A Review of High-Gain Free-Electron Laser Theory

Nicola Piovella 1,*, and Luca Volpe 2

1 Dipartimento di Fisica “Aldo Pontremoli”, Università degli Studi di Milano, Via Celoria 16, 20133 Milano, Italy
2 Centro de Laseres Pulsados (CLPU), Parque Científico, 37185 Salamanca, Spain; lvolpe@clpu.es
* Correspondence: nicola.piovella@mi.infn.it

Abstract: High-gain free-electron lasers, conceived in the 1980s, are nowadays the only bright sources of coherent X-ray radiation available. In this article, we review the theory developed by R. Bonifacio and coworkers, who have been some of the first scientists envisaging its operation as a single-pass amplifier starting from incoherent undulator radiation, in the so called self-amplified spontaneous emission (SASE) regime. We review the FEL theory, discussing how the FEL parameters emerge from it, which are fundamental for describing, designing and understanding all FEL experiments in the high-gain, single-pass operation.

Keywords: free-electron laser; X-ray emission; collective effects

1. Basic Concepts

The free-electron laser is essentially a device that transforms the kinetic energy of a relativistic electron beam (e-beam) into e.m. radiation [1–4]. The e-beam passing through a transverse periodic magnetic field oscillates in a direction perpendicular to the magnetic field and the propagation axis, and emits radiation confined in a narrow cone along the propagation direction. The periodic magnetic field is provided by the so-called undulator, an insertion device usually realized with two arrays of permanent magnets with alternating polarities (see Figure 1) or with two helical coils with current circulating in opposite directions. The wavelength of the emitted radiation depends on the undulator period, on the strength of the magnetic field and on the electron energy. This means that FELs can be continuously tuned in wavelength, ranging from microwaves (cm) to X-rays (Å); this is one of the main advantages of FELs with respect to atomic and molecular lasers, where the wavelength of the radiation field is fixed by the quantum transition between two atomic or molecular states, and it has, in general, a small tunability. Moreover, another advantage of the FEL with respect to atomic and molecular lasers is that its main processes happen in vacuum, with no thermal dispersion or breakdown effects in the active medium. The FEL is a powerful source of tunable, coherent e.m. radiation [5]. In general, the FEL radiation can be of two different kinds: basic spontaneous emission, i.e., synchrotron radiation, which comes from the direct interaction of the e-beam with the undulator magnetic field with no injected field, while stimulated emission occurs when also a seed radiation field co-propagates with the electron beam. We use the term "spontaneous emission" in analogy with the atomic laser physics (i.e., emission from the spontaneous decay of the atomic excited state) although its origin here is different: it corresponds to the classical incoherent undulator emission from accelerated charges.

1.1. FEL Spontaneous Emission

It is well-known that a free charge cannot radiate because of the energy–momentum conservation; the word “free” in FEL describes the state of the electrons used as active medium: the electrons are not bound as in atomic or molecular lasers. In the FEL, the electrons are not really “free” because they are accelerated by the periodic magnetic field. This
accelerated motion generates a radiation field which is particularly intense and directed
in a small cone around the instantaneous velocity vector. The FEL radiation is really a
synchrotron radiation, i.e., the radiation emitted by an electric charge moving at relativistic
speed when a transverse force is applied to it. This radiation power is greater than that
due to a parallel force by a factor $\gamma^2$, where

$$\gamma = \frac{1}{\sqrt{1 - \beta^2}},$$

is the Lorentz factor, $\gamma mc^2$ the electron energy, $m$ the electron mass unit and $\beta = v/c$.

![Diagram of free-electron laser](image)

**Figure 1.** Scheme of free-electron laser. Here the undulator is built by permanent magnets, where the
period corresponds to the distance between two red arrows. The electrons wiggle along the undulator
axis and emit a radiation burst in the forward direction.

The electrons inside the undulator are periodically deflected by the Lorentz force

$$\mathbf{F} = -e \mathbf{v} \times \mathbf{B}_w$$

exerted by the magnetic field $\mathbf{B}_w$ on electrons with charge $-e$, traveling at speed $v$. Since
the magnetic field has a periodic alternated polarization, the electrons will “wiggles”,
i.e., oscillate transversely along the undulator axis. Its main features are the following:

(i) The intensity is proportional to the electrons’ current, i.e., the radiation is incoherent
($I \propto N_e$ where $N_e$ is the number of electrons).

(ii) The emitted radiation is confined in a narrow cone along the direction of electrons
motion (that will be identify with the $z$-axis) within an angle of order of $\approx 1/\gamma$.

(iii) It is a narrow-band radiation, with on-axis spectral distribution

$$\frac{d^2 I}{d\Omega d\omega} = 4N_w^2 \gamma^2 a_w^2 \text{sinc}^2 \left( \frac{\pi N_w \omega - \omega_s}{\omega_s} \right)$$

where $\text{sinc}(x) = \sin(x)/x$,

$$a_w = \frac{e \lambda_w B_w}{2 \pi mc}$$

is the undulator parameter, $\lambda_w$ is the undulator period and $N_w$ is the number of undu-
lator periods. The spectrum is peaked around a spontaneous frequency $\omega_s = 2\pi c/\lambda_s$,
where, for on-axis radiation,

$$\lambda_s = \frac{1 - \beta_\parallel^2}{\beta_\parallel} \lambda_w$$

where $\beta_\parallel = \langle v_\parallel \rangle / c$ and $\langle v_\parallel \rangle$ is the average longitudinal velocity. The resonant
condition (5) can be derived requiring that during the time necessary for an electron
to travel an undulator period $\lambda_w$, the electromagnetic wave slips over it by a radiation wavelength $\lambda_s$,

$$\frac{\lambda_w}{v_\parallel} = \frac{\lambda_w + \lambda_s}{c}. \quad (6)$$

The radiation line-width, from Equation (3), is

$$\frac{\Delta \omega}{\omega_s} \simeq \frac{1}{N_w}. \quad (7)$$

The above result can be easily understood in the (average) longitudinal electron rest frame: here each electron “sees” an $N_w$-periods undulator magnetic field as an $N_w$-periods counter-propagating pseudo-radiation field (known as “Weizsacker–Williams Approximation” [6,7]), with Lorentz contracted wavelength $\lambda_w' = \lambda_w / \gamma_\parallel$. Hence, it oscillates $N_w$ times, emitting a sinusoidal wave train of length $N_w\lambda_w'$ at a wavelength $\lambda_s' = \lambda_w'$. In other terms, it acts as a “relativistic mirror” where the radiation is back-reflected. From this picture, we obtain the same result of Equations (3)–(7). In fact, by Lorentz-transforming the incident and reflecting wavelengths $\lambda_w'$ and $\lambda_s'$ back to the laboratory frame, we obtain the relation (5). Moreover, it is well-known that the Fourier transform of a plane-wave truncated after $N_w$ oscillations is a sinc-function with line-width $\Delta \omega/\omega_s = 1/N_w$.

From Equation (6), we can write $(k_w + k_s)dz/dt - \omega_s = 0$, (where $k_w = 2\pi/\lambda_w$ and $v_\parallel = dz/dt$) or equivalently $d[(k_w + k_r)z - \omega_rt]/dt = 0$, showing that the resonant relation can be obtained also by imposing the relative phase

$$\theta = (k_w + k_s)z - \omega_s t \quad (8)$$

of the electron in the undulator and e.m. fields to be constant. As we will show later, the module of electron transverse velocity $v_\perp = c\beta_{\perp}$ is approximately:

$$\beta_{\perp} \simeq \frac{a_w}{\gamma}. \quad (9)$$

Thus, from $1/\gamma^2 = 1 - \beta_\parallel^2 - \beta_\perp^2$, it follows

$$\frac{1}{\gamma_\parallel} = \frac{1 + a_w^2}{\gamma^2}. \quad (10)$$

where $\gamma_\perp^2 = 1/(1 - \beta_\perp^2)$. Finally, using the resonant condition (5) in the ultrarelativistic limit $\gamma_\parallel \gg 1$, we obtain:

$$\lambda_s = \frac{1 - \beta_\parallel}{\beta_\parallel} \lambda_w \simeq \frac{\lambda_w}{2\gamma_\parallel} \simeq \frac{\lambda_w}{2\gamma^2} \simeq \lambda_w \frac{1 + a_w^2}{2\gamma^2}. \quad (11)$$

This relation shows the high tunability of the FEL; in fact, the wavelength $\lambda_s$ can be tuned by varying either the electron energy $\gamma$ or the undulator magnetic field $B_w$.

1.2. FEL Stimulated Emission

Stimulated emission takes place when a radiation field with wavelength $\lambda \simeq \lambda_s$ co-propagates with the electron beam inside the undulator. From the resonant relation (11) one can define the resonant electron energy

$$\gamma_r = \sqrt{\frac{\lambda_w(1 + a_w^2)}{2\lambda}} \quad (12)$$
If the electron energy and the radiation wavelength preserve the relation (12), then the relative phase (8) between the transverse oscillations of the electron and the radiation remains constant. Depending on the value of this relative phase, one of these processes occurs for each electron:
(a) the electron gives energy to the field and decelerates, i.e., stimulated emission which provides “gain”,
(b) the electron takes energy from the field and accelerates, i.e., absorption.

If the first of these two processes dominates, then the injected radiation field is amplified, as in the first amplifier experiment [8]; moreover, if the undulator is long enough or if the process happens in an optical cavity, then the spontaneous emission is amplified, as in the first FEL oscillator experiment [9]. This oversimplified picture of the FEL gain process is at the single-particle level and leads to the Madey’s small-gain regime [10]: let us consider a “long” electron pulse so that the slippage of the radiation over the electrons can be neglected (steady-state regime). Then, the initial electron phases \( \theta \) are randomly distributed over a radiation wavelength. As a consequence, for a nearly mono-energetic and resonant electron beam, on average half of the electrons will decelerate and half of the electrons will accelerate, with the result that no net gain will occur. The Madey’s small-gain regime occurs when we inject, in a “short” undulator, an electron beam with average energy slightly above resonance, \( \langle \gamma \rangle_0 > \gamma_r \), such that gain (slightly) prevails over absorption [11].

1.3. High-Gain Regime and SASE

More generally, electrons will communicate with each other via the common radiation field. In fact, if the undulator is long enough and the electron current is high enough, then the electrons will start to bunch within an optical wavelength: electrons faster than \( \gamma_r \) will decelerate, slower electrons will accelerate, so that the electron energy will be driven toward resonance. This energy modulation, after a transient time called “lethargy”, will become a space modulation, i.e., the electrons will start to longitudinally bunch in pancakes on the scale of the radiation wavelength (micro-bunching), around a phase that produces gain. Since most of electrons will have, at this stage, nearly the same phase, they will emit coherently. The quantity measuring how strongly bunched the electrons are is the bunching parameter [12]:

\[
b \equiv \frac{1}{N_e} \sum_{j=1}^{N_e} e^{-i\theta_j} \equiv \langle e^{-i\theta} \rangle. \tag{13}\]

Given the physical meaning and mathematical definition of the phase \( \theta \), it should be clear how \( b \) is the measure of the longitudinal modulation of the electron beam on the scale of the radiation wavelength. A bunching equal to zero represents a completely uniform/random distribution of phases, while an ideal bunching of \( |b| = 1 \) can only be possible with all electrons perfectly in phase. We shall see the existence of a collective instability for the system, which leads to electron self-bunching (up to a value \( |b| \sim 0.8 \)) and to exponential growth of radiation until saturation, which set a limit on the conversion of kinetic electron energy into radiation energy; this is high-gain steady-state regime [12,13]. In this regime, the peak of the radiated power is proportional to \( N_e^{4/3} \) (see Equation (99)), whereas in the incoherent radiation case the power is only proportional to \( N_e \).

New effects appear when the propagation effect, due to the different velocity of the electrons and the radiation beam, is taken into account. As a consequence of this effect (known as slippage), the radiation pulse moves ahead with respect to the electron pulse by a radiation wavelength \( \lambda_r \) every undulator period \( \lambda_w \), i.e., by \( l_s = N_w \lambda_r \) at the end of a \( N_w \) period undulator. From analytical and numerical studies, it has been shown [14–17] that when the slippage length \( l_s \) is larger than a properly defined cooperation length \( l_c \) (defined as the slippage in a gain length, see later), an FEL can operate in the superradiant regime, in which the peak of the radiation power scales as \( N_e^2 \). Notice that radiation intensities scaling as \( N_e^2 \) may also arise from coherent synchrotron radiation emitted by electrons which have been pre-bunched by an external source. The radiation field emitted by these
pre-bunched electrons sum up coherently to give a $N_e^2$ scaling; this superradiant emission is typical from coherently prepared systems. In FEL the electrons enter the undulator in an unprepared state, so that emission is initially proportional to $N_e$ (incoherent emission). Then, the electrons begin to bunch on interacting with the spontaneous radiation and undulator field, which, with the additional slippage effect, leads to a radiated intensity proportional to $N_e^2$. Roughly speaking, the slippage introduces a dissipative mechanism for the radiation which, as soon it is emitted, it quickly escapes from the interaction volume inhibiting the re-absorption process. This feature is fundamental for the occurrence of superradiance [18]. Superradiance in a single-pass high-gain FEL has been observed experimentally since 2007 [19].

The presence of the collective instability (i.e., an exponential growth with a rate depending on $N_e$) allows an FEL to operate also in the absence of a seed signal (i.e., self-emission), starting from the fluctuations in the initial particle phases; this mode operation is called self-amplified spontaneous emission (SASE). In the SASE-mode, the radiation pulse contains a random superposition of narrow superradiant spikes [20]. In the time domain, the noisy pattern of the bunching along the electron beam leads to a superradiant spiking in the radiation pulse. The initial irregular spiking, seeded by non-uniformities on the scale of the radiation wavelength, cleans up and tends to a more regular pattern, with one spike every cooperation length. The position of the spikes, however, is random and depends strongly on the initial noise pattern of the bunching.

The SASE-FEL mode operation is important at wavelengths not accessible to coherent radiation sources, such as in the XUV and X region. The realization of a single-pass high-gain FEL in the SASE mode has been realized in the last decades and it is presently one (if not the only one) bright and powerful X-ray source in the world. Such sources emit X-ray pulses with a broad spectrum composed by many random superradiant spikes [20] (see the example shown in Figure 2). This drawback can be alleviated by different techniques, for instance by a seed signal injected at the entrance of the undulator, forcing the electrons to emit in the narrow-band spectrum imposed by the seed [21,22]. However, a more fundamental way to "clean" the SASE spectrum has been proposed operating an FEL in a quantum regime [23], where the SASE spectrum reduces to a single narrow line, strongly increasing the longitudinal coherence of a SASE-FEL-based X-ray source [24]. However, the price to pay is more stringent experimental requirements of the electron beam quality, in terms of emittance and energy spread, at the borderline of the present electron beam technology.

![Figure 2. Typical SASE spectrum, in this case from the simulations of an experiment at $\lambda = 2$ nm (from [25]).](image-url)
1.4. Quantum FEL

Most of the properties of FEL have been analyzed and found remaining in a strictly classical framework. Additionally, if the original proposal of Madey was based on a quantum description of FEL [10], the derived expression for the gain (known as “Madey gain”) is independent on the Planck constant \( h \), i.e., is classical; however, in attempt to extend the range of the FEL operation into the X-ray region, the quantum effects must be taken into account. In fact, at these wavelengths, the electron recoil due to the emission of a photon becomes comparable to the emission line-width, and a quantum mechanical treatment of the electron–photon interaction becomes compulsory. It is well known that the borderline between the classical and quantum world is defined by the Heisenberg Uncertainly Principle (HUP), which can be written in terms of the position and momentum uncertainty, \( \Delta p \Delta x \geq \hbar \), or in terms of the energy and time uncertainty, \( \Delta E \Delta t \geq \hbar \) (where \( \hbar = h / 2\pi \)).

When the Heisenberg inequality is strong (i.e., \( \gg \)) quantum effects can be neglected and the system behaves classically, conversely when the inequality tends to similarity (i.e., \( \sim \)) quantum effects become relevant and a quantum theory is necessary. Assuming that for a relativistic electron with energy \( E_e = mc^2 \gamma \), the interaction time is proportional to the inverse of the emitted photon frequency \( \Delta t \propto \omega_r^{-1} \) and the energy–time uncertainly relation can be written as follows

\[
\delta \gamma \geq \frac{\lambda_c}{2\pi \lambda_r} \tag{14}
\]

where \( \lambda_r = 2\pi \omega_r / \omega_r \) is the emitted photon wavelength, and \( \lambda_c = h / mc = 0.024 \, \text{Å} \) is the Compton wavelength. This relation shows that when reducing the emission wavelength, the strong inequality is not satisfied anymore. Therefore, for a complete understanding of the basic FEL process, we need a quantum theory which describes the interaction between the beam electrons and the emitted photons in terms of discrete momentum exchange. Since a quantum theory should tend to a classical description when the inequality in the HUP is strong (i.e., \( \gg \)), we should look for a parameter ruling the transition between the classical and the quantum regime. This parameter can be found starting from Equation (14) which can be written as follows

\[
\frac{mc\delta \gamma}{\hbar k_r} \geq 1 \tag{15}
\]

where \( k_r = 2\pi / \lambda_r \). The dimensionless left side of Equation (15) can be chosen as our transition parameter, because it represents the ratio between the maximum classical electron momentum spread \( mc\delta \gamma \) and the photon recoil momentum \( \hbar k_r \). The classical regime occurs when the electron momentum recoil is greater than the photons recoil, i.e., \( mc\delta \gamma \gg \hbar k_r \), whereas the quantum effects become important in the opposite. We observe that the energy spread \( \delta \gamma \) has a two-fold relevance: (a) if it refers to the initial energy spread, i.e., the energy spread of the electron beam injected in the undulator, the condition set a restriction on the quality of the electron beam; (b) if it refers to the electron energy spread induced by the momentum recoil due to the photon emission, the condition set a restriction on the strength of the coupling between electron and light. Since in the high-gain regime the maximum induced energy spread is \( \delta \gamma / \gamma \sim \rho \) (where \( \rho \) is the FEL parameter, see later), then the quantum FEL regime occurs for \( [24,26] \)

\[
\rho \equiv \frac{mc\gamma \rho}{\hbar k_r} \lesssim 1. \tag{16}
\]

The consequence that this condition implies for an experimental realization of the quantum FEL regime will be discussed in Section 4.3. However, we outline that this review presents mainly the well consolidated classical FEL theory, without discussing in details the relatively young theory of quantum FEL, which, due to some yet unsolved aspects, deserves a more appropriate presentation in an evolving research context.
2. Classical Model of Equations

From the discussion of the previous section, it results that a proper classical theory describing a high-gain FEL amplifier must be a many-particle theory. Moreover, the electron dynamics should be self-consistently related to the evolution of the e.m. field dynamics. One can set a self-consistent scheme for the FEL dynamics (similarly to that of the Maxwell–Bloch equations in laser physics [27]) which couple the Maxwell equations with the Newton–Lorentz equations for charged particles moving at relativistic speed in an e.m. field; alternatively one can use a Hamiltonian approach, as done in [28–30]. We will obtain the electron dynamics equations starting from a full 3D relativistic Hamiltonian of a single electron interacting with a circular polarized e.m. field and counter-propagating laser beam (e.m. undulator) (instead of the usual static undulator). This choice is motivated by recent increase of interest toward the laser undulator. However, it will be easy to adapt the obtained results to the usual static undulator. The evolution equations of the e.m. field are derived from the Maxwell equations in the slowly varying envelope approximation (SVEA, see later). The longitudinal electrostatic interaction between electrons (i.e., space-charge effects) will be also taken into account, but only in the 1D approximation.

3. 3D FEL Model

The circularly polarized laser undulator and the radiation fields are characterized by the transverse vector potential \( A \) which can be written in terms of the polarization vector \( \hat{e} = (\hat{x} + i\hat{y})/\sqrt{2} \) as:

\[
\begin{align*}
\mathbf{a} &= \frac{e}{mc} \mathbf{A} = \mathbf{a}_L + \mathbf{a}_r, \\
\mathbf{a}_L &= \frac{\hat{e}}{\sqrt{2}} a_L e^{-ik_L(z+ct)} + \text{c.c} \\
\mathbf{a}_r &= -i \frac{\hat{e}}{\sqrt{2}} a_r e^{ik_L(z-ct)} + \text{c.c} \\
\mathbf{a}_L, r &= \frac{e}{mc^2} k_L, r E_L, r (x, z) 
\end{align*}
\]

where \( E_{L,r} \) and \( k_{L,r} = 2\pi/\lambda_{L,r} \) are the electric field and the wave number of the laser undulator and of the radiation field with frequency \( \omega_{L,r} = ck_{L,r} \), respectively.

3.1. 3D Hamiltonian

The time-dependent Hamiltonian equations can be derived from the modified Hamiltonian principle [28]

\[
\delta \int_{t_1}^{t_2} \left( p_x dx/dt + p_y dy/dt + p_z dz/dt - H \right) dt = 0. \tag{20}
\]

Since we are interested in the systems evolution along the z axis, we change the independent variable from \( t \) to \( z \), and using \( H = E \) (where \( E \) is the total energy), we obtain:

\[
\delta \int_{z_1}^{z_2} \left( p_x dx/dz + p_y dy/dz - E dt/dz + p_z dz \right) dz = 0. \tag{21}
\]

In Equation (21), \((x, p_x), (y, p_y), (t, -E)\) appear as the canonical variables with respect to a new Hamiltonian \( H_1 = -p_z \). Hence, we can write

\[
\begin{align*}
\frac{dx}{dz} &= \frac{\partial p_z}{\partial p_x}, \quad \frac{dy}{dz} = \frac{\partial p_z}{\partial p_y}, \\
\frac{dt}{dz} &= \frac{\partial p_z}{\partial E}, \quad \frac{dE}{dz} = -\frac{\partial p_z}{\partial t}. 
\end{align*} \tag{22}
\]

\[
\begin{align*}
\frac{dp_z}{dz} &= \frac{\partial p_z}{\partial x}, \\
\frac{dp_x}{dz} &= \frac{\partial p_z}{\partial y}. 
\end{align*} \tag{23}
\]
Let $H$ be the relativistic Hamiltonian for an electron interacting with e.m. field

$$H = c \sqrt{(p + eA)^2 + m^2c^2} = mc^2\gamma - eV = E.$$  \hspace{1cm} (24)

where $E_z = -dV/dz$. The second equation in (23) yields:

$$\frac{d\gamma}{dz} = -\left(\frac{1}{mc^2} \frac{\partial P_z}{\partial t} - e_z\right),$$

$$e_z = -\frac{e}{mc^2} E_z = -\frac{d\Phi}{dz}, \quad \Phi = -\frac{e}{mc} V$$  \hspace{1cm} (25, 26)

and the new Hamiltonian is

$$H(x, P_x, y, P_y, ct, -\gamma; z) = -\sqrt{\gamma^2 - 1 - \frac{P_x^2}{\gamma^2} - \frac{P_y^2}{\gamma^2} - |a|^2 - 2 P_\perp \cdot a + \Phi}$$  \hspace{1cm} (27)

where $P_\perp = \gamma(dx_\perp/dz) + a$ and $H = -P_z$ are respectively the transverse and longitudinal momenta, in $mc$ units. The above Hamiltonian can be simplified with the following approximations:

1. the fast oscillating term $2P_\perp \cdot a$ can be neglected,
2. the ultra-relativistic limit, $\gamma \gg \sqrt{1 + |P_\perp|^2 + |a|^2}$ is assumed,
3. the small term $|a|^2 \ll |a|^2$ is neglected.

With 1–3, the Hamiltonian (27) is approximated by

$$H(x_\perp, P_\perp, ct, -\gamma; z) = -\gamma + \frac{1}{2\gamma} \left[1 + |P_\perp|^2 + |a|^2 - i\left(a_\perp^* \cdot a_\perp e^{i\theta} - c.c.\right)\right] + \Phi.$$  \hspace{1cm} (28)

where $\theta = (k_r + k_\perp)z - c(k_r - k_\perp)t$ is electron phase in the laser-undulator and radiation potential. From the Hamiltonian (28), we obtain the equations of motion:

$$\frac{dx}{dz} = \frac{\partial H}{\partial P_x} = \frac{P_x}{\gamma},$$  \hspace{1cm} (29)

$$\frac{dy}{dz} = \frac{\partial H}{\partial P_y} = \frac{P_y}{\gamma},$$  \hspace{1cm} (30)

$$\frac{dP_x}{dz} = -\frac{\partial H}{\partial x} = -\frac{1}{2\gamma} \frac{\partial}{\partial x} \left[|a|^2 - i\left(a_\perp^* a_\perp e^{i\theta} - c.c.\right)\right],$$  \hspace{1cm} (31)

$$\frac{dP_y}{dz} = -\frac{\partial H}{\partial y} = -\frac{1}{2\gamma} \frac{\partial}{\partial y} \left[|a|^2 - i\left(a_\perp^* a_\perp e^{i\theta} - c.c.\right)\right],$$  \hspace{1cm} (32)

$$\frac{dt}{dz} = \frac{\partial H}{\partial ct} = 1 + \frac{1}{2\gamma^2} \left[1 + \frac{P_x^2}{\gamma^2} + \frac{P_y^2}{\gamma^2} + |a|^2 - i\left(a_\perp^* a_\perp e^{i\theta} - c.c.\right)\right],$$  \hspace{1cm} (33)

$$\frac{d\gamma}{dz} = \frac{\partial H}{\partial (ct)} + e_z = -\frac{k_r}{2\gamma} \left[|a|^2 - c.c.\right] + e_z.$$

Notice that, using the definition of the total momenta of an electron in e.m. potential normalized to $mc$, $P_\perp \equiv \beta_\perp \gamma = (dx_\perp/dz)\gamma - a$, we obtain the transverse velocity of the electron

$$\beta_\perp = \frac{a}{\gamma} + \frac{dx_\perp}{dz}$$  \hspace{1cm} (35)

In the paraxial approximation, $dx_\perp/dz \ll 1$ and Equation (35) reduces to the simple expression $\beta_\perp = a/\gamma$. 
3.2. Maxwell Evolution Equations

The evolution of the e.m. transverse potential \( \mathbf{a}_r \) and of the longitudinal electrostatic field \( e_z = -\partial \Phi / \partial z \) can be determined from the rescaled Maxwell equation, expressed in the Coulomb gauge (\( \nabla \cdot \mathbf{a}_r = 0 \)):

\[
\left( \nabla_\perp^2 + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) \mathbf{a}_r = -\frac{\mu_0 c}{e} \mathbf{j}_\perp \tag{36}
\]

\[
\left( \nabla_\perp^2 + \frac{\partial^2}{\partial z^2} \right) \Phi = -\frac{e}{\varepsilon_0 mc} \varrho \tag{37}
\]

where \( \mathbf{j}_\perp \) and \( \varrho \) are the transverse density current and the charge density, respectively, for a beam of \( N_e \) electrons:

\[
\mathbf{j}_\perp = -e \sum_{j=1}^{N} \mathbf{b}_\perp \delta^{(3)}(\mathbf{x} - \mathbf{x}_j(t)), \quad \varrho = -e \sum_{j=1}^{N} \delta^{(3)}(\mathbf{x} - \mathbf{x}_j(t)). \tag{38}
\]

3.2.1. Vector Potential

The evolution of the transverse e.m. field can be obtained in the following way: substituting the electron current (38) in (36), we get

\[
\left[ D^+ D^- + \nabla_\perp^2 \right] \mathbf{a}_r = \frac{e^2}{\varepsilon_0 mc^2} \sum_{j=1}^{N} \mathbf{b}_\perp \mathbf{d}^2(\mathbf{x}_\perp - \mathbf{x}_{\perp,j}(t)) \delta(z - z_j(t)) \tag{39}
\]

where we have defined \( D^\pm = \partial / \partial z \pm (1/c) \partial / \partial t \). The radiation beam which propagates in the opposite direction with respect to the electron motion can be neglected because it does not interact resonantly with the electrons. The complex amplitude \( a_r(\mathbf{x}_\perp,z) \) is assumed to be a slowly varying function of \( z \) and \( t \); this means that the fast variation on the radiation wavelength scale does not affect significantly the field enveloped amplitude. This is known as slowly varying envelope approximation (SVEA):

\[
\left| \frac{\partial a_r(\mathbf{x}_\perp,z,t)}{\partial z} \right| \ll k_r |a_r(\mathbf{x}_\perp,z,t)| \tag{40}
\]

\[
\left| \frac{\partial a_r(\mathbf{x}_\perp,z,t)}{\partial t} \right| \ll \omega_r |a_r(\mathbf{x}_\perp,z,t)| \tag{41}
\]

Notice that this approximation is based on the presence, in the FELs dynamics, of two different scales; one of the order of the radiation wavelength, and the other on the order of the interaction length, normally three or four orders larger than the radiation wavelength. More precisely (as we will show in Section 3.7), the second scale length coincides with the “cooperation” length \( l_c \); moreover, the possibility to separate the two different scales can be done more rigorously by using a multiple scaling approach [31]. As a result of (SVEA), the second-order derivatives in longitudinal and temporal coordinates in Equation (39) can be neglected and

\[
\left[ D^+ D^- + \nabla_\perp^2 \right] \mathbf{a}_r \simeq \sqrt{2\delta k_r} e^{ik_r(z-c t)} \left[ D^+ + \frac{\nabla_\perp^2}{2ik_r} \right] a_r + \text{c.c.}
\]

Substituting Equation (35) for the electron transverse velocity \( \mathbf{b}_\perp \), and projecting the Equation (39) on the \( \mathbf{e}_r \) direction, we obtain

\[
\left[ D^+ + \frac{\nabla_\perp^2}{2ik_r} \right] \mathbf{a}_r = \frac{e^2}{\varepsilon_0 mc^2 k_r} \sum_{i=1 \atop \text{e}}^{N} \left\{ \frac{a_{\perp,j}(\mathbf{x}_\perp,z) e^{i\theta_j}}{r_{ij}} + \frac{\partial a_{\perp,j}(\mathbf{x}_\perp,z)}{\partial z} e^{i\delta(z-c t)} \right\} \frac{1}{r_{ij}} \times \delta(\mathbf{x}_\perp - \mathbf{x}_{\perp,j}(t)) \delta(z - z_j(t)) \tag{42}
\]
We average Equation (42) on the fast scale $\theta$, over a longitudinal dimension $l_b$ several wavelengths long $l_b = s \lambda_r$ (where $s$ is an integer number). In particular, we integrate both sides of Equation (42) by $1/l_b \int_R \chi_{[-s\pi,s\pi]} dz$ or similarly by $1/(l_b) \int_R \chi_{[-s\pi,s\pi]} d\theta$ where $\chi_{[a,b]}$ is equal to one within $[a,b]$ and zero otherwise; then the term on the right side of Equation (42) becomes:

$$\frac{1}{k_r l_b} \int_R dz \chi_{[-s\pi,s\pi]} \delta(z-z_j(t)) e^{-i\theta} = \frac{1}{l_b} \int_R d\theta \chi_{[-s\pi,s\pi]} \delta(\theta - \theta_j(t)) e^{-i\theta} = \frac{1}{l_b} \chi_{[-s\pi,s\pi]} e^{-i\theta}.$$

Then, defining the average over $N_e$ electrons of a generic function as:

$$\langle f(\theta, x_\perp) \rangle_\perp = \frac{\sum_{i=1}^{N_e} f(\theta_j, x_\perp)}{N_e} \langle \delta(x_\perp - x_\perp) \rangle_\perp$$

(43)

where $\Sigma$ is the transverse section of the electron beam and the total longitudinal electron density $n_\parallel = \chi_{[-s\pi,s\pi]} N_e / l_b$, we obtain:

$$\left[ D^+ + \frac{\nabla^2}{2ik_r} \right] a_r = \frac{k}{2} \left( \frac{\omega_p}{\omega_s} \right)^2 \left\{ \langle a_\parallel e^{-i\theta} \rangle_\perp - \langle \frac{a_\parallel}{\gamma} \rangle_\perp - \langle (dx_\perp/dz) e^{-i k_s z} \rangle_\perp \right\}$$

(44)

where $\omega_p \equiv \sqrt{e^2 n_e / e_0 m_e}$ is the plasma frequency and $n_e = n_\parallel / \Sigma$ is the total electron density. The first average in Equation (44) shows that the e.m. field evolution is ruled by the sum of all electron phases. As previously discussed, this average term is zero if the electron phases are homogeneously distributed. Conversely, if most of the electrons have the same phase, then the average is different from zero, and it may drive the radiation dynamics. The last two terms on the right side of Equation (44) are usually neglected; the second term is proportional to $a_r$, which is usually much smaller compared to $a_\parallel$; the third term is fast oscillating and can be neglected too. Then, the evolution equation for the self-consistent radiation field can be written as:

$$\left[ \frac{\partial}{\partial z} + \frac{1}{c} \frac{\partial}{\partial t} + \frac{\nabla^2}{2ik_r} \right] a_r = \frac{k}{2} \left( \frac{\omega_p}{\omega_s} \right)^2 a_\parallel \langle e^{-i\theta} \rangle_\perp$$

(45)

3.2.2. Space Charge Effects

Here we investigate the longitudinal micro-bunch space charge effects, i.e., the repulsion between neighbor electrons at the scale of the radiation wavelength. The macroscopic charge effects are not considerate because they become important only for very small e-beam energies. A full investigation of the space charge effect would be useful since the local repulsion between electrons could inhibit the establishment of the micro-bunching in the FEL process. However, a complete solution of the Equation (37) is difficult; rather, here we investigate the paraxial approximation of Equation (37), in which only the longitudinal micro-bunching space charge effects are taken into account, neglecting the transverse operator $\nabla_\perp$. Similarly to the evolution equation of the e.m. vector potential $a_r$, we assume the SVEA approximation in Equation (37), such that the evolution of the longitudinal electrostatic field in the paraxial approximation can be written as follows:

$$\frac{\partial e_z}{\partial z} = - \frac{e^2}{me_0 c^2} \frac{n_e(\theta)}{N_e} \sum_{j=1}^{N} \delta(\theta - \theta_j)$$

(46)

where $n_e(\theta)$ is the local electron density. Expanding $n_e(\theta)$ in a Fourier series with respect to $\theta$, we obtain

$$\frac{\partial e_z}{\partial z} = \frac{\omega_s^2}{c^2} \sum_{n=1}^{\infty} e^{i n \theta} \langle e^{-i n \theta} \rangle + \text{c.c.}$$

(47)
where we omitted the term \( n = 0 \) in the sum, since we neglect the macroscopic space-charge effect. Finally, the longitudinal electrostatic field is:

\[
e^z(\theta) = k_f \left( \frac{\omega_p}{\omega_r} \right)^2 \sum_{n=1}^{\infty} \frac{e^{in\theta}}{in} + c.c.
\]

(48)

Notice that the first harmonic term (i.e., \( n = 1 \)) is proportional to the bunching factor \( b = \langle e^{-i\theta} \rangle \). Clearly, the longitudinal space-charge force is appreciable only if the electron density is large enough. Similarly to the transverse e.m. field, the \( e_z \)-field is proportional to the sum over the electron phases. This fact shows that the space-charge effects are negligible also if the electron phases are homogeneously distributed. We observe also that if we sum the space-charge force over the electrons, the result is zero; this means that the space charge force is an internal force with respect to the electron system once that the macroscopic effects have been neglected. Then, the main contribution to the space-charge force comes from its first harmonic term, and we write [32]:

\[
e^z(\theta) = 2k_r \left( \frac{\omega_p}{\omega_r} \right)^2 \frac{1}{N_e} \sum_{j=1}^{N_e} \sin(\theta - \theta_j)
\]

(49)

Equation (45) and Equations (29)–(34) together with Equation (49) form a self-consistent system of equations for the 3D dynamics of \( N_e \) electrons interacting with radiation and laser undulator fields. The basic physics of FEL can be understood more clearly in the 1D scheme approximation. For this reason, before discussing the full 3D model, we review the one-dimensional theory and its results.

3.3. 1D FEL Model

The 1D approach is exhaustive for almost everything of the FELs physics. Moreover, at the same time it offers a very clear simple picture of the basic mechanism of the FEL process. In the previous section, a closed set of 3D equations representing the evolution of the whole system has been obtained. The one-dimensional approximation consists of neglecting any dependence on transverse spatial coordinates (Equations (29)–(32)), so that the dynamical variables, i.e., the electron phases \( \theta_j \), the electron energy \( \gamma_j \) and the dimensionless radiation field amplitude \( a_r \), depend only on \( z \) and \( t \). The electron dynamic Equations (33) and (34) in the 1D approximation become:

\[
c \frac{d\theta_j}{dz} = (k_r + k_w) - k_r c \frac{d\gamma_j}{dz}
\]

\[
c \frac{d\gamma_j}{dz} = \frac{a_L k_r}{2\gamma_j} \left( a_r e^{i\theta_j} + c.c. \right) + e^z(\theta_j).
\]

(51)

with \( j = 1, \ldots, N_e \). The 1D field evolution is obtained neglecting the transverse dependence in the radiation field \( a_r(z, x_\perp, t) \approx a_r(z, t) \) and assuming a uniform magnetic undulator \( a_L(z, x_\perp) = a_w \); then, by integrating both side of Equation (45) by \( 1/\Sigma \int_{\Sigma} dx_\perp \), where \( \Sigma \) is the transverse section of the beam, we obtain

\[
\left( \frac{\partial}{\partial z} + \frac{1}{c} \frac{\partial}{\partial t} \right) a_r = \frac{k}{2} \left( \frac{\omega_p}{\omega} \right)^2 a_w \left( e^{-i\theta} \right)
\]

(52)

where the plasma frequency is \( \omega_p \equiv \sqrt{e^2 n_e/\epsilon_0 m_e} \). The evolution equation for each electron phase can be obtained from Equation (50)

\[
c \frac{d\theta_j}{dz} = (k_r + k_w) - k_r c \frac{d\gamma_j}{dz}
\]

\[
c \frac{d\gamma_j}{dz} = \frac{a_L k_r}{2\gamma_j} \left( a_r e^{i\theta_j} + c.c. \right) + e^z(\theta_j).
\]

(53)
where $\gamma_r = \sqrt{k_r (1 + a_w^2)} / 2k_w$ is the resonant energy introduced in the previous section, (Equation (12)). We assume that during the interaction with the e.m. field, the energy of each electron remains close to the resonant energy, $\gamma \simeq \gamma_r$ (this assumption is known as the “Compton limit”), i.e.,

$$\frac{\gamma_j^2 - \gamma_r^2}{\gamma_j^2} \approx 2 \frac{\gamma_j - \gamma_r}{\gamma_r} \ll 1$$

(54)

With this assumption, we define a new energy-variable $\eta_j = \frac{\gamma_j - \gamma_r}{\gamma_r}$. The self-consistent equations for the electron phases and energies and the radiation field are:

$$\frac{d\theta_j}{dz} = 2k_w \eta_j$$

(55)

$$\frac{d\eta_j}{dz} = -\frac{a_w k}{2\gamma_r^2} [a_r e^{i\theta_j} + c.c.] - i \frac{k_r}{\gamma_r} \left( \frac{\omega_p}{\omega} \right)^2 \left[ \langle e^{-i\theta} \rangle e^{i\theta_j} - c.c. \right]$$

(56)

$$\left( \frac{\partial}{\partial z} + \frac{1}{c} \frac{\partial}{\partial t} \right) a_r = \frac{k_{aw}}{2\gamma_r} \left( \frac{\omega_p}{\omega} \right)^2 \langle e^{-i\theta} \rangle.$$  

(57)

3.4. Universal Scaling

One of the most useful features of the 1D FEL model is the possibility to introduce in Equations (55)–(57) dimensionless variables so that any physical parameters will not appear explicitly in the equations. This allows a general analysis of the FEL scaling laws and a simple interpretation of the basic physics of the process. We start defining the fundamental FEL-parameter [12]

$$\rho = \frac{1}{\gamma_r} \left( \frac{a_w \omega_p}{4ck_w} \right)^{2/3}.$$  

(58)

In terms of $\rho$, the system of coupled evolution Equations (55)–(57) can be set in a dimensionless form by introducing the following variables and parameters.

$$\bar{p}_j = \frac{\eta_j}{\rho}, \quad A_r = \frac{\omega_r}{\omega_p \sqrt{\rho \gamma_r}} a_r,$$

(59)

$$z = \frac{z}{l_g}, \quad \bar{t} = \frac{ct}{l_g}, \quad l_g = \frac{1}{2k_w \rho} = \frac{\lambda_w}{4\pi \rho}$$

(60)

where $l_g$ is the gain length. Note that in this universal scaling

$$\rho |A_r|^2 = \frac{\epsilon_0 |E_0|^2}{2mc^2\gamma_r n_e} = \frac{P_{rad}}{P_{beam}}$$

(61)

is the ratio between the e.m. and the electron power, i.e., is the FEL efficiency. Using this “universal scaling”, we obtain:

$$\frac{d\theta_j}{dz} = \bar{p}_j$$

(62)

$$\frac{d\bar{p}_j}{dz} = - \left[ A_r e^{i\theta_j} + c.c. \right] + \sigma \left[ \langle e^{-i\theta} \rangle e^{i\theta_j} - c.c. \right]$$

(63)

$$\left( \frac{\partial}{\partial z} + \frac{1}{c} \frac{\partial}{\partial \bar{t}} \right) A_r = \langle e^{-i\theta} \rangle$$

(64)

where $\sigma = 4\rho (1 + a_w^2) / a_w^2$ is the space-charge parameter. Note that $\sigma$ is proportional to $\rho$, so it can be neglected in the Compton limit $\rho \ll 1$ (this is because $\eta_j = \rho \bar{p}_j \ll 1$). As we have anticipated before, the 1D FEL equations in the Compton limit assume a dimensionless form, with the advantage of being solvable without specifying any operating parameters. Once solved, the scaling can be reversed to find the real physical quantities needed for a particular experimental set-up. The 1D FEL Equations (62)–(64) form a Maxwell-pendulum
model: in fact, writing the complex field as \( A_r = |A_r| \exp(i \phi_r) \) and neglecting the space-charge term, we obtain

\[
\frac{d^2 \theta_j}{d \bar{z}^2} = -2|A_r| \cos(\theta_j + \phi_r)
\]  

(65)

Of course, this equation greatly differs from an ordinary pendulum equation since amplitude and phase of the field are not constant, but their evolution is determined by Equation (64).

3.5. Steady State Regime

The set of Equations (62)–(64) can be written in a simpler way by a transformation of coordinates

\[
\bar{z} = z, \quad \bar{z}_1 = \frac{z - \langle v \parallel \rangle \bar{t}}{1 - \langle \beta \parallel \rangle}
\]  

(66)

The differential operators of the FEL equations change as follows:

\[
\frac{d}{d \bar{z}} \approx \frac{\partial}{\partial \bar{z}} + \frac{1}{\langle v \parallel \rangle} \frac{\partial}{\partial \bar{t}} \Rightarrow \frac{\partial}{\partial \bar{z}}
\]  

(67)

\[
\frac{\partial}{\partial \bar{z}} + \frac{1}{c} \frac{\partial}{\partial \bar{t}} \Rightarrow \frac{\partial}{\partial \bar{z}} + \frac{\partial}{\partial \bar{z}_1}
\]  

(68)

where in the first step of the Equation (67) we assumed \( \bar{z} \approx \langle v \parallel \rangle \bar{t} \) with (\( \langle v \parallel \rangle \) the longitudinal electron average velocity). Relation (68) shows that the time derivative can be neglected if the difference between the electrons velocity and the speed of light is unimportant \((1 - \langle \beta \parallel \rangle) \approx 0 \Rightarrow \langle v \parallel \rangle \approx c\)) or alternatively if the interaction time is small compared with the slippage time. This limit can be easily understood from Equation (68) by normalizing \( \bar{z} \) to the undulator length \( L_w \) and \( \bar{t} \) to the electron pulse duration \( \tau_b = l_b / \langle v \parallel \rangle \); the ratio between the time and space derivative coefficients is:

\[
\frac{L_w (1 - \langle \beta \parallel \rangle)}{l_b \langle \beta \parallel \rangle} = \frac{N_w \lambda_r}{\lambda} \frac{l_b}{l_s}
\]  

(69)

where \( l_s = N_w \lambda_r \) is the slippage length and we have used the resonant condition \( \lambda_r = \lambda_w (1 - \langle \beta \parallel \rangle) / \langle \beta \parallel \rangle \). If \( l_b \gg l_s \), the time derivative can be neglected (steady-state regime):

\[
\frac{d \theta_j}{d \bar{z}} = \bar{p}_j
\]  

(70)

\[
\frac{d \bar{p}_j}{d \bar{z}} = -\left[ A_r e^{i \theta_j} + c.c. \right]
\]  

(71)

\[
\frac{d A_r}{d \bar{z}} = \langle e^{-i \theta_j} \rangle.
\]  

(72)

3.5.1. Constants of Motion

The set of Equations (70)–(72) admits two constants of motion. The first is obtained by differentiating the average momentum

\[
\frac{d}{d \bar{z}} \left( \frac{1}{N_e} \sum_{j=1}^{N_e} \bar{p}_j \right) = -\frac{d}{d \bar{z}} |A_r|^2 \rightarrow \langle p \rangle + |A_r|^2 = C
\]  

(73)

where \( C \) is a constant. The meaning of this conservation law becomes clear if we return to physical quantities, using the relation (59) and (61):

\[
\langle p \rangle + |A_r|^2 = \frac{1}{N_e} \sum_{j=1}^{N_e} \Delta \gamma_j + \frac{e_0 |E_0|^2}{2mc^2 \gamma_j \rho e \rho} \rightarrow mc^2 \sum_{j=1}^{N_e} \gamma_j + \frac{e_0}{2} V |E_0|^2 = C'
\]  

(74)
where \( n_e = N_e / V \). The above relation describes the energy conservation law (the energy of the undulator has been supposed constant) and shows that the energy of the emitted radiation field is provided by the kinetic energy of the electron beam.

The second constant of motion is the total Hamiltonian of the system with \( 2N_e + 1 \) variables:

\[
H(\theta_j, p_j, \text{Re}(A_r), \text{Im}(A_r))_{j=1,N_e} = \sum_{j=1}^{N_e} \frac{p_j^2}{2} - i \sum_{j=1}^{N_e} [A_re^{i\theta_j} + \text{c.c.}] 
\]

(75)

This Hamiltonian can be written in a more compact form, in which the real and imaginary part of the radiation field are included as canonical variables \((\theta_0, \bar{p}_0)\). In fact, defining \( A_r = (\theta_0 + i\bar{p}_0) / \sqrt{2N_e} \), we obtain:

\[
H(\theta_j, \bar{p}_j)_{j=0,N_e} = \sum_{j=1}^{N_e} \frac{\bar{p}_j^2}{2} - \sqrt{2} \sum_{j=1}^{N_e} \bar{p}_j \sum_{j=1}^{N_e} \left[ \theta_0 \sin(\theta_j) + p_0 \cos(\theta_j) \right] 
\]

(76)

from which the equations of motion (70)–(72) follow as

\[
\frac{d\theta_j}{d\bar{z}} = \frac{\partial H}{\partial \bar{p}_j}, \quad \frac{d\bar{p}_j}{d\bar{z}} = -\frac{\partial H}{\partial \theta_j}, \quad \text{for } j = 0, N_e 
\]

(77)

3.6. Linear Analysis

In FELs at the undulator entrance \((\bar{z} = 0)\), the electron beam exhibits a narrow longitudinal energy distribution. For simplicity, we assume that each electron has the same initial energy \( \gamma_j(\bar{z} = 0) = \gamma_0 \), so that the dimensionless momentum \( \bar{p} \) at time \( \bar{z} = 0 \) is given by

\[
\bar{p}(0) = \frac{\gamma_0 - \gamma_r}{\rho \gamma_r} \equiv \delta 
\]

(78)

We define this value as the detuning parameter \( \delta \). It is particularly useful to redefine our variables so that the initial condition for \( \bar{p} \) is zero:

\[
\bar{p}_j' = \bar{p}_j - \delta \\
\theta_j' = \theta_j - \delta \bar{z} \\
A_r' = A_r e^{i\delta \bar{z}} 
\]

(79)

In this way, the detuning parameter appears explicitly in the equations which (dropping the primes) read:

\[
\frac{d\theta_j}{d\bar{z}} = \bar{p}_j 
\]

(80)

\[
\frac{d\bar{p}_j}{d\bar{z}} = -[A_r e^{i\theta_j} + \text{c.c.}] 
\]

(81)

\[
\frac{dA_r}{d\bar{z}} = \langle e^{-i\theta} \rangle + i\delta A_r 
\]

(82)

This set of equations can be linearized in terms of three collective variables \([12]\)

\[
\mathcal{A} = A_r, \quad \text{field amplitude} \\
\mathcal{B} = \langle e^{-i\theta} \rangle, \quad \text{bunching} \\
\mathcal{P} = \langle \bar{p} e^{-i\theta} \rangle, \quad \text{momentum bunching} 
\]

(83–85)
Neglecting the second-order terms $\langle e^{-i\theta} \rangle$ and $\langle \bar{p}^2 e^{-i\theta} \rangle$, we obtain a closed set of linear equations

\[
\frac{dB}{dz} = -iP \quad (86)
\]
\[
\frac{dP}{dz} = -A \quad (87)
\]
\[
\frac{dA}{dz} = B + i\delta A \quad (88)
\]

which can be reduced to

\[
\frac{d^3A}{dz^3} - i\delta \frac{d^2A}{dz^2} - iA = 0. \quad (89)
\]

Assuming a solution of the form $A \propto e^{i\lambda z}$, one obtains the following dispersion relation which rules the stability of the system:

\[
\lambda^3 - \delta \lambda^2 + 1 = 0. \quad (90)
\]

When Equation (90) has three real roots, then the system is stable, but if it has one real root and two complex-conjugate ones, then one of the latter will cause an exponential growth of the field until non-linear effects come into play. If instead of choosing an ideal cold beam, we assume an initial energy distribution $f(p_0)$ with a finite width, then (90) generalize into

\[
\lambda - \delta + \int_{-\infty}^{+\infty} f(p_0) \frac{d}{dp_0} \lambda \left( \lambda^2 - \mu^2 \right) = 0. \quad (91)
\]

The above integral can be analytically solved, for instance in the case of a rectangular or Lorenz distribution [16]. For the rectangular case with half-width $\delta\gamma$, we obtain:

\[
(\lambda - \delta)(\lambda^2 - \mu^2) + 1 = 0 \quad (92)
\]

where $\mu \equiv \delta\gamma / \rho\gamma_r$ is the energy spread parameter.

As the exponential behavior of $A_r(z)$ is determined by the imaginary part of the complex root of Equation (92), it is relevant to plot it as a function of the detuning parameter $\delta$ for different values of the energy spread $\mu$. Figure 3 suggests some immediate considerations:

1. given a spread $\mu$, the optimal gain occurs for the specific detuning shift;
2. energy spread ($\mu > 0$) lowers the growth rate, and shift the resonance to $\delta = \mu$;
3. the width of the gain curve shrinks as $1/\sqrt{\mu}$.

Coming back to physical variables, this means that in order to preserve the exponential gain, for $\mu = 0$, the e-beam must satisfy:

\[
\frac{\delta\gamma}{\gamma_r} \leq \rho \quad (93)
\]

The solution for the cold beam case ($\mu = 0$) at resonance ($\delta = 0$) is:

\[
\lambda^3 = -1 \Rightarrow \lambda_1 = 1 , \lambda_2 = \frac{-1 + \sqrt{3}i}{2} , \lambda_3 = \frac{-1 - \sqrt{3}i}{2} \quad (94)
\]

and the scaled field is

\[
A(z) = \frac{1}{3} \sum_{k=1}^{3} \left( A(0) - iB(0) \frac{1}{\lambda_k} + iP(0) \right) e^{i\lambda_k z} \quad (95)
\]

For $z \gg 1$ (exponential growth regime) the growing mode $\lambda_3$ dominates over the oscillatory $\lambda_1$ and decaying $\lambda_2$ modes, so that
\[ \mathcal{A}(z) \simeq \frac{1}{3} \left( A(0) - i \frac{B(0)}{\lambda^3} + i \frac{P(0)}{\lambda^3} \right) e^{i\lambda z} \] (96)

The first term in the bracket corresponds to the amplification of an external input signal, while the second and the third terms correspond to some initial bunching. As it results from the linear analysis, an exponential instability of the emitted radiation field takes place in the FEL, until non-linear effects saturate this growth. The linear solution of the exponential growing mode fits well the numerical solution of the full non-linear system (70)–(72).

In Figure 3, \(|\text{Im}\lambda|\) vs. \(\delta\) and different values of \(\mu\): (a) 0, (b) 0.5, (c) 3, (d) 5, (e) 7 and (f) 10.

In Figure 4, the radiated power is plotted, as obtained from the numerical integration of the 1D non-linear equations. The figure shows that after an initial lethargy, the field power grows exponentially, reaching a saturation after a certain number of gain lengths (depending on the initial conditions). After saturation, the amplification process is replaced by an oscillatory energy exchange between the electrons and the radiation field. In the case of a seed signal intensity \(|A_0|^2\), the asymptotic linear solution is given by the

\[ |A|^2 \simeq \frac{1}{9} |A_0|^2 \exp \left[ \sqrt{\frac{2}{l_g}} \right], \] (97)

so that the \(l_g\) as defined in Equation (60) corresponds effectively to the FEL gain length. Since saturation occurs at \(|A|^2 \simeq 1\), from the Equation (97) we can estimate the saturation length as

\[ z_{sat} \simeq \frac{l_g}{\sqrt{3}} \ln \left( \frac{9}{|A_0|^2} \right) \propto \frac{\lambda_w}{\rho}, \] (98)

Furthermore, the result \(|A|^2 \simeq 1\) means that the saturation value is independent of the initial conditions. Since we know that \(|A|^2\) is proportional to \(|E_0|^2 / \rho n_e\), and \(\rho \propto n^{1/3}\) then

\[ |A|^2 \propto \frac{|E_0|^2}{n_e^{4/3}} \Rightarrow |E_0|^2 \propto n_e^{4/3} \] (99)

i.e., the intensity is proportional to \(n_e^{4/3}\) instead of \(n_e\): this implies the existence of a collective behavior in the electron beam. The independence of the saturated field amplitude
on its initial value gives us another important information, related again to the importance of the $\rho$ parameter. Since in Equation (61) we have defined the efficiency $\eta$ of the FEL as $\eta = \rho |A|^2$, then the measure of FEL efficiency is given by $\rho$. In this section we have shown that in the context of the 1D theory, $\rho$ is the only parameter which determines the constraints on the FEL:

$$\begin{align*}
\Delta \gamma / \gamma & \approx \rho & \text{Energy spread} \\
\eta &= P_{\text{rad}} / P_{\text{beam}} \approx \rho & \text{Efficiency} \\
l_g &\propto 1 / \rho & \text{Gain length}
\end{align*}$$

(100)

3.7. Superradiant Regime

The steady-state regime that we just described is based on the assumption that the slippage is negligible ($l_b \ll l_s$, see Equation (69)): the undulator is not long enough to appreciate the difference in velocity between the electrons and the radiation, so that all sections of the electron beam evolve almost identically and the peak power of radiation scale as $n^{4/3}$ (see Equation (99)). When slippage is taken into account, the FEL can operate in a different regime of cooperative emission, the superradiant regime [14–16,33,34], where the peak power scale as $n^2$. The slippage modifies substantially the interaction process between the radiation and the electrons: in fact, the radiation propagates with respect to the electrons, interacting with different slices of the electron beam, such that there is a region, near the trailing edge of the electron pulse and of length $l_s$, where the electrons emit radiation without being affected by the radiation produced by the other electrons behind them. Let us introduce a new characteristic length which is useful for the analysis of the propagation effect, the cooperation length,

$$l_c = \frac{\lambda_r}{4\pi \rho}.$$ 

(101)

Its meaning can be better understood using the resonant relation (5), which allows to write it in terms of the relativistic parameter $\beta ||$:

$$l_c = l_g \frac{1 - \langle \beta || \rangle}{\langle \beta || \rangle}.$$ 

(102)

so that it can be interpreted as the slippage in a gain length $l_g$. We define the electron pulse to be long or short with respect to the cooperation length. In general, if the electron pulse
is long enough \((l_b \gg l_c)\), the superradiant and the steady state instability will be present together: superradiance occurs in the region near the trailing edge of the electron pulse (slippage region) while the steady-state emission occurs in the remaining part of the beam: this is named “strong superradiance” because the peak power is greater than that in the steady-state regime.

Figure 5 shows the results of a numerical simulation integrating the 1D FEL equations with slippage:

\[
\frac{\partial \theta_j}{\partial \bar{z}} = p_j \\
\frac{\partial p_j}{\partial \bar{z}} = -\left[ A_r e^{i\theta_j} + c.c. \right] \\
\left( \frac{\partial}{\partial \bar{z}} + \frac{\partial}{\partial z_1} \right) A_r = \langle e^{-i\theta} \rangle + i \delta A_r
\]

where \(\bar{z} = z/l_b\) and \(z_1 = (z - c(\beta_\parallel) t)/l_c\). In Figure 5, the electron beam length is \(l_b = 30 l_c\), with a flat current profile in the interval \(0 < z_1 < 30\) and the detuning \(\delta = 0\). We observe a steady-state intensity in \(\bar{z} < z_1 < 30\) and a superradiant peak growing near the trailing edge \((z_1 = 0)\), in the slippage region \(0 < z_1 < \bar{z}\) (see Figure 5b). The intensity observed in the region \(30 < z_1 < 30 + \bar{z}\) is the trace of the radiation propagating forward, in front to the electron leading edge \(z_1 = 30\). The peak of the superradiant pulse in the slippage region strongly exceeds the maximum value of the steady-state intensity, \(|A_r| \sim 1.4\).

Figure 5. Strong superradiance, with \(l_b = 30 l_c\), \(\delta = 0\) and different values of \(\bar{z}\): (a) \(\bar{z} = 8\); (b) \(\bar{z} = 17\); (c) \(\bar{z} = 34\).
In the short-bunch regime \( l_b \ll l_c \) the radiation emitted by electrons escapes from the electron pulse in a length shorter than the gain length, so that the steady-state saturation never occurs: this regime is called “weak superradiance” because the peak power is lower than that in the steady-state regime.

Figure 6 shows the results of a simulation with \( l_b = 0.1 l_c, \delta = 0 \) and different values of \( \bar{z} \). In this case, almost all the radiation intensity is outside of the electron pulse, propagated from the trailing edge \( z_1 = 0.1 \), and no steady-state emission occurs. Weak superradiance in high-gain single-pass FELs has been observed experimentally in the past years by several groups \[19,35\].

![Figure 6. Weak superradiance, with \( l_b = 0.1 l_c, \delta = 0 \) and different values of \( \bar{z} \): (a) \( \bar{z} = 50 \); (b) \( \bar{z} = 75 \); (c) \( \bar{z} = 125 \).](image)

It is possible also to observe pure superradiance by tuning the system out of resonance: in fact, while the steady-state regime needs resonance (i.e., \( \delta = 0 \)) to produce exponential gain, superradiance weakly depends on the detuning, since it has a much larger gain bandwidth. Therefore, when the system is detuned in such a way to prevent steady-state radiation, the superradiant instability travels forward over unperturbed electrons, extracting energy from them with an even greater efficiency than in the steady-state regime.
This case is shown in Figure 6, with the same parameters as Figure 5 but with a detuning $\delta = 2$, such that the steady-state emission is inhibited. In this case the superradiant pulse, born near the trailing edge of the electron beam, travels unperturbed over the electrons toward the leading edge, with a peak intensity growing as $z^2$. It has been demonstrated [15,34,36] that the superradiant pulse, as clearly observed in Figures 6 and 7, is described by a self-similar solution of Equations (103)–(105). It is obtained by setting $\theta_j(z,z_1) = \theta_{1j}(y)$, $p_j(z,z_1) = p_{1j}(y) / \sqrt{z_1}$ and $A_j(z,z_1) = z_1 A_1(y)$, where $y = \sqrt{z_1}(\bar{z} - z_1)$ and $\theta_{1j}(y)$, $p_{1j}(y)$ and $A_1(y)$ are the solutions of the following ordinary differential equations:

\[
\frac{d\theta_{1j}}{dy} = p_{1j} \tag{106}
\]
\[
\frac{dp_{1j}}{dy} = -\left[A_1 e^{i\theta_{1j}} + c.c.\right] \tag{107}
\]
\[
\frac{y}{2} \frac{dA_1}{dy} + A_1 = \langle e^{-i\theta_1} \rangle \tag{108}
\]

Figure 8 shows $|A_1|^2$ as it results from the numerical solution of Equations (106)–(108).
Figure 8. Superradiant self-similar solution $|A_1|^2$ vs. $y = \sqrt{z_1}(z - z_1)$.

From the definition of the dimensionless radiation amplitude, the power of the pulse is
$$P_r = \rho P_{\text{beam}} |A_1|^2 \propto \rho^3 P_{\text{beam}} \propto N_e^2,$$
where $N_e$ is the number of electrons, i.e., superradiant, and its width decreases as $1/\sqrt{N_e}$. In the weak superradiance, $z_1 = l_b/l_c$ and $y \approx \sqrt{l_b/l_c} z$, so that $|A_1|^2 \sim (l_b/l_c)^2 A_1(\sqrt{l_b/l_c} z)^2$ (see Figure 6). The peak power is proportional to $l_b^2$ and the pulse width is proportional to $1/\sqrt{l_b}$. In the strong superradiance, the peak pulse grows as $z^2$ and shrinks as $1/\sqrt{z}$ when it propagates over the electron beam (see Figure 7).

3.8. SASE Operation

The self-amplified spontaneous emission (SASE) operation for an FEL is made up of three basic ingredients [20]:

1. high gain instability;
2. propagation effects, i.e., "slippage";
3. start-up from noise.

The first experimental observation of the high-gain regime, also starting from noise, was carried out in the microwave range using a wave guide in the Livermore experiment [37]. Presently, short wavelength FEL, which amplify incoherent shot-noise via SASE are of greater interest worldwide as a source of ultra-bright coherent X-ray radiation. SASE FEL in the X-ray region has been demonstrated firstly in LCLS project facility [38] at Standford, CA, USA, in the European X-ray FEL [39] at Hamburg, Germany and in Riken, Japan [40]. Actually, several X-ray FELs operating in the SASE regime have been and will be realized over the world [5]. The most important features of the SASE FEL operation are determined by the following characteristic lengths [20]:

$$\begin{align*}
  l_b & \quad \text{bunch length} \\
  l_c & \quad \text{cooperation length} \\
  l_g & \quad \text{gain length}
\end{align*}$$

An important parameter determining the evolution of the system is given by the ratio between $l_b$ and $l_c$ (the number of cooperation lengths in a bunch length). If the ratio $l_b/l_c > 2\pi$ (long bunch case), then the radiation pulse contains many spikes, each one having a maximum duration corresponding to about $2\pi l_c$ with a width of order of $1/l_b$ and large intensity fluctuations (see Figure 9).
Figure 9. SASE mode operation: typical temporal pattern of the intensity $|A_r|^2$ vs. $z_1$ for a long beam with $l_b = 50 l_c$, at different undulator lengths $\bar{z}$: (a) $\bar{z} = 0$; (b) $\bar{z} = 50$; (c) $\bar{z} = 100$.

On the contrary, if the ratio $l_b/l_c \leq 2\pi$ (short bunch case), only a single radiation pulse is present in this case, with no inner spikes. In both cases, superradiance occurs for sufficiently long undulators. The final result is an almost chaotic temporal pulse structure, with a broad spectral width and with a number of spikes of the order of the number of cooperation lengths in the electron bunch (see Figure 10 for the power spectrum of the intensity profile shown in Figure 9 vs. $\bar{\omega} = (\omega - \omega_s)/2\rho\omega_s$, where $\omega_s = c k w \gamma^2 / (1 + \alpha_w^2)$ is the spontaneous frequency).

Finally, in order to model correctly this spiking behavior, we must take into account not only the slippage between radiation electron pulse, but also the finite bunch length; the electron bunch behaves as a $l_b/l_c$ statistical independent zones, each of them giving rise to a superradiant spike which grows and narrows extracting energy from electrons within a cooperation length [20,41].

In the “so-called” quantum SASE regime [24], a completely different behavior occurs, the “classical” random spiking behavior almost disappears and a strong narrowing of the spectrum occurs. This new phenomenon is called “quantum purification”, also if its experimental evidence is yet to come.
Figure 10. SASE mode operation: power spectrum of the intensity shown in Figure 9 vs. $\tilde{\omega} = \frac{(\omega - \omega_s)}{2\rho\omega_s}$; (a) $\tilde{z} = 0$; (b) $\tilde{z} = 50$; (c) $\tilde{z} = 100$.

4. From 1D to 3D

The description of an $N_e$ electron beam can be made introducing, in the transverse phase space $(x_{\perp}, \eta_{\perp})$ (where $\eta_{\perp} = \frac{dx_{\perp}}{dz}$ and $z$ is the longitudinal coordinate along the beam direction), the second-order moments, as for instance the rms beam size $\sigma^2_x = \langle |x_{\perp}|^2 \rangle$, the rms beam angular divergence $\sigma^2_{\eta_x} = \langle |\eta_{\perp}|^2 \rangle$ and the mixed terms $\langle x_{\perp} \cdot \eta_{\perp} \rangle$. The rms beam size (for example in $x$ direction) evolves in the free space as:

$$\sigma_x(z) = \sqrt{\epsilon_x \left( \beta^+_x + \frac{z^2}{\beta^*_x} \right)}$$

(110)

where $\beta^+_x = \langle x^2 \rangle / \epsilon_x$ is one of the Twiss parameters and $\epsilon_x = \sqrt{\langle x^2 \rangle \langle \eta^2_x \rangle - \langle x \eta_x \rangle^2}$ is the rms emittance along $x$, which is conserved in the free space evolution and for a linear transport system; $z = 0$ corresponds to the point where $\langle x \eta_x \rangle = 0$ and the beam size is minimum ("beam waist"). The radiation beam can be described in a similar way: in fact,
using the free space diffraction formula, it is possible to show that the size of a Gaussian beam evolves as:

\[ \sigma_r(z) = \sqrt{\frac{\lambda_r}{4\pi} \left( Z_r + \frac{\sigma^2_r}{Z_r} \right)} \]  

where \( Z_r = \frac{4\pi \sigma^2_r}{\lambda_r} \) is the Rayleigh range. The evolution equations of the electron and radiation beam sizes are formally equivalent if we identify \( \epsilon_{rad} = \lambda_r / 4\pi \) as the rms emittance of the Gaussian radiation beam and \( \beta^* \) as the Rayleigh range of the electron beam. Furthermore, the free space evolution of the counter-propagating laser beam size \( \sigma_L(z) \) follows the same relation of Equation (111) where instead of \( Z_r \) we substitute \( Z_L = \frac{4\pi \sigma^2_L}{\lambda_L} \) where \( \lambda_L \) is the laser beam wavelength.

4.1. Transverse Effects

From this picture, we can set some reasonable condition that must be satisfied in a 3D geometry:

1. The matching between electron and radiation beam requires that the beam waist and the Rayleigh range of each other must be comparable:

\[ \sigma_e \simeq \sigma_r \]  

\[ \beta^* \simeq Z_r \rightarrow k_r \epsilon_r \simeq \frac{1}{2} \left( \frac{\sigma_e}{\sigma_r} \right)^2 \]  

2. The electron beam should be contained in the laser beam and the electron beam should not diverge appreciably in a Rayleigh range \( Z_L \):

\[ \sigma_e \leq \sigma_L \]  

\[ \beta^* \leq Z_L \rightarrow k_L \epsilon_r \leq \frac{1}{2} \left( \frac{\sigma_e}{\sigma_L} \right)^2 \]  

The condition (113) is known as the “Pellegrini criterium” [42] and it can be relaxed in particular conditions, depending on the ratio between the electron and the radiation beam size (112). Conditions (114) and (115) are more stringent and concern the matching between the e-beam and the laser undulator profile. If conditions (114) and (115) are violated, inhibition of the gain process occurs. If we use a magnetic undulator instead of the laser undulator, then the homogeneous condition (115) can be written as \( k_w \epsilon_r < 1 \), which is usually satisfied. In fact, for an undulator period of the order of \( \lambda_w = 1 \) cm and emittance \( \epsilon_r \simeq 10^{-7} \) m-rad, \( k_w \epsilon_r \simeq 10^{-4} \). An FEL experiment operating with a laser undulator does not require a focusing system for the electron beam, since the interaction between electrons and laser beam occurs in free space near the beam waist zone, i.e., the interaction length \( L_{int} \) is of the order of few laser Rayleigh ranges \( (L_{int} \simeq Z_L) \). Hence, the gain process occurs if the condition (115) is satisfied. Furthermore, if a magneto-static undulator instead of a laser undulator is used, then the interaction length is usually more than one meter, and the electron beam must be conveniently focused to keep a nearly constant beam size. A linear magneto-static undulator provides a “natural” focusing effect in one transverse direction and needs external focusing (by quadrupole magnets) on the other transverse direction. This focusing force induces an oscillation of the electron motion with a wavelength much longer than the undulator period \( \lambda_w \). For an electron beam matched in two planes, the beam size (averaged over the undulator period \( \lambda_w \)) remains constant along the undulator, while individual electrons perform a periodic transverse motion, called “betatron motion”.

A potential deleterious effect for an FEL is the apparent energy spread induced by the betatron motion, i.e., by the electron transverse velocities. The relation between energy spread and transverse velocity can be made more transparent taking into account the
transverse dimension in the resonant condition; it straightforward to demonstrate that an observer far away in the \( \eta \perp \) direction from the axis sees a slightly different period, i.e.,

\[
\lambda_{\perp} = \frac{\lambda_{w}}{2\gamma_{r}^{2}} \left( 1 + a_{w}^{2} + \eta_{\perp}^{2} \gamma_{r}^{2} \right)
\]

(116)

where \( \lambda_{w} \) is replaced by \( \lambda_{L}/2 \) for a laser undulator. The above relation agrees with the resonance condition (11) when the observer angle tends to zero. From the transverse resonant relation (116), using the relation

\[
\Delta \gamma \approx \frac{\Delta \lambda}{2\lambda}
\]

(117)

we can argue that the energy spread normalized to the FEL parameter \( \rho \) has the following contributions:

1. different longitudinal momentum distribution, (see Equation (93))

\[
\frac{\Delta \gamma}{\rho \gamma_{r}} \left|_{1D} \right. \approx 1
\]

(117)

2. off-axis variation of the undulator parameter

\[
\frac{\Delta \gamma}{\rho \gamma_{r}} \approx \frac{1}{2\rho} \frac{\Delta a_{w}^{2}}{1 + a_{w}^{2}}
\]

(118)

3. angular divergence of the beam

\[
\frac{\Delta \gamma}{\rho \gamma_{r}} \left|_{\perp} \right. \approx \frac{1}{2\rho} \frac{\eta_{\perp}^{2} \gamma_{r}^{2}}{1 + a_{w}^{2}}
\]

(119)

These effects are called “non homogeneous effects” and give rise to a broadening of the resonant condition.

4.2. Full 3D Model

A most useful scheme for modeling a beam of \( N_{e} \) electrons interacting with e.m. field is the Maxwell–Vlasov scheme. Following this scheme, we introduce the transverse phase space variables \((x_{\perp}, \eta_{\perp})\) and the longitudinal phase space variables \((\theta, \bar{p})\). The transverse phase-space distribution, for a given phase and longitudinal momentum, can be described by a distribution function \( F(x_{\perp}, \eta_{\perp}) \) with a non-negative value equal to the number of electrons per unit area at the transverse phase-space point \((x_{\perp}, \eta_{\perp})\). A Gaussian distribution function is normally adopted for its simplicity. Hence, the number of electrons within a \( dx_{\perp} d\eta_{\perp} \) at the beam waist (such that \( \langle x_{\perp} \cdot \eta_{\perp} \rangle = 0 \)) is

\[
F(x_{\perp}, \eta_{\perp}) \ dx_{\perp} d\eta_{\perp} = \frac{N_{e}}{2\pi \epsilon_{r}} \exp \left( -\frac{x_{\perp}^{2}}{2\sigma^{2}} - \frac{\eta_{\perp}^{2}}{2\sigma^{2}} \right) \ dx_{\perp} d\eta_{\perp}
\]

(120)

The evolution equation of the electron beam distribution function can be obtained using the Liouville theorem \( \partial f/\partial z = \{H, f\} \) where \( \{H, f\} \) is the Poisson bracket. In the previous sections, we have obtained a three dimensional Hamiltonian with its associated equations of motion, Equations (29)–(34), and an evolution equation for the radiated e.m. field (45). Now, we introduce the “3D Universal Scaling”:

\[
\theta = (k_{r} + k_{L}) z - c(k_{r} - k_{L}) t, \quad z = \frac{z}{\tau_{g}}, \quad \bar{x}_{\perp} = \frac{x_{\perp}}{\sigma},
\]

\[
\rho = \frac{\gamma - \gamma_{r}}{\rho \gamma_{r}}, \quad \bar{p}_{\perp} = \frac{\sigma}{\epsilon_{r}} \eta_{\perp}, \quad \bar{\eta}_{\perp} = \frac{d\bar{x}_{\perp}}{dz}, \quad a_{L} = a_{0} \delta(z, x_{\perp})
\]

(121)
where \( a_0 = a_{0r} \), \( g(z, \mathbf{x}) \) is the transverse laser profile, \( \sigma \) is the rms electron beam radius at the beam waist and \( \epsilon_r = \sigma \sigma' \) is the rms beam emittance. With this universal scaling the equations of motion (29)–(34) can be written:

\[
\begin{cases}
\frac{dx}{dz} = b \ (1 - \rho \bar{p}) \bar{p}_\perp \\
\frac{dp}{dz} = -\frac{2\rho}{X} (1 - \rho \bar{p}) \nabla_{\perp} \left[ \frac{\bar{\gamma}}{2\rho} |g|^2 - i \left( g^* A e^{i\theta} - c.c. \right) \right] \\
\frac{d\theta}{dz} = \rho + \left[ \frac{\bar{\gamma}}{2\rho} (1 - |g|^2) - \frac{bX}{4} \bar{p}_\perp^2 \right] \\
\frac{d\bar{p}}{dz} = -(1 - \rho \bar{p}) \left( g^* A e^{i\theta} + c.c. \right) + \mathcal{E}.
\end{cases}
\] (122)

Whereas the 1D theory is ruled only by the FEL parameter \( \rho \), instead, in the 3D theory, we must introduce new parameters which characterize the transversal dynamics:

\[
X = 2k_r \epsilon_r, \quad b = \frac{l_g}{\bar{\beta}_c} = \frac{l_g \epsilon_r}{\sigma^2}, \quad \bar{\gamma} = \frac{a_0^2}{1 + a_0^2}.
\] (123)

In the Compton limit, the terms proportional to \( \rho \) (as for instance in \( (1 - \rho \bar{p}) \simeq 1 \)) can be neglected, since the value of \( \rho \) ranges from about \( 10^{-5} \) to \( 10^{-3} \). With this simplification, Equation (122) becomes:

\[
\begin{cases}
\frac{dx}{dz} = b \ \bar{p}_\perp \\
\frac{dp}{dz} = -\frac{\bar{\gamma}}{\rho X} \nabla_{\perp} |g|^2 \\
\frac{d\theta}{dz} = \rho + \left[ \frac{\bar{\gamma}}{2\rho} (1 - |g|^2) - \frac{bX}{4} \bar{p}_\perp^2 \right] \\
\frac{d\bar{p}}{dz} = -(g^* A_r e^{i\theta} + c.c.) + \mathcal{E}.
\end{cases}
\] (124)

They can be obtained by the following Hamiltonian

\[
\bar{H} = \frac{b^2}{2} + b \left( \frac{\bar{\gamma}}{2\rho} (1 - |g|^2) - \frac{bX}{4} \bar{p}_\perp^2 \right) + \frac{\bar{\gamma}}{\rho X} |g|^2 - i \left( g^* A_r e^{i\theta} - c.c. \right) + \Phi.
\]

where \( \Phi = \gamma_r / (\rho^2 (1 + a_0^2)) \Phi \) is such that \( \mathcal{E} = -\partial \Phi / \partial \theta \). Now, using the Liouville theorem, the Hamiltonian (125), the Maxwell Equation (45) and 3D universal scaling (121), we obtain the following Maxwell–Vlasov system of equations for an e-beam interacting with a laser undulator and a radiation field:
Here $a = l_Z/\gamma = b/X$ and $Z_r = 4\pi \sigma^2/\lambda_r$ are respectively the diffraction parameter and the Rayleigh range of the emitted radiation, with a transverse radius equal to the electron beam radius. The distribution function $f$ and the radiation field $A_r$ have the following dependence $f = f(\theta, \bar{p}, \bar{x}_\perp, p_\perp, z, \bar{z})$, $A_r = A_r(\theta, \bar{x}_\perp, z, \bar{z})$, and $g = g(\bar{x}_\perp, \bar{z})$ is the laser undulator profile. The term $b \bar{p}_\perp \nabla_{\bar{x}_\perp}$ corresponds, with the unscaled variables, to $\eta_\perp \nabla_{\bar{x}_\perp}$ and describes the transverse drift of the beam, responsible for instance of the beam size increase away from the beam waist; the comparison of the two diffraction terms $a \simeq b$ (which corresponds to Equation (113)) rules the matching condition between the electron and radiation beam. In fact, assuming the same initial spot size, if $b < a$ the electron beam is contained into the radiation beam, which corresponds to the “Pellegrini condition” $X < 1$ [42], otherwise the radiation is confined into the electron beam size. The phase term of Equation (125) contains three terms which, with the unscaled variables, can be written as:

$$
\begin{align*}
\dot{\bar{p}} &= \frac{\gamma - \gamma_r}{\rho^2 \gamma_r} = \left( \frac{\Delta \gamma}{\rho \gamma_r} \right)_{1D} \simeq 1 \\
\frac{\bar{\xi}}{2 \rho} (1 - |g|^2) &\simeq \frac{1}{2 \rho} \frac{\Delta a_{\perp}^2}{1 + a_{\perp}^2} \simeq \left( \frac{\Delta \gamma}{\rho \gamma_r} \right)_{u_w} \\
\frac{bX}{4} \mathbf{p}_\perp^2 &\simeq \frac{1}{2p} \frac{\theta^2 \gamma_r^2}{1 + a_{\perp}^2} \simeq \left( \frac{\Delta \gamma}{\rho \gamma_r} \right)_{\perp}
\end{align*}
$$

and are responsible of the “non homogeneous effects” (117)–(119). Finally, the last term $d \mathbf{p}_\perp / dz = \dot{\mathbf{p}} / (2\rho X) \nabla_{\bar{x}_\perp} |g|^2$ corresponds with unscaled variables to a focusing force due to the laser undulator profile $\dot{\bar{\eta}}_\perp = -a_{\perp}^2 / (2\gamma_0)^2 \nabla_{\bar{x}_\perp} |g|^2$. Notice that $\mathbf{p}_\perp \cdot \nabla_{\bar{x}_\perp} = \rho \partial / \partial x + \bar{p}_y \partial / \partial \bar{y}$ and the average in the field evolution equation is replaced by the average over an ensemble, $\langle O \rangle = \int O f(\xi) d\xi$ with $\int f(\xi) d\xi = 1$. This closed set of equations can be solved numerically choosing a Gaussian initial condition at $f(\bar{z} = 0) = f_0$:

$$
\begin{equation}
\begin{aligned}
f_0(\theta, \bar{p}, \bar{x}_\perp, p_\perp) &\propto \exp \left\{ - \frac{|\bar{x}_\perp + b \bar{z}_0 p_\perp|^2}{2} - \frac{\mathbf{p}_\perp^2}{2} \right\} \\
\end{aligned}
\end{equation}
$$

where $\bar{z}_0$ is the waist position. Typically, a Gaussian phase-space profile corresponds to a thermal distribution.

4.2.1. Optical Guiding

One of the most peculiar results of the high-gain FEL theory is the presence of optical guiding, i.e., the existence of exponentially growing modes which have a profile independent of the longitudinal coordinate [43,44]. This is particularly important in the case of long undulators, where the radiation tends to affected by diffraction. Neglecting emittance, energy spread, betatron oscillation and slippage, the linear FEL equations taking into account diffraction in the paraxial approximations are obtained by combining Equations (86)–(88) and (125) to give

$$
\begin{align*}
\frac{\partial^2 B}{\partial z^2} &= iA \\
\left\{ \frac{\partial}{\partial z} - i\sigma \nabla_\perp^2 \right\} A &= u(\bar{r}) B + i\delta A
\end{align*}
$$

(127)

where $A$ and $B$ depend on the transverse position $\bar{r}$ and $u(\bar{r})$ is the normalized transverse electron distribution. Looking for solutions which depend exponentially on $z$ as $\exp(i\lambda z)$ and are azimuthally symmetric, Equation (127) becomes

$$
\begin{align*}
\left\{ (\lambda - \delta) A - \alpha \frac{1}{\bar{r}} \frac{d}{d\bar{r}} \left( \bar{r} \frac{dA}{d\bar{r}} \right) \right\} &= -\frac{u(\bar{r})}{\lambda^2} A
\end{align*}
$$

(128)
Outside the electron distribution ($u = 0$), the solution (normalized to unity) is

$$\mathcal{A} = H_0(\phi \bar{r}) \quad (129)$$

where $H_0$ is the Hankel function of zero order of the first kind, and $\phi^2 = (\delta - \lambda)/a$. For large $\bar{r}$, the asymptotic limit is

$$\mathcal{A} \approx \left( \frac{2}{\pi \phi} \right)^{1/2} \exp(i\phi \bar{r} - i\pi/4). \quad (130)$$

In order that $\mathcal{A} \to 0$ for $\bar{r} \to \infty$, we require that $\text{Im}\phi > 0$. Since we want an exponentially growing mode, then $\text{Im}\lambda < 0$ and $\text{Im}\phi^2 > 0$, so that it must be also $\text{Re}\phi > 0$. Considering an uniform sharp-edge distribution, with $u = 1/\pi$ for $\bar{r} < 1$ and $u = 0$ for $\bar{r} > 0$, the mode amplitude satisfying Equation (128) inside the electron beam and normalized to unity is

$$\mathcal{A} = J_0(\chi \bar{r}) \quad (131)$$

where $J_0$ is the Bessel function of zero order of the first kind and

$$\chi^2 = \frac{1}{a} \left( \delta - \lambda - \frac{1}{\lambda^2} \right) = \phi^2 - \frac{1}{a^2(\phi^2 - \delta)^2} \quad (132)$$

looking for values of $\phi$ and $\chi$ which are simultaneous solutions of Equations (132) and (133). A detailed numerical study has been done in [43], showing that the condition $a = 1$ (i.e., $Z_r \gg l_\phi$) measures the distance over which diffraction establishes transverse coherence in the high-gain FEL. If $Z_r \gg l_\phi$, one-dimensional theory gives the correct gain, but there is not full transverse coherence. If $Z_r < l_\phi$, the fundamental transverse mode has gain substantially larger than that of the other modes, the laser mode area is much larger than the electron beam area and a full transverse coherence is established.

### 4.3. Quantum Regime of FEL

We conclude with some remarks about the experimental constraints of the quantum FEL regime, as described in [24,45]. It has been proposed [46,47] that one can use, instead of a magneto-static undulator, a high-power laser beam as undulator for the X-ray emission region. In fact, since $\lambda_r \propto \lambda/2\gamma^2$, the wavelength of the emitted radiation can be reduced or increasing the electron energy or, alternatively, reducing the undulator period. An infrared laser, coming from high-power Nd or CO$_2$ lasers, with a wavelength $\lambda_L$ of 1 or 10 $\mu$m could yield X-ray FEL radiation with electrons energy about a hundred MeV.

$$\lambda_r \propto \frac{\lambda_L}{4\gamma^2} \simeq 0.1 \text{ nm} \quad (134)$$

Now, with a simple argument, we show why a laser undulator is advisable for X-ray FEL emission in the classical scheme, while its is compulsory in the quantum regime. As we have discussed, the transition between quantum and classical regime is ruled by the quantum FEL (QFEL) parameter [24]

$$\bar{\rho} = \rho \frac{mc\gamma}{\hbar c} = \rho \gamma \frac{\lambda_r}{\lambda_c} \quad (135)$$
where $\lambda_c = h/mc$ is the Compton wavelength and $h$ is the Planck constant. The QFEL parameter is proportional to the ratio between the electron energy and the single-photon energy and the quantum regime occurs when $\rho \leq 1$. Using the resonant relation and the definition of $\rho$ in Equation (135), the condition for the quantum regime can be written in terms of $\rho$:

$$\rho \leq \frac{\sqrt{2\lambda_c}}{\sqrt{\lambda_r \lambda_w (1 + a_w^2)}}$$  \hspace{1cm} (136)

We have seen that to reach the high-gain regime, a number of period of the order of $N_w \propto \frac{1}{\rho}$ is required, so that the undulator length must be:

$$L_w = N_w \lambda_w \simeq \frac{\lambda_w}{\rho} \geq \frac{\sqrt{\lambda_r \lambda_w^3 (1 + a_w^2)}}{\lambda_c}$$  \hspace{1cm} (137)

For an X-ray FEL experiment with $\lambda_r \approx 1\text{Å}$, using a magnetic undulator with a period $\lambda_w \sim 1\text{cm}$ and $E = 3.5$ Gev, we obtain from Equation (137) $L_w \geq 3$ km, which is possible but impracticable. If we substitute the magneto-static undulator with a counter-propagating laser beam, we must substitute in the resonant relation and otherwise $\lambda_w \to \lambda_L/2$, and for a typical laser with wavelength of 1 $\mu$m and an electron energy $E = 2.5$ Mev, we obtain $L_w \geq 2$ mm. However, the use of a laser undulator would require an exceptional stability in intensity and frequency in the interaction region, which makes it still very challenging [48].

5. Conclusions

We reviewed the theory of high-gain FEL, following the approach developed by R. Bonifacio and coworkers in the 1980s, which underpins the major advances in understanding that led to the next generation of X-ray FELs. The review started with a brief introduction to the basic concepts and quickly developing the equations for “spontaneous” FEL radiation, followed by a description of stimulated FEL emission, which gives an intuitive understanding of the role of bunching, which is also necessary to understand the high-gain regime. Finally, the introductory section introduces the basic concepts of the quantum FEL where electron recoil resulting of photon emission dominates the resonant interaction. The theoretical model, based on the many-particle evolution of the electrons moving in the undulator and self-consistently coupled to the emitted radiation field, has the advantage of showing the emergence of few fundamental parameters describing the main features of the FEL devices. In particular, the 1D dynamics of a Compton (i.e., high-energy) classical FEL is all described by the fundamental FEL parameter $\rho$, measuring efficiency, gain length $l_g = \frac{\lambda_w}{4\pi \rho}$ (where $\lambda_w$ is the undulator period), energy spread and gain bandwidth in the high-gain regime.

A 3D model of the FEL is developed based on a 3D Hamiltonian and Maxwell’s equations. This is expanded to include the role of space-charge effects, which become important for FELs operating in the Raman regime. To gain an intuitive understanding of FEL theory, we discussed in detail the 1D model, using suitably scaled parameters to develop a “universal” set of equations. Then we examined the steady-state regime, where slippage is neglected, and we reviewed the results of the linear analysis and a general dispersion relationship, showing the exponential collective instability. Including the slippage effect, we discussed the FEL superradiance, which is characterized by the cooperation length $l_c = (\lambda_r/\lambda_w)l_g$. Superradiance in FELs can be classified as weak superradiance for short electron bunches ($l_b \ll l_c$) and strong superradiance for long electron bunches ($l_b \gg l_c$). At the core of the superradiant emission is the existence of a self-similar solution, i.e., a solution which maintains the same profile growing and shrinking when it propagates along the electron bunch. From the self-similar solution, the main scaling laws with the electron beam length, current and undulator length follow naturally. Superradiance is also the core of the self-amplified spontaneous emission (SASE) operation, very important for the development of the X-ray sources. We described the main
features of SASE in the 1D limit, showing some typical temporal profiles and spectra of the radiation intensity. Next, we described three-dimensional effects, as optical guiding and transverse electron motion due to emittance. These effects are characterized by two other dimensionless parameters, $a = l_\gamma / Z_r$ and $b = l_b / \beta^*$ for radiation and electron diffraction, where $Z_r$ is the Rayleigh range and $\beta^*$ the betatron length. Finally, we presented the main concept of the quantum regime of FEL, which is characterized by the quantum FEL parameter $\bar{\rho} = \rho (mc^2 / \hbar \omega)$, that can be interpreted as the average number of photons emitted by each electron. The quantum regime of FEL occurs when $\bar{\rho} < 1$. The major constraints for the experimental realization of a quantum FEL have been discussed.

In this review, we preferred to present only the classical theory of FEL, since it is well established and confirmed by many experiments and is at the base of several numerical codes developed for assisting the experiments [49–52]. About the quantum FEL, it has not been experimentally proved yet and its studies are still under progress [53,54], leaving the subject still in an advanced research context, with the hope that in the near future, due to the fast progresses in particle-beam and laser technology, it may become a reality. However, since the quantum FEL stems from an original idea of Rodolfo Bonifacio, we have presented the main basic concepts of the quantum FEL, but without entering into a detailed description of its quantum theory.

Since FEL has been studied since more than fifty years, many other authors have contributed to the advance of the FEL physics, both theoretically and experimentally, and it would be a hard task to make reference of all of them. About that, several excellent reviews and books on FELs exist [1–4]. The main aim of this paper was to review the main contributions made to the FEL physics by Rodolfo Bonifacio, to which this work is dedicated.

Author Contributions: All authors conceived the study. N.P. directed the research. All authors contributed to writing of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References


