite the continuing pumping.

In contrast, for initial values N_n of the SP population large enough (for instance, for $N_n = 5$, as shown by the blue curves in figures 4(e) and (f)), the spaser tends to the 'logical 1' state where the stationary SP population reaches the value of $N_n \approx 60$. Due to the relaxation oscillations, it actually exceeds this level within a short time of $\lesssim 100$ fs after the seeding with the initial SPs. As the SP population N_n reaches its stationary (CW) level, the gain medium inversion n_{21} is pinned down at a low level of a few per cent, as is typical for the CW regime of the spaser. This 'logical 1' state also persists indefinitely, as long as the inversion is supported by the pumping.

There is a critical curve (separatrix) that divides the two stable dynamics types (leading to the logical levels of 0 and 1). For the present set of parameters this separatrix starts with the initial population of $N_n \approx 1$. For a value of the initial N_n slightly below 1, the SP population N_n experiences a slow (hundreds of fs in time) relaxation oscillation but eventually relaxes to zero (figure 4(e), black curve), while the corresponding chromophore population inversion n_{21} relaxes to the high value $n_{21} = 0.65$ (panel (f), black curve). In contrast, for a value of N_n slightly higher than 1 (light blue curves in panels (e) and (f)), the dynamics is initially close to the separatrix but eventually the initial slow dynamics tends to the high SP population and low chromophore inversion through a series of relaxation oscillations. The dynamics close to the separatrix is characterized by a wide range of oscillation times due to its highly nonlinear character. The initial dynamics is slowest (the 'decision stage' of the bistable spaser that lasts $\gtrsim 1$ ps). The 'decision time' is diverging infinitesimally close to the separatrix, as is characteristic of any threshold (logical) amplifier.

The gain (amplification coefficient) of the spaser as a threshold (logical or bistable) amplifier is the ratio of the high CW level to the threshold level of the SP population N_n. For this specific spaser with the chosen set parameters, this gain is ≈ 60, which is more than sufficient for the digital information processing. Thus this spaser can make a high gain, ~10 THz bandwidth logical amplifier or dynamical memory cell with excellent prospects of applications.

The last but probably the most important regime to consider is that of the pulsed pumping of the bistable spaser. In

this case, the population inversion ($n_{21} = 0.65$) is created by a short pulse at $t = 0$ and simultaneously initial SP population N_n is created. Both are emulated as the initial conditions in equations (4)–(6). The corresponding results are displayed in figures 4(g) and (h).

When the initial SP population exceeds the critical one of $N_n = 1$ (the blue, green, and red curves), the spaser responds by generating a short (duration less than 100 fs) pulse of the SP population (and the corresponding local fields) within a time of $\lesssim 100$ fs (panel (g)). Simultaneously, the inversion is rapidly (within ∼100 fs) exhausted (panel (h)).

In contrast, when the initial SP population N_n is less than the critical one (i.e., $N_n < 1$ in this specific case), the spaser rapidly (within a time of $\lesssim 100$ fs) relaxes as $N_n \to 0$ through a series of relaxation oscillations—see the black and magenta curves in figure 4(g). The corresponding inversion decays in this case almost exponentially with a characteristic time ∼1 ps determined by the enhanced energy transfer to the SP mode in the metal—see the corresponding curves in panel (h). Note that the SP population decays faster due to the threshold nature of spasing.

Gold as the spaser nanoplasmonic core metal, which is used in experiments [16], has a relaxation rate an order of magnitude higher than that of silver. Therefore it is plausible that a gold-core spaser as an amplifier will have a bandwidth of ∼100 THz. We will consider such a system elsewhere. Taking into account that the local fields produced by the spaser are concentrated within a few tens of nanometers in its vicinity, these are parameters that hold promise of many applications in the fundamental science and coming digital and analog femtosecond technologies.

5. Discussion and conclusion

We have demonstrated the principal possibilities of the spaser functioning as an ultrafast nanoamplifier of the local optical fields. In doing so we have had to overcome a principal and formidable problem. Any spaser has inherent feedback. Consequently, the spaser as proposed initially [4] and observed experimentally [16] rapidly, on the femtosecond timescale, will generate the coherent population of the SPs and approach the CW regime. In this regime, the net gain (the saturated amplification minus loss) is exactly zero (the same as for lasers). Thus CW spasers (or CW lasers, for that matter) cannot fundamentally be used as amplifiers. In the conventional optical amplifiers, in contrast to lasers, the feedback is deliberately removed and any parasitic feedback (scattering, etc) is carefully minimized to prevent the spontaneous generation. Such a regime is not possible for the spaser as it is currently known. In this article, we have demonstrated two ways to set a spaser as a nanoamplifier: (i) the conventional monostable spaser can amplify in the transient ultrafast (femtosecond) regime before the CW generation is established; (ii) the spaser with the saturable absorber can become bistable and function as a threshold (logical) amplifier with ultrafast (femtosecond) switching times.

We have developed quantum theory of the spaser based on density matrix (optical Bloch) equations. The spaser

described is a nanoparticle that consists of a metal core and a gain (active) medium. The metal core plays the role of the resonator ('cavity') whose states are the localized SPs, which we have quantized. The gain medium is described completely quantum mechanically taking into account the excitation, decay, and dephasing processes. The main approximation that we have employed is that the SP population number N_n is treated as a classical variable. This is the same semiclassical approximation as is most often used in the physics of lasers. Given the fact that the typical number of SP quanta per spasing mode is still large enough ($N_n \sim 100$), this approximation appears reasonable and reliable.

Pursuing the goal of the fastest possible operation and the nanometric size (on the same order as that of microelectronic MOSFETs), we have deliberately considered spasers whose size is much less than the radiation wavelength and whose metal thickness is less than the skin depth. This has allowed us to use the quasistatic SP eigenmodes as full counterparts of the electrostatic fields in the transistors. Note that such a nanoscopic spaser has recently been demonstrated experimentally [16].

Though the main emphasis of this article is on the ultrafast dynamics of the spaser as a nanoamplifier of the local plasmonic fields, the CW regime is described first as the basis for understanding the ultrafast dynamics. The corresponding results are illustrated in figure 2. The 'spasing curve' (the dependence of the SP population N_n on the pumping rate) shows a pronounced threshold and a linear dependence of the coherent SP population N_n on the pumping rate g (figure 2(a)). This linear dependence $N_n(g)$ is due to the extremely strong feedback in the spaser caused by the very strong SP mode localization. The same effect is responsible for the clipping of the population inversion in the spasing state to very low levels, as illustrated in figure 2(b). With the increased pumping, a higher SP population leads to lower quantum fluctuations and decreased spectral width of the spasing line, as shown in figure 2(c). Gradually, a very narrow and intense spasing line appears in the emission spectrum of the spaser—see figures 2(d)–(f). This narrow line is indicative of the intense optical fields that are excited in the vicinity of a spaser—cf figure 1(a) where the field must be multiplied by a factor of $\sqrt{N_n}$. Qualitatively, the linear spasing curve and the narrowing of the spasing line with the pumping are in full agreement with the recent experiment [16].

Adding a saturable absorber to the spaser shell, we have shown that it can work as a bistable device on the basis of quantum amplification. The corresponding results are illustrated in figure 3 for a CW operation mode.

The main results of this article concern with ultrafast processes of amplification in the spaser. These are illustrated in figure 4. We have shown that a monostable spaser (without a saturable absorber) performs amplification of the injected SPs in a transient regime. However, this process only manifests itself in shorter times to establish the high SP population—see figures 4(a)–(d) and the corresponding discussion. In this sense, such a spaser performs the function of an amplitude to time converter but not a true nanoamplifier.

The most significant results are those related to the ultrafast amplification in the bistable spaser containing a

saturable absorber within its gain shell. This amplification has many similarities with that for semiconductor threshold (logical) nanoamplifiers based on MOSFETs. Such a spaser with stationary pumping behaves as an ultrafast nanoscale bistable device—see the results in figures 4(e) and (f). For initial conditions not too close to the threshold, it amplifies and switches with a characteristic time of ~100 fs. The bistable spaser with pulsed pumping is a high gain ultrafast pulse nanoamplifier—see figures 4(g) and (h) and the corresponding discussion in the previous section. Its typical response time and pulse length are $\lesssim$100 fs, corresponding to the 10 THz bandwidth.

For the gold-based spaser, in contrast to the present silver-based one, the typical switching time should be much shorter, presumably ~10 fs corresponding to ~100 THz bandwidth. However, the gold-based spasers will work at a lower SP population due to the higher loss in the metal. Correspondingly, they will possess lower signal/noise ratio and wider spectral lines. We will consider such spasers elsewhere.

Concluding, the spaser is the size of the MOSFET [13] (~10 nm) and can perform the same function of signal amplification on the nanoscale. Having a high enough gain, the spaser can be an active element foundation of highly integrated nanoplasmonic devices, including ultrafast processors. Based on metals, it is inherently much faster, by a factor of ~1000, than silicon devices, as we have shown above. For the same reason, it also highly robust environmentally: it can work at high temperatures, in the presence of microwave and ionizing radiations, etc. This article has presented quantum theory of the spaser as a nanoamplifier, which is critically needed for understanding and utilization of the spaser.

Acknowledgments

References

[1] Novotny L and Hecht B 2006 *Principles of Nano-Optics* (Cambridge: Cambridge University Press)
[2] Stockman M I, Kling M F, Kleineberg U and Krausz F 2007 *Nat. Photon.* **1** 539
[3] Stockman M I 2008 *New J. Phys.* **10** 025031
[4] Bergman D J and Stockman M I 2003 *Phys. Rev. Lett.* **90** 027402
[5] Atwater H A 2007 *Sci. Am.* **296** 56
[6] Anker J N, Hall W P, Lyandres O, Shah N C, Zhao J and Duyne R P V 2008 *Nat. Mater.* **7** 442
[7] Israel A, Mrejen M, Lovsky Y, Polhan M, Maier S and Lewis A 2007 *Laser Focus World* **43** 99

[8] Tang L, Kocabas S E, Latif S, Okyay A K, Ly-Gagnon D S, Saraswat K C and Miller D A B. 2008 *Nat. Photon.* **2** 226
[9] Challener W A *et al* 2009 *Nat. Photon.* **3** 220
[10] Kim S, Jin J H, Kim Y J, Park I Y, Kim Y and Kim S W 2008 *Nature* **453** 757
[11] Nagatani N, Tanaka R, Yuhi T, Endo T, Kerman K and Tamiya Y T E 2006 *Sci. Technol. Adv. Mater.* **7** 270
[12] Hirsch L R, Stafford R J, Bankson J A, Sershen S R, Rivera B, Price R E, Hazle J D, Halas N J and West J L 2003 *Proc. Natl Acad. Sci. USA* **100** 13549
[13] Kahng D 1963 *US Patent Specification* 3,102,230
[14] Tsividis Y 1999 *Operation and Modeling of the MOS Transistor* (New York: McGraw-Hill)
[15] Stockman M I 2008 *Nat. Photon.* **2** 327
[16] Noginov M A, Zhu G, Belgrave A M, Bakker R, Shalaev V M, Narimanov E E, Stout S, Herz E, Suteewong T and Wiesner U 2009 *Nature* **460** 1110
[17] Hill M T *et al* 2007 *Nat. Photon.* **1** 589
[18] Hill M T *et al* 2009 *Opt. Express* **17** 11107
[19] Oulton R F, Sorger V J, Zentgraf T, Ma R-M, Gladden C, Dai L, Bartal G and Zhang X 2009 *Nature* **461** 629
[20] Stockman M I, Faleev S V and Bergman D J 2001 *Phys. Rev. Lett.* **87** 167401
[21] Larkin I A and Stockman M I 2005 *Nano Lett.* **5** 339
[22] Gordon J A and Ziolkowski R W 2007 *Opt. Express* **15** 2622
[23] Bergman D J and Stroud D 1992 *Solid State Physics* vol 46, ed H Ehrenreich and D Turnbull (Boston: Academic) pp 148–270
[24] Bergman D J and Stroud D 1980 *Phys. Rev.* B **22** 3527
[25] Kirakosyan A S, Stockman M I and Shahbazyan T V 2009 arXiv:0908.0647
[26] Plum E, Fedotov V A, Kuo P, Tsai D P and Zheludev N I 2009 *Opt. Express* **17** 8548
[27] Seidel J, Grafstroem S and Eng L 2005 *Phys. Rev. Lett.* **94** 177401
[28] Noginov M A, Zhu G, Mayy M, Ritzo B A, Noginova N and Podolskiy V A 2008 *Phys. Rev. Lett.* **101** 226806
[29] Li K, Li X, Stockman M I and Bergman D J 2005 *Phys. Rev.* B **71** 115409
[30] Dong Z G, Liu H, Li T, Zhu Z H, Wang S M, Cao J X, Zhu S N and Zhang X 2008 *Opt. Express* **16** 20974
[31] Wegener M, Garcia-Pomar J L, Soukoulis C M, Meinzer N, Ruther M and Linden S 2008 *Opt. Express* **16** 19785
[32] Fang A, Koschny T, Wegener M and Soukoulis C M 2009 *Phys. Rev.* B **79** 241104(R) (Rapid Communications)
[33] Chang S W, Ni C Y A and Chuang S L 2008 *Opt. Express* **16** 10580
[34] Zheludev N I, Prosvirnin S L, Papasimakis N and Fedotov V A 2008 *Nat. Photon.* **2** 351
[35] Protsenko I E, Uskov A V, Zaimidoroga O A, Samoilov V N and O'Reilly E P 2005 *Phys. Rev.* A **71** 063812
[36] Ambati M, Nam S H, Ulin-Avila E, Genov D A, Bartal G and Zhang X 2008 *Nano Lett.* **8** 3998
[37] Zhou Z K, Su X R, Peng X N and Zhou L 2008 *Opt. Express* **16** 18028
[38] Noginov M A, Podolskiy V A, Zhu G, Mayy M, Bahoura M, Adegoke J A, Ritzo B A and Reynolds K 2008 *Opt. Express* **16** 1385
[39] Adamo G, MacDonald K F, Fu Y H, Wang C M, Tsai D P, de Abajo F J G and Zheludev N I 2009 Light well: a tunable free-electron light source on a chip *Phys. Rev. Lett.* **103** 113901
[40] Noginov M A, Zhu G, Bahoura M, Adegoke J, Small C, Ritzo B A, Drachev V P and Shalaev V M 2007 *Appl. Phys.* B **86** 455
[41] Noginov M A 2008 *J. Nanophoton.* **2** 021855
[42] Johnson P B and Christy R W 1972 *Phys. Rev.* B **6** 4370
[43] Schawlow A L and Townes C H 1958 *Phys. Rev.* **112** 1940
[44] Shalaev V M 2007 *Nat. Photon.* **1** 41