A quantum mechanical analysis of Smith–Purcell free-electron lasers

Hesham Fares a,b,*, Minoru Yamada a,c

a Faculty of Electrical and Computer Engineering, Institute of Science and Engineering, Kanazawa University, Kakuma-machi, Kanazawa 920–1192, Japan
b Department of Physics, Faculty of Science, Assiut University, Assiut 71516, Egypt
c Department of Electronic System Engineering, Malaysia-Japan International Institute of Technology (MJJIT), Universiti Teknologi Malaysia (UTM), Malaysia

A R T I C L E   I N F O

Article history:
Received 2 October 2014
Accepted 17 February 2015
Available online 26 February 2015

Keywords:
Smith–Purcell radiation
Free-electron lasers

A B S T R A C T

The paper presents a quantum mechanical treatment for analyzing the Smith–Purcell radiation generated by charged particles passing over a periodic conducting structure. In our theoretical model, the electrons interact with a surface harmonic wave excited near the diffraction grating when the electron velocity is almost equal to the phase velocity of the surface wave. Then, the surface harmonic wave is electromagnetically coupled to a radiation mode. The dynamics of electrons are analyzed quantum mechanically where the electron is represented as a traveling electron wave with a finite spreading length. The conversion of the surface wave into a propagating mode is analyzed using the classical Maxwell’s equations. In the small-signal gain regime, closed-form expressions for the contributions of the stimulated and spontaneous emissions to the evolution of the surface wave are derived. The inclusion of the spreading length of the electron wave to the emission spectral line is investigated. Finally, we compare our results based on the quantum mechanical description of electron and those based on the classical approach where a good agreement is confirmed.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

In 1953, S.J. Smith and E.M. Purcell firstly demonstrated that an optical light is emitted when an electron beam moves parallel and close over a metallic diffraction grating in vacuum [1]. The process was understood in terms of a simple model based on oscillations of the image charges induced on the metallic surface by electrons. The Smith–Purcell (SP) effect is widely considered as a possible mechanism for free-electron laser (FEL) operating over a wavelength range extending from the millimeter [2,3] to the optical region [4,5]. In recent experiments [6–8], it has been shown that the SP radiation provides a promising candidate to realize compact and tunable radiation source in the THz region. FELs based on the SP effect are operated in the amplifier and oscillator configurations where the optical feedback is required in the latter case [9–14]. Such FELs could be made with much more compact device structure compared to other FELs (e.g., undulator FELs), and therefore may be interesting for application.

Many theoretical analyses have been developed to analyze the dynamics of the Smith–Purcell radiation [15–22]. In most of these analyses, the metallic waveguide with periodic corrugation behaves as a slow wave structure through which a slow space harmonic of a transverse magnetic (TM) Floquet mode is propagated. The TM mode has a longitudinal electric field component and interacts most strongly with the traveling electrons. The working principle involves the synchronism between the velocity of the electron beam and the phase velocity of the TM surface modes of the periodic structure. The SP radiation contains a broad continuous frequency band and the radiation wavelength is determined by the observed angle, period of grating, and electron beam energy. Since the minimum corrugation period can be obtained currently is \( \lesssim 0.1 \mu m \), the generation of SP radiation operating up to the ultraviolet is commonly accomplished by using nonrelativistic electron beams. In these cases, the radiation wavelength is a fraction of the corrugation period.

In all forgoing analyses, the interacting electron is considered as a point particle where its spreading size is assumed to be much shorter than the period of grating as well as the wavelength of emitted radiation. M. Yamada in [23] developed a quantum mechanical treatment for calculating the amplification gain of an optical amplifier in which an electron beam passes above a dielectric planar waveguide. This amplifier is basically one type of the Cherenkov FELs. The theoretical analysis was performed basing on the density matrix formalism which is a quantum statistical treatment [24–26]. In the theoretical model of [23,27,28], the electron is represented by an electron wave with a finite spreading length. In [27,28], the validity of the theoretical model is examined by comparing the experimentally measured data of intensities and spectrum profiles of optical radiation with those predicted theoretically. It is confirmed that the spectral
profile of emission is characterized by the spreading length of the electron wave.

In this paper, we present a new theoretical analysis for the SP radiation in the small-signal low gain regime. In our analysis, the dynamic of the EM wave is described using the Maxwell’s wave equations. On the other hand, the dynamic of electrons are quantum mechanically analyzed using the density matrix formalism. We derive a generalized expression for the dispersion function that determines the spectral profile of emitted radiation. In this expression, the inclusion of the finite spreading length of the electron wave is studied. Our analysis is devoted for the non-relativistic electron energies (≤ 100 keV). Also, since low-density beams are utilized in the SP experiments constructed to date, the space charge effects are neglected.

In Section 2, the analytical representations of the EM waves are shown. The EM wave is assumed to compose of surface modes that propagate along the corrugated surface and radiation modes that are emitted from the corrugated surface. Formulations of the radiation power and the stored energy are presented. In Section 3, the electron dynamics are described on the basis of the density matrix method. In Section 4, the gain coefficient of amplification by the stimulated emission and the radiation rate by the spontaneous emission are derived. In Section 5, the resonance condition for the electron dynamics are described using the Maxwell’s wave equation as

\[ \nabla \cdot \mathbf{E} = \frac{\partial \mathbf{B}}{\partial t} \]

\[ \nabla \times \mathbf{H} = \frac{\partial \mathbf{E}}{\partial t} + \mathbf{J} \]

where \( \epsilon(x,y,z) \) and \( \mu(x,y,z) \) are the electrical conductivity and dielectric constant, respectively, which have different values in the metal and vacuum regions. \( \mathbf{J}(x,y,z) \) is the current density of the electron beam where the moving electrons produces an EM wave on the grating. The interaction between the EM field localized close to the grating surface and the electron beam to obtain radiation power is counted through this current density \( \mathbf{J} \). The evaluation of \( \mathbf{J} \) is performed with the help of quantum mechanical treatment assuming the electron wave representation as will be given in following sections. In our model, it is assumed that a sufficiently intense magnetostatic field in the direction of the beam flow is applied. Then, the electron beam is considered to be thin where the transverse velocities of electrons in the direction normal to the electrons propagation can be neglected. In the limit of a thin electron beam, we can neglect the effects of the self-magnetic fields in the transverse directions on the longitudinal modulations of electrons. Therefore, the EM wave is restricted to be a TM mode having \( H_r \) and \( E_z \) components where the z-component of the electric field \( E_z \) interacts most strongly with the electrons.

We define the spatial wavenumber of the corrugation corresponding to the grating period \( \Lambda \) as

\[ G = \frac{2\pi}{\Lambda} \]

\[ y_0 = \sum_{m \geq 0} d_{m0} \cos(mGz) = \sum_{m \geq 0} d_{m0} \left( e^{jmGz} + e^{-jmGz} \right) \]

In Eq. (4), \( m \) is an integer and the Fourier coefficients \( d_{m0} \) determine the peak-to-peak depth of the grating whereas \( d_{m0} \) refers to the average depth of the periodic structure. As shown in Fig. 1, we set \( y = 0 \) at the top boundary of the corrugation, the interaction length is \( L_z \), and the distribution width of the EM wave is \( L_z \). A part of the energy associated with the surface waves is transformed as radiation waves. The wavenumbers of both types of waves along z direction are modified by the corrugation, and then the fields’ distribution is critically modified by the corrugation.

The magnetic field component \( H_r \) can be written as

\[ H_r(x,y,z,t) = H_t(x,y,z,t) + H_e(x,y,z,t) \]

In Eq. (5), \( H_t \) is a radiation wave component and is given by

\[ H_t = \bar{E}(t) \sqrt{\frac{\rho_0}{\mu_0}} \left( R^+ e^{-j\beta R^+} + R^- e^{-j\beta R^-} \right) e^{jkz} + c.c. \]

and \( H_e \) is a surface field expressed as a superposition of plane waves of different frequencies as

\[ H_e = \bar{E}(t) \sqrt{\frac{\rho_0}{\mu_0}} U_l(x,y,z) e^{jkz} + c.c. \]

where

\[ U_l(x,y,z) = \sum_{m = -\infty}^{m = \infty} \left\{ A_{l+}^m(y,z) e^{-j\beta R^+} + A_{l-}^m(y,z) e^{-j\beta R^-} \right\} \]

(8) where \( \omega \) is the angular frequency, \( \bar{E}(t) \) is the temporal field amplitude and c.c. indicates the complex conjugate of the preceding terms, \( \beta \) and \( \gamma \) are the wavenumbers (i.e., propagation constants) in the z and y directions of the radiation wave, respectively. Superscripts (+) and (−) indicate the forward and backward propagating components along z direction, respectively. \( R^{\pm} \) and \( A_{l \pm}^m \) are amplitude coefficients and amplitude functions, respectively.

The radiation field \( H_r \) exists only in the vacuum region where \( y \geq 0 \). By putting \( \epsilon = \epsilon_0 \), \( \sigma = 0 \), and \( J = 0 \) in Eq. (2), the relation...
between wavenumbers is
\[ p^2 + q^2 = \mu_0 H_0 \sigma^2 = \frac{\sigma^2}{c^2}. \] (9)

The amplitude functions of the surface wave \( A_{m}^{(\pm)}(y,z) \) are assumed to be varied much slower than the spatial phase variation in the \( z \) direction \( e^{-j\omega t + j\beta z} \). Then, the surface wave must be an evanescent wave for \( \pm \gamma \) direction where \( |mG \pm \beta| > \omega/c \), and satisfies the relation of
\[ A_{m}^{(\pm)}(y,z) \rightarrow 0 \quad \text{for} \quad |y| \rightarrow \infty \quad \text{and} \quad m \neq 0. \] (10)

The principle surface mode (i.e., \( m = 0 \)) having the amplitude \( A_{0}^{(\pm)}(y,z) \) is rather exceptional. This mode directly couples to the wave radiation at \( y = 0 \), because this term has identical propagation constant \( \pm \beta \) along \( z \) direction. Hence, we can write
\[ A_{0}^{(\pm)}(y,z) = 0 \quad \text{for} \quad y > 0 \] (11a)
\[ A_{0}^{(\pm)}(y,z) = R^{(\pm)} \quad \text{for} \quad y = 0 \] (11b)
\[ |A_{0}^{(\pm)}(y,z)| \rightarrow 0 \quad \text{for} \quad y \rightarrow -\infty. \] (11c)

By separating the EM wave into radiation and surface waves, we also represent the electric field component \( E \) as
\[ E(x,y,z,t) = E_{r}(x,y,z)e^{i\omega t} + c.c. \] (12)
where \( E_{r} \) and \( E_{s} \) are the radiation and surface field components, respectively, and are given by
\[ E_{r} = \hat{E}_{r}(Q(x,y,z)e^{i\omega t} + c.c. \] (13)
\[ E_{s} = \hat{E}_{s}(P(x,y,z)e^{i\omega t} + c.c. \] (14)
\[ Q(x,y,z) \] and \( F(x,y,z) \) are field distribution functions and are derived from the magnetic component \( H_{s} \) using the relations of \( E_{r} = \{1/(j\omega + \sigma)\}(\partial H_{s}/\partial z) \) and \( E_{s} = \{-1/(j\omega + \sigma)\}(\partial H_{s}/\partial y) \), resulting in
\[ Q_{r}(x,y,z) = \frac{\partial P_{s}}{\partial z} \left\{ e^{(\omega + \sigma)z} - (R_{1}^{+})e^{-j\sigma y - j\beta z} + (R_{1}^{-})e^{-j\sigma y + j\beta z} \right\} \] (15)
\[ Q_{s}(x,y,z) = \frac{\partial P_{s}}{\partial y} \left\{ e^{(\omega + \sigma)z} - (R_{1}^{+})e^{-j\sigma y - j\beta z} + (R_{1}^{-})e^{-j\sigma y + j\beta z} \right\} \] (16)
\[ F_{r}(x,y,z) = \frac{-1}{(\omega - j\sigma)} \left\{ \sum_{m=-\infty}^{\infty} \left\{ (mG + \beta)A_{m}^{(+)\dagger}(y,z)e^{-j(mG + \beta)z} + (mG - \beta)A_{m}^{(-)\dagger}(y,z)e^{-j(mG - \beta)z} \right\} \right\} \] (17)
\[ F_{s}(x,y,z) = \frac{j}{(\omega + j\sigma)} \left\{ \sum_{m=-\infty}^{\infty} \left\{ \frac{\partial A_{m}^{(+)\dagger}(y,z)}{\partial y} e^{-j(mG + \beta)z} + \frac{\partial A_{m}^{(-)\dagger}(y,z)}{\partial y} e^{-j(mG - \beta)z} \right\} \right\}. \] (18)

In the vacuum region where \( y > 0 \), the amplitude functions must be given by
\[ A_{m}^{(\pm)}(y,z) = A_{m}^{(\pm)}(0,z)e^{-\gamma_{m}^{(\pm)}y} \quad \text{for} \quad y > 0 \] (19)
with the relation of
\[ (mG + \beta)^2 - y_{m}^{(\pm)2} = \frac{\sigma^2}{c^2}. \] (20)
Therefore, the distribution functions are given by
\[ F_{r}(x,y,z) = \frac{1}{\omega} \sum_{m=-\infty}^{\infty} \left\{ (mG + \beta)A_{m}^{(+)\dagger}(0,z)e^{-\gamma_{m}^{(+)y} - j(mG + \beta)z} \right\} \] (21)
\[ F_{s}(x,y,z) = \frac{1}{\omega} \sum_{m=-\infty}^{\infty} \left\{ (mG - \beta)A_{m}^{(-)\dagger}(0,z)e^{-\gamma_{m}^{(-)y} - j(mG - \beta)z} \right\}, \quad \text{for} \quad y > 0 \] (22)

2.2. Condition of the Smith–Purcell radiation

Here, we examine the SP radiation condition basing on our notation. Assuming the radiation angle is \( \theta \) as shown in Fig. 1, we get
\[ \pm \beta = \frac{\omega}{c} \cos(\theta) \] (23)
\[ \gamma = \frac{\omega}{c} \sin(\theta) \] (24)
where \( +\beta \) and \( -\beta \) correspond to the cases of forward and backward propagations when \( 0 \leq \theta \leq \pi/2 \) and \( \pi/2 \leq \theta \leq \pi \), respectively.
To generate the SP radiation, the velocity of electron \( v_{e} \) must coincide with the phase velocity of the induced wave, that is \( (mG \pm \beta)v_{e} = \omega \).

By substituting Eq. (23) into Eq. (25), we get the relation of
\[ \omega = \frac{2mc}{A} \left\{ \frac{1}{|c/v_{e} - \cos(\theta)|} \right\}. \] (26)

Eq. (26) can be rewritten in terms of the emitted wavelength as
\[ \lambda = \frac{A}{m} \frac{c}{v_{e}} \cos(\theta) \] (27)
which is the well-known resonance condition of the SP radiation.

The relation between the angular frequency and the EM wavenumber along \( z \) direction is illustrated in Fig. 2. The lines \( \omega = \pm \beta \) correspond to the case of \( \gamma = 0 \) at which \( \theta = 0 \). When the operating point is located between the light lines (i.e., \( \omega = \pm \beta \)), the radiation into vacuum space is generated. The beam line \( \omega = (mG \pm \beta)v_{e} \) indicates the condition of the interaction between the EM field and the electron beam. If the interaction is occurred at a point \( P \), other operating points \( P' \), \( P' \), \( P'' \), given with different numbers of \( m \), are also induced. Then the radiation is generated at the point \( P' \) in Fig. 2.

A schematic illustration of the analytic model used in this paper is described in Fig. 3. The electron beam interacts with one field harmonic having the modified wavenumber of \( mG \pm \beta \). This
interaction will be analyzed in Section 2.2 on the basis of the density matrix method. Although the EM wave is separated into different modes, but they are coupled to each other, and then the radiation wave is generated. This coupling is explained by Maxwell’s wave equation.

### 2.3. Radiating power and stored energy

The Poynting vector of the radiating wave for \( \theta \) or \( \pi - \theta \) direction is

\[
\mathbf{S}_r(\theta) = 2 \sqrt{\frac{\varepsilon_0}{\mu_0}} |\mathbf{E}(\theta)|^2 \sin(\theta).
\]

Then, the total radiation power \( P \) emitted from the corrugated surface is

\[
P = \int_{0}^{L_z} \int_{-\infty}^{\infty} \int_{0}^{\infty} \mathbf{S}_r(\theta) d\theta d\xi d\zeta = 2L_\text{L}_x \int_{0}^{L_z} \int_{-\infty}^{\infty} \int_{0}^{\infty} |\mathbf{E}(\theta)|^2 \sin(\theta) d\theta d\xi d\zeta.
\]

where

\[
\varepsilon_{av} = \frac{1}{2} \int_{0}^{L_z} \int_{-\infty}^{\infty} \int_{0}^{\infty} \varepsilon |\mathbf{E}(x,y,z)|^2 d\theta d\xi d\zeta
\]

is an averaged dielectric constant for the surface wave.

### 2.4. Equation for evolution of the EM wave

The surface evanescent wave loses energy due to the transformation into radiation and the unavoidable loss of metallic grating. Therefore, we define a decay time \( \tau_{\text{rad}} \) as an eigen value of the following equation

\[
\left\{ \nabla^2 - \mu_0 \sigma(x,y,z) \frac{\partial}{\partial t} - \mu_0 \varepsilon(x,y,z) \frac{\partial^2}{\partial t^2} \right\} \mathbf{F}(x,y,z) e^{i(\omega - (1/2\tau_{\text{rad}}))\tau} = 0.
\]

Assuming \( \omega > 1/\tau_{\text{rad}} \), we get from Eq. (33)

\[
\left\{ \nabla^2 - \mu_0 \sigma(x,y,z) + \mu_0 \varepsilon(x,y,z) \right\} \mathbf{F}(x,y,z) = -\frac{\mu_0}{\tau_{\text{rad}}} \left( \sigma \int_{0}^{\infty} \mathbf{F}(x,y,z) e^{i(\omega - (1/2\tau_{\text{rad}}))\tau} d\tau \right).
\]

By substituting Eqs. (12)–(14) into Eq. (1), using the relations of Eqs. (9) and (34), and neglecting the second derivative of \( \mathbf{E}(t) \) (i.e., \( \partial^2 \mathbf{E}/\partial t^2 \approx 0 \)), we get the following equation:

\[
\frac{dE}{dt} = -\frac{1}{2\varepsilon_{av}} \int_{0}^{L_z} \int_{-\infty}^{\infty} \int_{0}^{\infty} \mathbf{F}^* e^{i\omega t} dx dy dz - \frac{\tilde{E}(t)}{2\tau_{\text{rad}}}
\]

with

\[
\tilde{E}(t) = \int_{0}^{L_z} \int_{-\infty}^{\infty} \int_{0}^{\infty} e^{i(\omega - (1/2\tau_{\text{rad}}))\tau} \mathbf{F}(x,y,z) dx dy dz.
\]

In Eq. (36), \( \langle \cdots \rangle \) refers to the operation of the quantum statistical average.

To get a sufficient gain for laser oscillation, the radiating power should be larger than the power lost by the grating absorption, and hence, we approximately write

\[
\frac{1}{\tau_{\text{rad}}} \approx \frac{P}{W}
\]

where \( P \) is the total radiation power given by Eq. (29) and \( W \) is the stored energy of the surface wave as given in Eq. (32).

### 3. Electron dynamics

From the SP condition shown in Eq. (27), the radiated wavelength \( \lambda \) can be reduced by decreasing the corrugation period \( \Lambda \). The metallic corrugation assists in establishing the phase variation of the EM field. It is noticed that the corrugated metallic surface behaves as the static magnetic field in the undulator FEL. As SP experiments move toward the generation of shorter wavelength radiation, shorter corrugation periods are required. The minimum \( \lambda \) obtained currently can be in the order of few tenths of micrometers. When the corrugation period becomes shorter than the spreading length of an electron wave, which is in the order of tens of micrometers (\( \sim 40 \mu m \)) [27,28], the electron sees non-uniform dielectric constant along a single electron. Equivalently, an electron wave is subjected to a spatially variant electric force from the electron beam.
surface EM wave. In this paper, we take into account the quantum mechanical representation of an electron. Then we examine the validity of the quantum model by examining the operating conditions at which this model reveals similar results of the classical treatments.

By assigning a number \( \nu \) for each electron in the electron beam, the electron wave function is given by

\[
\psi(t, \mathbf{r}) = \sum_j c_j(t) \phi_j(\mathbf{r}) e^{-i\omega_j t} \tag{39}
\]

where \( \phi_j(\mathbf{r}) \) is a spatial distribution function of an electron wave at the \( \nu \)-th energy state. The coefficient \( c_j(t) \) weights the contribution of the \( \nu \)-th electron. In the quantum mechanical framework, an electron must be considered as a wave packet with a finite length. The electron packet can be expressed in the form of linear superposition of plane waves. Since the summation over all possible wavenumbers is inherent in calculating the expectation values, it is mathematically favorable to express the electron wave function in the form of a plane wave. In this work, we use the formulation of the electron wave as introduced in Ref. [29], where the electron wave is assumed to be a plane wave which spreads in a rectangular box whose volume is \( \epsilon^3 \)

\[
|f\rangle = \phi_f(\mathbf{r}) = \frac{1}{\sqrt{\epsilon^3}} e^{ik_F \cdot \mathbf{z}} \tag{40}
\]

Eqs. (39) and (40) state that the electron travels in the form of \( e^{ik_F z} \) in Ref. [30]. The expectation value of an arbitral operator or function \( \langle \mathcal{F} \rangle \) is then obtained by (i.e., see Eq. (A8) in Appendix A)

\[
\langle \mathcal{F} \rangle = \sum_{f \rightarrow g} \rho_{fg}(t) \mathcal{F}_{fg} = \sum_f \langle f | \mathcal{F} | g \rangle = \text{Tr}(\rho \mathcal{F}) \tag{41}
\]

where \( \text{Tr}(...) \) is the trace operation, \( \rho \) is the density matrix in which the dynamic of an electron is involved, and \( |f\rangle \) is an eigenstate of an electron. By introducing the identity operator

\[
1 = \sum_g |g\rangle |g\rangle \tag{42}
\]

the first term in the right-hand side of Eq. (36) is given as

\[
\langle \mathcal{F}^*_e e^{-i\omega_j t} \rangle = \sum_{f \rightarrow g} \langle f | \rho_g | f \rangle \mathcal{F}^*_e e^{-i\omega_j t} \langle f | \mathcal{F} | g \rangle \tag{43}
\]

As shown in Eq. (A12), the dynamics of an electron are described using the dynamic equation of the density matrix

\[
\frac{d\rho}{dt} = \frac{i}{\hbar} \left[\hat{H}_0 + \hat{H}_{\text{int}}, \rho\right] - \frac{1}{2} \left\{ \rho, \hat{\Gamma} \hat{\Gamma}^\dagger - \hat{\Gamma}^\dagger \hat{\Gamma} \rho \right\} \tag{44}
\]

where \( \hat{H}_0 \) and \( \hat{H}_{\text{int}} \) are the principal “free” and interaction Hamiltonians of electrons, respectively. \( \rho \) is a distribution function of electrons at the thermal equilibrium. \( \hat{\Gamma} \) is an operator representing the relaxation effect of the electron wave due to the mutual scattering of electrons.

In Eq. (44), the principal Hamiltonian \( \hat{H}_0 \) has the property

\[
\langle f | \hat{H}_0 | g \rangle = W_{fg} \delta_{fg} \tag{45}
\]

where

\[
W_{fg} = \frac{\hbar^2 k_F^2}{2m_0} - \frac{m_0 v_F^2}{2} = \hbar y. \tag{46}
\]

\[
H_{\text{int}} \text{ is given by}
\]

\[
H_{\text{int}} = -\frac{e^2}{2m_0 \epsilon_0} \left( E_z E_x + E_x E_y + H.c. \right) \tag{47}
\]

where \( e \) is the electron charge unit, \( m_0 \) is the rest mass of an electron, \( E_z \) is the \( z \)-component of the external electric field, \( p \) is the momentum operator, and \( H.c. \) means the Hermite conjugate of the preceding terms. \( \Gamma \) is assumed to have only diagonal elements and whose expectation value is \( 1/\tau_e \), where \( \tau_e \) is the electron relaxation time,

\[
\langle f | \mathcal{F} | g \rangle = \frac{1}{\tau_e} \delta_{fg} \tag{48}
\]

In our study, for weakly pumped SP FELs operated in the low gain regime (i.e., the electron beam current is below 1 A), we neglect the space-charge (collective) effects. At high electron densities, the inclusions of the space-charge effects can be considered by taking into account the static electric field of neighboring electrons in the dynamics formulations. This can be done by considering the scalar potential \( U \), generated by the surrounding modulated electrons, in the electron Hamiltonian given by Eq. (47) (i.e., the term \( -eU \) appears as an additional term in the left-hand side of Eq. (47)). Detailed discussions on how to take into count the space-charge effects can be found in Refs. [25,31].

The electron current density \( \mathbf{J} \) is given in terms of the electron density \( \rho \) and the quantum mechanical operator of \( p \) (or \( v \)),

\[
\mathbf{J} = -e \mathbf{v} \rho = -\frac{e}{m_0} \mathbf{p} \tag{49}
\]

Since the momentum operator \( p \) has diagonal elements only where

\[
\langle f | | g \rangle = \hbar k \delta_{fg} = m_0 v_F \delta_{fg} \tag{50}
\]

the expectation value of \( \langle \mathcal{F}^*_e e^{-i\omega_j t} \rangle \) in Eq. (43) can be rewritten as

\[
\langle \mathcal{F}^*_e e^{-i\omega_j t} \rangle = \frac{1}{Z} \sum_{g} \mathcal{N}_g \langle f | \rho | g \rangle \mathcal{F}^*_e e^{-i\omega_j t} \tag{51}
\]

where \( \mathcal{N}_g \) is the spatially averaged electron density in the electron beam. From Eq. (A7), the diagonal element of the density matrix \( \rho_{fg}(t) \) corresponds to the transition state of an electron where the electron transits from level \( f \) to level \( g \) (i.e., the mixed state). In the time interval of the electron transition, the electron wave has a beating vibration of \( e^{-i(\omega_f - \omega_g) t} \). Note that \( \rho_{fg}(t) \) is also proportional to \( e^{-i(\omega_f - \omega_g) t} \) as shown in Eq. (A7). When this beating vibration coincides with the phase vibration of the EM wave \( e^{-i\omega_j t} \) (i.e., \( \omega_f - \omega_g = \omega \)) corresponding to the energy conservation rule, the interaction occurs. From the above discussion, the electron density can be treated as a quantum mechanical operator in the form of \( \rho \). For the double summations over energy levels \( f \) and \( g \) in Eq. (51), the case of \( f = g \) represents a dynamic motion of an electron without making any transition from the initial energy level \( f \) to other energy levels. In this case, the term \( \mathcal{N}(f | \rho | f) = \mathcal{N}_{ff} \) in Eq. (51) denotes the average component of the electron density (or the average current density). In the case of \( f \neq g \) in the double summations of Eq. (51), the term \( \mathcal{N}(f | \rho | g) = \mathcal{N}_{fg} \) denotes the modulated component of the electron density (or the modulated term current density).
By substituting Eqs. (45)–(48) into Eq. (44), the dynamic equation for the off-diagonal element of the density matrix \( \rho_{f} \) is given by
\[
\frac{d\rho_{f}}{dt} = -j \omega_{f} \rho_{f} + \frac{e}{2m_{0} \hbar \omega_{0}} \left( \langle f | p E_{r} | g \rangle \langle g | E_{r} p | f \rangle - \langle f | p E_{z} | g \rangle \langle g | E_{z} p | f \rangle - \langle f | E_{r} p | f \rangle - \langle f | E_{z} p | f \rangle \right) - \frac{\rho_{f}}{\tau_{e}}.
\] (52)

In Eq. (52),
\[
\langle f | E_{r} p | g \rangle = m_{0} v_{f} \langle f | E_{z} | g \rangle / \rho_{fg},
\] (53)
\[
\langle f | E_{z} p | g \rangle = \langle f | E_{z} | g \rangle m_{0} v_{fg} / \rho_{fg},
\] (54)
\[
\langle f | p E_{z} | g \rangle = \rho_{f} m_{0} v_{fg} \langle f | E_{z} | g \rangle,
\] (55)
\[
\langle f | E_{z} p | f \rangle = \rho_{f} \langle f | E_{z} | g \rangle m_{0} v_{fg}.
\] (56)

For the electric field interacting with the electron beam and is given by Eqs. (12), (14) and (22), we define the off-diagonal element of the field distribution of the surface wave \( F_{p}(x, y, z) \) as
\[
\langle f | F_{p}(x, y, z) | g \rangle = \sum_{m = -\infty}^{\infty} \left[ F_{+, m}^{(1)} / \lambda_{m}^{(1)} \right] \left\{ \frac{1}{\tau_{e}} \int_{0}^{t} \int_{0}^{t} e^{i [\omega t - \beta c z / \lambda^{(1)}]} d\tau_{1} d\tau_{2} \right\}.
\] (57)

with
\[
F_{+, m}^{(1)} = -\frac{e}{\hbar \omega_{0}} \left( 1 - e^{-\beta c z / \lambda^{(1)}} \right) \frac{e^{i \omega_{0} t}}{\lambda^{(1)}} a_{k_{g}} - k_{f} - m \gamma / \beta c / 2 \}
\] (58)

Using Eq. (40) of the electron wave function, the inner product in Eq. (58) will be
\[
\langle f | e^{-i \omega_{0} t} e^{i \omega_{0} t} - j \gamma t \hat{n} \cdot \mathbf{G} + \beta c / 2 \rangle = 1 / \tau_{e} \int_{0}^{t} \int_{0}^{t} e^{-i \omega_{0} t} e^{i \omega_{0} t} - j \gamma t \hat{n} \cdot \mathbf{G} + \beta c / 2 \rangle dx dy dz.
\] (59)

Hence, Eq. (58) becomes
\[
F_{+, m}^{(1)} = -\frac{e}{\hbar \omega_{0}} \left( 1 - e^{-\beta c z / \lambda^{(1)}} \right) \omega_{0} a_{k_{g}} - k_{f} - m \gamma / \beta c / 2 \}
\] (60)

where \( \sin(X) = \sin(X)/X \). From Eq. (60), it is shown that \( F_{+, m}^{(1)} \) becomes maximum when \( k_{g} - k_{f} - m \gamma / \beta c = 0 \) which implies the momentum conservation rule for the electron transition. This transition is caused by the coupling between the electron and EM waves for a specific geometry of metallic corrugation.

Another inner product for the complex conjugate of the field distribution function is
\[
\langle f | F_{p}^{*}(x, y, z) | g \rangle = \langle g | F_{p}(x, y, z) | f \rangle^{*} = \sum_{m = -\infty}^{\infty} \left[ F_{+, m}^{(1)} / \lambda_{m}^{(1)} \right] \left\{ \frac{1}{\tau_{e}} \int_{0}^{t} \int_{0}^{t} e^{i \omega_{0} t} e^{i \omega_{0} t} - j \gamma t \hat{n} \cdot \mathbf{G} + \beta c / 2 \rangle dx dy dz \right\}.
\] (61)

Using Eqs. (57) and (61), we obtain
\[
\langle f | E_{f} | g \rangle = \tilde{E}(t) \sum_{m = -\infty}^{\infty} \left[ F_{+, m}^{(1)} / \lambda_{m}^{(1)} \right] e^{i \omega_{0} t} + \tilde{E}^{*}(t) \sum_{m = -\infty}^{\infty} \left[ F_{+, m}^{(1)} / \lambda_{m}^{(1)} \right] e^{-i \omega_{0} t}.
\] (62)

Then the dynamic equation depicted by Eq. (52) is rewritten as
\[
\frac{d\rho_{f}}{dt} = -j \omega_{f} \rho_{f} + \frac{e(v_{f} + v_{g}) \rho_{fg} - \rho_{ff}}{2 \hbar \omega_{0}} \left\{ \tilde{E}(t) \sum_{m = -\infty}^{\infty} \left[ F_{+, m}^{(1)} / \lambda_{m}^{(1)} \right] e^{i \omega_{0} t} + \tilde{E}^{*}(t) \sum_{m = -\infty}^{\infty} \left[ F_{+, m}^{(1)} / \lambda_{m}^{(1)} \right] e^{-i \omega_{0} t} \right\} - \frac{\rho_{f}}{\tau_{e}}.
\] (63)

The solution of Eq. (63) takes the form
\[
\rho_{f} = \frac{e(v_{f} + v_{g}) \rho_{fg} - \rho_{ff}}{2 \hbar \omega_{0}} \left\{ \tilde{E}(t) \sum_{m = -\infty}^{\infty} \left[ F_{+, m}^{(1)} / \lambda_{m}^{(1)} \right] e^{i \omega_{0} t} + \tilde{E}^{*}(t) \sum_{m = -\infty}^{\infty} \left[ F_{+, m}^{(1)} / \lambda_{m}^{(1)} \right] e^{-i \omega_{0} t} \right\}.
\] (64)

Using Eq. (64) in Eq. (43), Eq. (43) becomes
\[
\langle F_{p}^{*}(x, y, z) e^{-i \omega_{0} t} \rangle = \sum_{m = -\infty}^{\infty} \sum_{f = 0}^{\infty} e^{2 \pi n_{f} v_{g} / \lambda_{m}^{(1)}} \rho_{fg} / \rho_{ff} \left\{ \tilde{E}(t) \sum_{m = -\infty}^{\infty} \left[ F_{+, m}^{(1)} / \lambda_{m}^{(1)} \right] e^{i \omega_{0} t} + \tilde{E}^{*}(t) \sum_{m = -\infty}^{\infty} \left[ F_{+, m}^{(1)} / \lambda_{m}^{(1)} \right] e^{-i \omega_{0} t} \right\}.$$ (65)

4. Radiation gain and spontaneous emission

By substituting Eq. (65) into Eq. (36), we get the following equation for the temporal variation of the field amplitude:
\[
\frac{d\tilde{E}}{dt} = \int_{0}^{t} \int \int_{-\omega_{0} / 2}^{\omega_{0} / 2} \sum_{m = -\infty}^{\infty} \sum_{f = 0}^{\infty} e^{2 \pi n_{f} v_{g} / \lambda_{m}^{(1)}} \rho_{fg} / \rho_{ff} \left\{ \tilde{E}(t) \sum_{m = -\infty}^{\infty} \left[ F_{+, m}^{(1)} / \lambda_{m}^{(1)} \right] e^{i \omega_{0} t} + \tilde{E}^{*}(t) \sum_{m = -\infty}^{\infty} \left[ F_{+, m}^{(1)} / \lambda_{m}^{(1)} \right] e^{-i \omega_{0} t} \right\}.$$ (66)

From Eq. (66), the variation of the EM wave intensity is given by
\[
\frac{d\tilde{E}^{2}}{dt} = \frac{d\tilde{E}^{2}}{dt} \tilde{E}^{*} \tilde{E} + \frac{d\tilde{E}^{2}}{dt} \tilde{E} \tilde{E} = \left( G_{it} - \frac{1}{\tau_{rad}} \right) \tilde{E}^{2},$$ (67)

where \( G_{it} \) is the gain coefficient given by
\[
G_{it} = \int_{0}^{t} \int \int_{-\omega_{0} / 2}^{\omega_{0} / 2} \sum_{m = -\infty}^{\infty} \sum_{f = 0}^{\infty} e^{2 \pi n_{f} v_{g} / \lambda_{m}^{(1)}} \rho_{fg} / \rho_{ff} \left\{ \tilde{E}(t) \sum_{m = -\infty}^{\infty} \left[ F_{+, m}^{(1)} / \lambda_{m}^{(1)} \right] e^{i \omega_{0} t} + \tilde{E}^{*}(t) \sum_{m = -\infty}^{\infty} \left[ F_{+, m}^{(1)} / \lambda_{m}^{(1)} \right] e^{-i \omega_{0} t} \right\}.$$ (68)
with
\[
\left| P_{mg}^{f} \right|^2 = \left( \frac{C}{\omega} \left( 1 - e^{-i\theta/m} \right) \right)^2 A_{n0}^{*} A_{m0} + \left( 0,2 \right)^2 \sin^2 \left\{ \left( k_{b} - k_{f} - mG \mp \beta \right) \epsilon / 2 \right\}.
\]

As shown in Fig. (4), we assume that the initial energy level of electron is b where it transits to a lower energy level a resulting in an emission and transits to a higher energy level c resulting in absorption. The emission and absorption process satisfy the relations of
\[
\omega_{b} - \omega_{a} = \omega
\]
and
\[
\omega_{c} - \omega_{b} = \omega
\]
with the condition of
\[
\rho_{aa} = \rho_{cc} = 0.
\]

Eq. (72) state that the there is a zero probability of finding an electron at final energy levels a and c assuming the energy separation of \( \hbar \omega \) is much larger than the thermal energy of the electron beam.

Using Eqs. (70)–(72), the gain coefficient shown in Eq. (68) becomes
\[
G_{st} = \int_{\theta}^{\beta} \int_{\gamma}^{\delta} e^{\sum_{n} N_{e} \rho_{nb}} \left\{ \sum_{m=0}^{\infty} \left[ \frac{e^{i\omega}}{\alpha} \sum_{m=0}^{\infty} \left[ (2m+1) \right] \frac{E_{m}^{2}}{E_{m}^{2}} \right] \right\} dx dy dz
\]
\[
+ \left[ \begin{array}{c} \left\{ \begin{array}{c} (m_n+1) \frac{E_{m}^{2}}{E_{m}^{2}} \right\} - \left\{ \begin{array}{c} (m_n+1) \frac{E_{m}^{2}}{E_{m}^{2}} \right\} \end{array} \right\} \end{array} \right] \right\} dx dy dz \times \left[ 1 - e^{-\hbar / \tau_{c}} \right].
\]

The average current density \( J \) can be represented in terms of the diagonal element \( \rho_{nb} \) and the initial electron velocity \( v_{e} = v_{b} \) as
\[
J = e(N_{e}) = e \sum_{n} N_{e} \rho_{nb} = e N_{e}
\]
(74)

Here, it is instructive to point that a forward/or backward component with a specific \( m \) can satisfy the condition \( k_{b} - k_{f} - mG \mp \beta \approx 0 \) according to the momentum conservation rule. The radiation wave can never interact with the electron beam in the absence of the corrugation since \( k_{b} - k_{f} + \beta \approx 0 \).

We define a spatial coupling \( \xi_{m} \) as
\[
\xi_{m} = \int_{\theta}^{\beta} \int_{\gamma}^{\delta} e^{\sum_{n} N_{e} \rho_{nb}} \left\{ \sum_{m=0}^{\infty} \left[ \frac{e^{i\omega}}{\alpha} \sum_{m=0}^{\infty} \left[ (2m+1) \right] \frac{E_{m}^{2}}{E_{m}^{2}} \right] \right\} dx dy dz
\]
(75)
and a dispersion function \( D_{stm} \) as
\[
D_{stm} = \sin^2 \left\{ \left( k_{b} - k_{a} - mG \mp \beta \right) \epsilon / 2 \right\} - \sin^2 \left\{ \left( k_{c} - k_{b} - mG \mp \beta \right) \epsilon / 2 \right\}.
\]

Since our model is restricted to the low gain regime of SP FELs, we can neglect the variations in the refractive indices of the electron beam and metal regions due to the presence of the electron beam. In this limit, to evaluate the coupling coefficient \( \xi_{m} \) given in Eq. (75), possible solutions for the amplitude functions of the surface wave \( A_{m}^{\pm} \) can be found in [32].

Combining Eqs. (74)–(76) with Eq. (73), the gain coefficient is rewritten in a simple form as
\[
G_{st} = \frac{e^{2} \hbar \omega C_{sp}}{x_{m} h_{\omega}} \xi_{m} D_{stm}.
\]
(77)

In Eq. (77), \( t \) is considered as the interaction time counted as the time interval for passing the interaction region. The term \( \tau_{c} \left[ 1 - e^{-\hbar / \tau_{c}} \right] \) shows that the gain coefficient is influenced by the relaxation time \( \tau_{c} \) when \( \tau_{c} \ll t \), while it depends on the interaction time \( \Delta t \) when \( t \ll \tau_{c} \) [33,34].

The spontaneous emission, which is the seed of the emission process, can be investigated from the quantization of the EM field. In the stored energy of the surface wave \( W \) expressed by Eq. (31), the energy of the magnetic field is not equal to that of the electric field in general, but they are almost same. The field quantization of the stored energy can be realized by introducing the photon number \( S \) such as
\[
W = 2e_{\omega} \left| \tilde{E}(t) \right|^{2} = \hbar \omega S.
\]
(78)

Then, the EM wave intensity is given by
\[
\left| \tilde{E}(t) \right|^{2} = \frac{\hbar \omega S}{2 e_{\omega}}
\]
(79)

The electron transition to a lower energy level (from \( b \) to \( a \)) corresponds the emission process and is proportional to \( S + 1 \). On the other hand, the electron transition to a higher energy level (from \( b \) to \( c \)) corresponds to the absorption process and is proportional to \( S \). In the emission process consisting of stimulated and spontaneous emissions, the spontaneous emission has almost the same characteristics of the stimulated emission except that it is independent of the existing radiation field and corresponds to the zero-point energy for which \( S = 1 \). Using Eqs. (67) and (79) and putting \( S = 1 \), we get a term \( C_{sp} \) to indicate spontaneous emission by counting the electron transition from \( b \) to \( a \) and such as
\[
\frac{d\left| \tilde{E}(t) \right|^{2}}{dt} = \left( C_{st} - \frac{1}{\tau_{rad}} \right) \left| \tilde{E}(t) \right|^{2} + C_{sp}
\]
(80)
where
\[
C_{sp} = \frac{e^{2} \hbar \omega C_{sp}}{x_{m} h_{\omega}} \xi_{m} D_{stm}
\]
and \( D_{stm} \) is a dispersion function for the spontaneous emission and is given by
\[
D_{stm} = \sin^2 \left\{ \left( k_{b} - k_{a} - mG \mp \beta \right) \epsilon / 2 \right\} - \sin^2 \left\{ \left( k_{c} - k_{b} - mG \mp \beta \right) \epsilon / 2 \right\}.
\]

From Eq. (67), the intensity of the surface wave at the steady state is given by
\[
\left| \tilde{E}(t) \right|^{2} = \frac{C_{sp}}{\left( 1 / \tau_{rad} \right) - C_{st}}.
\]
(83)

Then the radiation power given in Eq. (29) becomes
\[
P = 2L \int_{\phi} \sin \left( \theta \right) \sqrt{ \frac{\hbar \omega C_{sp}}{x_{m} h_{\omega}} \left( \left| R_{+} \right|^{2} + \left| R_{-} \right|^{2} \right) } \left( 1 / \tau_{rad} \right) - C_{st}.
\]
(84)

When the amplification gain \( C_{st} \) is not large enough to compensate the loss to the radiation modes (i.e., \( C_{st} < 1 / \tau_{rad} \)), the EM radiation will be only a spontaneous emission, and then
\[
P = 2L \int_{\phi} \sin \left( \theta \right) \sqrt{ \frac{\hbar \omega C_{sp}}{x_{m} h_{\omega}} \left( \left| R_{+} \right|^{2} + \left| R_{-} \right|^{2} \right) } \tau_{rad}.C_{sp}.
\]
(85)

It is noted that the stimulated emission is difficult to be realized without setting an external cavity to reflect back the radiation field for which \( 1 / \tau_{rad} \approx 0 \).
5. Representation of the electron wavenumbers with the electron velocity

The dispersion functions $D_{\text{trm}}$ and $D_{\text{rpm}}$ are given in terms of the electron wave numbers as shown in Eqs. (76) and (82). In this section, we rewrite these functions in more simple and practical form.

The energy of the electron at the initial state $b$ is

$$W_b = \frac{\hbar k_b^2}{2m_0} = \frac{m_0v_b^2}{2} \tag{86}$$

and the electron energy at the final state $a$ is

$$W_a = \frac{\hbar k_a^2}{2m_0} = \frac{m_0v_a^2}{2} - \hbar \omega. \tag{87}$$

Then the wave number $k_a$ is represented with the initial electron velocity $v_a$ as

$$k_a = \frac{m_0v_a}{\hbar} \left(1 - \frac{2\hbar\omega}{m_0v_a^2}\right)^{1/2} \tag{88}$$

and the difference between the initial and final wave numbers is

$$k_b - k_a = \frac{m_0v_b}{\hbar} \left\{1 - \left(1 - \frac{2\hbar\omega}{m_0v_b^2}\right)^{1/2}\right\} = \frac{\omega}{v_b} \left\{1 + \frac{\hbar\omega}{2m_0v_b^2}\right\}. \tag{89}$$

For clarity, we put $v_b = v_c$ to refer to the electron velocity, then

$$\frac{(k_b - k_a - mG + \beta)\epsilon}{2} = -\left\{(mG + \beta)\omega - \omega^2\right\} \frac{\epsilon}{2V_c}. \tag{90}$$

Similarly, the electron energy at the final state $c$ is expressed as

$$W_c = \frac{\hbar k_c^2}{2m_0} = \frac{m_0v_c^2}{2} + \hbar \omega \tag{91}$$

and then

$$\frac{(k_c - k_b - mG + \beta)\epsilon}{2} = -\left\{(mG + \beta)\omega - \omega^2\right\} \frac{\epsilon}{2V_c}. \tag{92}$$

Using Eqs. (90) and (92), the dispersion function is rewritten as

$$D_{\text{trm}} = \text{Sinc}^2 \left\{\left(mG + \beta\right)\omega - \omega^2\right\} \frac{\epsilon}{2V_c} \tag{93}$$

The critical difference between $k_b - k_a$ and $k_c - k_b$ is required to obtain $D_{\text{trm}}$, while the difference $k_b - k_a$ is only counted to obtain $D_{\text{rpm}}$ where

$$D_{\text{rpm}} = \text{Sinc}^2 \left\{\left(mG + \beta\right)\omega - \omega^2\right\} \frac{\epsilon}{2V_c}. \tag{94}$$

Numerical examples of the functions $D_{\text{trm}}$ and $D_{\text{rpm}}$ are shown in Fig. 5, where $\lambda = 0.5 \, \mu m$, $\Lambda = 0.4 \, \mu m$, $m = 3$, and $\epsilon = 30 \, \mu m$. In this example, we use the value of the coherent length of the electron wave $\ell$ obtained in the experiments of Refs. [27,28]. The electron velocity is $v_e \approx 1.15 \, m/s$ which corresponds the acceleration voltage of $V = 37.3 \, kV$.

Since the SP radiation is usually measured as a spectrum profile for fixed electron velocity $v_e$ and radiation angle $\theta$, we rewrite the functions $D_{\text{trm}}$ and $D_{\text{rpm}}$ as functions of measuring wavelength $\lambda$. By using Eqs. (3) and (23), the dispersion coefficient $(mG + \beta)\omega - \omega$ becomes

$$(mG + \beta)\omega - \omega = 2\pi \frac{m}{\Lambda} \left(1 - \frac{\epsilon}{V_c} - \cos(\theta)\right). \tag{95}$$

We refer to the center wavelength of the emission is $\lambda_0$, which is obtained by putting Eq. (94) equal to 1, as

$$\lambda_0 = \frac{\Lambda}{m} \left(1 - \frac{\epsilon}{V_c} - \cos(\theta)\right). \tag{96}$$

Then, the function $D_{\text{rpm}}$ is rewritten as

$$D_{\text{rpm}} = \text{Sinc}^2 \left\{\frac{\pi \epsilon}{\Lambda} \left(1 - \frac{\lambda_0}{\lambda}\right)\right\} \tag{97}$$

and $D_{\text{trm}}$ is rewritten as

$$D_{\text{trm}} = \text{Sinc}^2 \left\{\frac{\pi \epsilon}{\Lambda} \left(1 - \frac{\lambda_0}{\lambda}\right) - \frac{hc^2}{2m_0v_c^2} \left(\frac{\lambda_0}{\lambda}\right)\right\}$$

$$- \text{Sinc}^2 \left\{\frac{\pi \epsilon}{\Lambda} \left(1 - \frac{\lambda_0}{\lambda}\right) + \frac{hc^2}{2m_0v_c^2} \left(\frac{\lambda_0}{\lambda}\right)\right\}. \tag{98}$$

where $h$ is Planck’s constant.

Numerical examples of the dispersion functions $D_{\text{trm}}$ and $D_{\text{rpm}}$ shown in Eqs. (97) and (98) are shown in Fig. 6. In this example, it is assumed that $\Lambda = 0.5 \, \mu m$, $m = 2$, $\epsilon = 30 \, \mu m$, $V = 30 \, kV$, and $\theta = \pi/2$ to generate a spontaneous emission with $\lambda_0 = 0.728 \, \mu m$. In this example, since the emission is almost spontaneous in
nature, the full width at half maximum of the spectral width is almost 8% of the central wavelength.

It is instructive to point out that the dispersion functions $D_{2\omega m}$ and $D_{\omega m}$ of the stimulated and spontaneous emissions are defined in terms of the so-called coherent length of an electron $\ell$. Due to collisions among electrons, the spreading length must not be infinitely long (i.e., shorter than the interaction length). Therefore, it should be the main cause of the spectrum broadening. In the experiments shown in Ref. [35], a grating with a 0.416 $\mu$m pitch and the electron acceleration voltage of 30 kV are used to obtain SP spontaneous radiation with a peak wavelength of 1.3 $\mu$m at the observation angle $\theta = \pi/2$. The grating length (i.e., the interaction length) is 2.5 cm. In Ref. [35], it is shown that the spectral full width at half maximum (FWHM) is $\Delta \lambda \sim 0.1 \mu$m in the wavelength domain and is $\Delta \omega \sim 1.1 \times 10^{14}$ s$^{-1}$ in the angular frequency domain. It is obvious that the amounts of the broadening $(\Delta \omega) = 2\pi/\ell$ and $(\Delta \lambda) = 2\pi/\ell$ due the relaxation and interaction times, respectively, are much smaller than the broadening $\Delta \omega \sim 1.1 \times 10^{14}$ s$^{-1}$ observed experimentally. Thus, the spectrum broadening can only be predicted by the finite length of electron spreading, which corresponds to the coherent length of the electron wave. Here, it is also very interesting to point out that $\Delta \lambda/\Delta \omega = 0.076$ in the experiment in [35] with an excellent agreement to the result shown in Fig. 6 where the spreading length of electron in the order of few tens of micrometers (i.e., $\sim 30 \mu$m). This value of the spreading length also matches with that predicted from the Cherenkov FELS experiments in [27,28].

6. Discussions and conclusions

In this section, we compare our obtained results with those given by Yariv and Shih [36] as a typical treatment for the FELS based on the longitudinal coupling configuration such as Cher- enkov and SP FELS.

In Ref. [36], a periodic dielectric corrugation is used (i.e., $\sigma \approx 0$) and the authors assumed the more efficient interaction with the surface modes of the corrugation (i.e., $G_{st} \gg 1/\tau_{rad}$). Also, the inclusion of the spontaneous emission is neglected (i.e., $G_{sp} = 0$). In the nonrelativistic limit, the exchange power between the electron beam and the EM wave during the interaction time $t$ is given by [36]

$$\Delta P_c = \frac{e^2\omega^2}{2m_0} |E|^2 f(\theta),$$

(99)

where $I = e^2/\hbar^2 \omega^2/\omega^2 (\overline{N}_e) dx dy$ is the electron beam current and $f(\theta)$ is the well-known dispersion function in previous theories of FELS given by

$$f(\theta) = \frac{d}{d\theta} \left( \frac{\sin(\theta/2)}{\theta/2} \right)^2,$$

(100)

where $\theta = \beta_{eq} \sqrt{c - \omega}/(L_x/\nu)$ and $\beta_{eq} = (mG/\pm)\beta$ in the case of the SP effect. Symbols in Eqs. (99) and (100) have the same meanings as in our present paper.

In this work, we state that the spreading length of the coherent electron wave $\ell$ determines the broadness profile of the dispersion function. Assuming that the collisions among electrons do not exist, $\ell$ could be expanded to be the full length of the interaction region $L_x$. This case (i.e., when $\ell \approx L_x$) represents the classical limit at which the electron is described as a point like particle moving without any perturbations from the neighboring electrons. In this limit, the dispersion function given by Eq. (93) is rewritten in terms of the classical dispersion function $f(\theta)$ as

$$D_{2\omega m} = \text{Sinc}^2 \left\{ \frac{\theta}{2} - \Delta \theta \right\} - \text{Sinc}^2 \left\{ \frac{\theta}{2} + \Delta \theta \right\} \approx f(\theta) \times 2\Delta \theta$$

(101)

where $\Delta \theta = (\hbar \omega^2/m_0\nu^2)(L_x/2\nu)$ and represents the electron recoil energy. As experiments move toward the generation of short-wavelength radiation, corrections to the classical dispersion function $f(\theta)$ become significant. In our analysis, the quantum corrections are invoked in the effects of the electron's recoil energy and the finite length of the electron wave as shown in Eqs. (93) and (94). In the emission and absorption processes, the electron's recoil energy is $(\hbar^2\omega^2/2m_0\nu^2)$, which is quite small for most of practical cases. Then, the spreading length of the electron wave is the essential parameter in determining the gain line shape in our quantum mechanical treatment. However, if the electronic recoil is assumed to be zero, no gain is found. Approximately, the term of the electron recoil $(\hbar^2\omega^2/2m_0\nu^2)$ in Eqs. (93) and (94) should be considerable when $\hbar \omega^2/2m_0\nu^2 > 0.001$ (i.e., when $\lambda < 1$ nm in the x-ray range or in shorter range). From Eq. (27), the quantum corrections of the electron recoil could be essential when the electron velocity becomes very close to the speed of light in vacuum (i.e., $\nu_c \sim c$) and $\theta \approx 0$ where SP wavelengths will be shortest at $\theta \approx 0$. To date, the SP FEL operated in the x-ray region cannot be obtained using the current technology. This is because the SP radiation observed in the forward direction (at $\theta \approx 0$) is very weak. Another main technical restriction is that SP x-ray production requires that the electron beam must be steered extremely close to the grating (within few tens of nanometers) without striking its surface. In this paper, it is assumed that counting the electron recoil results in negligible quantum corrections to the gain in comparison with that of the spreading length of the electron wave as discussed in the end of Section 5.

Using Eq. (67) and assuming $G_{st} \gg 1/\tau_{rad}$, $G_{sp} = 0$, and $\sigma = 0$, we can write the radiation power transferred to the SP radiation $\Delta P$ after the transit time $t$ as

$$\Delta P = G_{st}tP.$$  

(102)

Using Eqs. (29), (77), (99), and (100), we find that the power transferred to the radiation $\Delta P$ obtained in our model is one half of that obtained in Ref. [36] such as $\Delta P = (1/2)\Delta P_c$. It is noted that Planck's constant shown in the denominator of Eq. (77) is canceled with that appears in Eq. (101) where $\Delta \theta$ contains Planck's constant in the numerator. Therefore, the quantum correction is mainly invoked in determining the dispersion relation by the coherent length of the electron wave instead of the interaction length.

We end this section by confirming that the coherent length of the electron wave should be adopted to fit the experimental spectral line of the emitted radiation. This length is influenced by the electron density of the electron beam and the initial thermal broadness of the electron velocity. However, in the ideal cases of negligible electron collisions, the coherent length might be equal to the interaction length of the device and our analytical model reduces to the classical one.

In conclusion, in our analysis of the SP radiation based on the small-signal approximation, the electron beam interacts with the surface wave of the metallic corrugation whose propagation velocity is delayed by the corrugated structure. The surface wave produces an EM radiation through the diffraction effect from the corrugated structure. The description of the surface and radiation waves is introduced basing on Maxwell's wave equation. We derive a generalized expression for calculating the low gain of single-pass radiation. In this expression, the profile of the gain distribution is influenced by the finite length of an electron wave and the electron's recoil due to the emission and absorption processes. The compatibility of our analysis with already known classical analyses is confirmed.
Appendix A

To take into account the statistical properties of electrons, let us assign a number $\nu$ for each electron in the electron beam. The expectation value of an arbitrary quantum mechanical operator $\mathcal{R}$ is given by

$$\langle \mathcal{R} \rangle = \sum_{\nu} \rho^{(\nu)}(r, t) |\mathcal{R}| |\psi^{(\nu)}(r, t)\rangle$$  \hfill (A1)

where $\rho^{(\nu)}$ is the probability of finding the $\nu$-th electron in the ensemble where

$$\sum_{\nu} \rho^{(\nu)} = 1.$$  \hfill (A2)

$$|\psi^{(\nu)}(r, t)\rangle$$ is the state vector of the $\nu$-th electron and satisfies the normalization condition

$$\langle \psi^{(\nu)}(r, t) | \psi^{(\nu)}(r, t) \rangle = \int |\psi^{(\nu)}(r, t)\rangle^2 d^3 r = 1.$$  \hfill (A3)

Also, the eigen vector $|\phi_{g}(r)\rangle$ satisfies the normalization and orthogonal conditions

$$|\langle \phi_{g}(r) | \phi_{f}(r) \rangle\rangle = \int |\phi_{g}(r)\rangle^* \phi_{f}(r) d^3 r = \delta_{gf}. \hfill (A4)$$

Using Eq. (39), Eqs. (A3) and (A4), we obtain

$$\sum_{f} |c^{(f)}_{\nu}(t)|^2 = 1.$$  \hfill (A5)

By substituting Eq. (39) into Eq. (A1), the expected value of the operator $\mathcal{R}$ is given by

$$\langle \mathcal{R} \rangle = \sum_{\nu} \sum_{f} \sum_{g} \left\{ \rho^{(\nu)}(t) e^{i\delta^{(\nu)}}(t) e^{i\delta^{(\nu)}}(t) \right\} |\phi_{g}(r)\rangle \langle \phi_{f}(r) | \mathcal{R} | \phi_{f}(r) \rangle.$$  \hfill (A6)

By defining the matrix $\rho$ whose matrix element is given by

$$\rho^{(\nu)}_{gf}(t) = \sum_{\nu} \rho^{(\nu)}(t) e^{i\delta^{(\nu)}}(t) e^{i\delta^{(\nu)}}(t) \delta_{gf}, \hfill (A7)$$

the expectation value of $\mathcal{R}$ can be written as

$$\langle \mathcal{R} \rangle = \sum_{\nu} \sum_{f} \sum_{g} \rho^{(\nu)}_{gf}(t) \mathcal{R} \delta_{gf}. \hfill (A8)$$

By using Eq. (39), we rewrite Eq. (A7) in the form of

$$\rho^{(\nu)}_{gf}(t) = |\langle \psi^{(\nu)} | \phi_{f}(r) \rangle|^2 \left\{ \sum_{\nu} |\psi^{(\nu)}\rangle^* \rho^{(\nu)} |\psi^{(\nu)}\rangle \right\} |\phi_{f}(r)\rangle.$$  \hfill (A9)

Therefore, the abstract density matrix $\rho$ is given by

$$\rho^{(\nu)} = \sum_{\nu} |\psi^{(\nu)}\rangle^* \rho^{(\nu)} \langle \psi^{(\nu)} \rangle. \hfill (A10)$$

The Schrödinger equation of the eigen vector $|\psi^{(\nu)}(r, t)\rangle$ is

$$i \hbar \frac{\partial |\psi^{(\nu)}(r, t)\rangle}{\partial t} = \left( \hat{H} - \frac{i}{\hbar} \frac{\partial}{\partial t} \right) |\psi^{(\nu)} \rangle.$$  \hfill (A11)

where the expectation value of the operator $\hat{I}$ is the electron relaxation time $\tau$.

Using Eq. (A10) with the help of Eq. (A11), the dynamic equation of the density matrix is

$$\frac{d\rho}{dt} = \frac{i}{\hbar} [\hat{H}, \rho] - \frac{1}{2} \{ (\rho - \rho^\dagger) \Gamma + (\rho - \rho^\dagger) \}.$$  \hfill (A12)

The second term on the right-hand side of Eq. (A12) is a phenomenological term that represents the relaxation of electrons to their equilibrium positions.

References