Dispersive effects on the main-wave modulational instability in free-electron lasers

F. B. Rizzato

Institute for Physical Science and Technology, University of Maryland, College Park, Maryland 20742

(Received 26 June 1989; revised manuscript received 1 September 1989)

The main-wave one-dimensional modulational instability due to the dispersive terms of the wave equation in a free-electron laser is introduced and analyzed. We derive the appropriate dispersion relation and compare its associated growth rate with one due to wave-particle energy exchange alone, as obtained by Davidson and Wurtele under the deep trapping assumption [Phys. Fluids 30, 557 (1987)]. It is found that, depending basically on the relation between some characteristic parameters, the modulational instability may be governed by wave dispersion. We also discuss the effects of waveguides on these instabilities and the behavior of the unstable modes as a function of the mentioned characteristic parameters.

It is known that the initial parametric instability in free-electron lasers (FEL) saturates when particles begin to be trapped in the ponderomotive wells formed by the beating of the main electromagnetic wave and the wiggler field.\(^1,2\) As was discussed by Davidson and Wurtele,\(^2\) one essential point is to understand the stability of such a state because, in general, the space-time variation of the main signal may affect its monochromaticity degrading FEL efficiency (of course, we are not talking about a proper tapering of FEL’s parameters, which may enhance its gain\(^3,4\)).

In the present work, we reconsider the problem of the main-wave stability in the sense defined by Davidson and Wurtele.\(^2\) In other words, given an initial steady-state wave, we wish to know if slow perturbations on the phase and amplitude of this wave are unstable functions of time.

An important difference, however, will be introduced. Davidson and Wurtele\(^2\) discarded all the slow second derivatives in Maxwell’s equations for the main wave based on the reasonable (but not always correct) supposition that their effect is smaller than the one due to the first derivatives. In their case, the modulational instability (MI) was originated by energy exchange between waves and particles. In our treatment we will keep these second derivatives. Putting in another way, we may say that our interest will be the analysis of the instabilities that come out as a combined result of these second derivatives and the already mentioned wave-particle energy exchange. The resultant instability will be called modulational instability due to kinetic effects (MIK) when the second derivatives are not important. On the opposite limit, when they are decisive in determining the signal of the perturbation squared frequency, the instability will be referred as due to dispersive effects (MID). Modulational instability due to dispersive effects has been discussed by many authors in many contexts,\(^6,7\) and here we intend to show under which conditions it may dominate the instability processes.

By definition, two electromagnetic waves \(\mathbf{A}\) and \(\mathbf{A}_w\) are present in our system. They are described by the following functions:

\[
\mathbf{A} = \frac{1}{2} \mathbf{a}'(z,y,t)e^{(i(kz -\omega t)) + \text{c.c.}},
\]

\[
\mathbf{A}_w = \frac{1}{2} \mathbf{a}_w'(y)e^{(-k_w z -\omega_{aw} t)} + \text{c.c.},
\]

with \(\mathbf{a}'(z,y,t) = a'(z,y,t)e, \mathbf{a}_w'(y) = a'_w(y)e, \) and \(e = \sqrt{\gamma} \mathbf{x}\) being the vector representing the polarization.

The field \(\mathbf{A}\) is the one corresponding to the main signal and the field \(\mathbf{A}_w\) is the one describing the (given) wiggler. It is seen that as the wiggler wave is moving to the left (or right, depending on the signal of \(k_w), we are not working in the lab frame. It will be clear, however, that the frame we are using (the ponderomotive frame) offers a decisive number of advantages. Transformations to the lab frame are easily done and will not be analyzed here. We also assume that all transverse \((y)\) dependence of the field \(\mathbf{a}\) on the coordinates is due to the presence of a waveguide aligned with our \(z\) axis. The precise three-dimensional structure of this waveguide will not be specified in this paper, although it may be relevant in a variety of situations; in our work, the waveguide will manifest itself basically through the dispersion relation between \(\omega\) and \(k\), which will be derived by assuming the presence of two conducting plates at \(y = \pm \pi/(2k_w)\) and \(a'' = a \cos(k_w y)\), with \(k_w\) representing the transverse scale of our system (which may be equivalent to the longitudinal one).\(^5\) It is still to be noted that we will not be considering the presence of harmonics of the main signal and electrostatic potentials, which may be justified if the plasma is rather tenuous\(^2\) \(\omega_p^2/\omega^2 < 1\) with \(\omega_p^2\) as the plasma frequency of our system.

To analyze the particle orbits, we assume \(ea'/mc^2 < 1\), and use the force law in the form

\[
mdv/\gamma v = -\frac{e}{c} (\mathbf{A}^{\text{tot}} + \mathbf{v} \times \nabla \times \mathbf{A}^{\text{tot}}),
\]

with \(\mathbf{A}^{\text{tot}} = \mathbf{A}_w + \mathbf{A}, \gamma \equiv [1 + (ea'/mc^2)^2]^{1/2}, m\) as the electron rest mass, \(e\) as the electron charge, and \(c\) as the velocity of light.

To solve the above equation we take its \(x\) component, substituting the solution into the longitudinal one. This gives
\[
v_i = -\frac{e}{\gamma mc} A^{\text{tot}} ,
\]
where
\[
\phi_p \equiv \frac{e^2 A}{\gamma^2 mc^2} a'_w \hat{a}' \cos[(k + k_w)z + \sigma] ,
\]
with \(A\) as a number of order 1 that represents the high-frequency effects generated by linearly polarized waves (in circularly polarized systems with \(e = \hat{\mathbf{x}} + i \hat{\mathbf{y}}, \Lambda = 1\)). Here, \(\Lambda\) will not be important; soon it will be absorbed by other physical quantities.

In the set (2), we have defined, without loss of generality, \(a'_w\) as being purely real, and \(a' = \hat{a}' \cos\theta\) with \(\hat{a}'\) and \(\sigma\) as (real) slow functions of their argument \([(\omega, k)/(a', \sigma) \ll (\Delta z, \Delta z)/(a', \sigma)]\). We suppress motions along the \(y\) axis by assuming the fields to produce transverse ponderomotive focusing forces. This can be achieved by using wigglers of the form \(a'_w = a_w \cosh(k_w y)\) (Ref. 9), for example. In that case all the particles are attracted toward the plane \(y = 0\) where all the fields are locally homogeneous. It is precisely this local homogeneity that allows us to consider the \(y\) component of the transverse momentum as a conserved quantity which, besides simplifying the model as much as possible (without affecting its physical content) and to compare it with previous treatments, 2 eventually we will assume deep trapping of particles inside the integration region of Eq. (7) and \(N_0\) is the linear density (constant) of particles along the \(x\) axis inside each ponderomotive well \([N_0\) has a dimension of \((\text{length})^{-1}\)]

With these comments in mind we can write an expression for the density of particles \(n'\):
\[
n' = n(z, t) \delta(y) .
\]

As it is known, we see from the set (2) that the longitudinal motion of the particle is driven by an "effective potential," the longitudinal ponderomotive potential. To simplify the model as much as possible (without affecting its physical content) and to compare it with previous treatments, 2 eventually we will assume deep trapping of particles in the ponderomotive wells. This situation may not be the real one, but may provide very useful information in what follows.

We need now an adequate evolution equation for the field \(A\). To deduce it, we begin with the wave equation for the field \(A\)
\[
[\partial_z^2 - c^2 \partial_t^2 + \nabla^2] A = 4\pi e J ,
\]
where \(J\) is the transverse current,
\[
J = n' v_i .
\]

Writing \(A\) in the form \(1\), we obtain from Eq. (3) and from the assumed forms of \(a', n',\) and \(a'_w\) (upon an integration along the \(y\) axis),
\[
\{ -[\omega^2 - c^2 g(k^2)] - 2i(\omega a' + c^2 k \partial_z a') + \partial_z^2 - c^2 \partial_t^2 \} a
\]
\[
= 4\pi e v_i \left( \frac{1}{2} a'_w \right)_{\omega, k} ,
\]
where \((a'_w)_{\omega, k}\) represents the \((\omega, k)\) Fourier component of the physical quantity \(a'_w\) (computed at \(y = 0\)). The function \(g(k^2)\) incorporates all the transverse effects of this waveguide. It has the form
\[
g(k^2) = k^2 + k_w^2 .
\]

The next step is to calculate the density \(n\) in term of the field \(A\) (and \(A_w\)), following the above rule to select only the components of the product \(a'w\) which are in phase with the main signal. Taking advantage of the periodicity of our system, let us represent \(n\) in the form
\[
n = \sum_{s = -\infty}^{s = +\infty} n_s e^{i(k + k_w)z} .
\]

Each \(n_s\) is to be obtained by assuming its dependence on space to be weak
\[
\frac{2\pi}{|k + k_w|} n_s = \int z^2 - 2iv(k + k_w) c^{-1} s e^{-i(k + k_w)z} dz .
\]

The point \(z\) is one around which we wish to know the value of \(n_s\). Of course, were \(n\) a strictly periodic function, \(n_s\) would be \(z\) independent; here we are only assuming \(\Delta n_s << (k + k_w)n\).

Equation (7) tells us that
\[
\frac{2\pi}{|k + k_w|} n_s = N_0 e^{-i(k + k_w)z} ,
\]
where \(z\), a function of the point \(z\), is the location of the deeply trapped particles inside the integration region of Eq. (7) and \(N_0\) is the linear density (constant) of particles along the \(x\) axis inside each ponderomotive well \([N_0\) has a dimension of \((\text{length})^{-1}\)].

By inserting Eq. (6) in the right-hand side of Eq. (4), and by taking its components in phase with \(A\), we get
\[
\{-\Delta - 2i \left( \frac{\partial_z + V/c}{c} \right) + \partial_z^2 - \partial_t^2 \} a
\]
\[
= -\partial_z^2 a_w e^{-i(a'(k + k_w)/d)||z_0 + z^*} ,
\]
where we are adopting, from now on, the following normalization conditions and definitions: \(t \rightarrow \omega t, z \rightarrow (\omega/c)z, (a, a'_w) \rightarrow (e/mc^2)(a, a_w), \partial_z^2 \equiv (4\pi N_0 e^2/\gamma) \left((k + k_w)/2\pi \omega^2\right)(k_z/2), \Delta \equiv (\omega^2 - g(k^2)) / \omega^2 \partial_z^2, V \equiv c^2 k / \omega,\) and \(z_0\) as the equilibrium position of the particles inside the ponderomotive wells, it being \(z^*\) just the departure from this equilibrium \((z = z_0 + z^*)\). We notice that \(e^{i(k + k_w)/\omega)||z_0 = 1\) (independent of \(z\)). This implies that \(z^*\) must be a slow function of \(z\) which, as we shall see, is true.

As it was said, the amplitude \(a\) depends weakly on \(z\) and \(t\). Let us rewrite this condition
\[
(a, k)a >> \partial_z a > \partial_t a
\]
\[
(\partial_z + V/c \partial_z) a >> \partial_z a, V/c \partial_z a .
\]

In this case, we see that the dispersive term of Eq. (9) begins to play a role only in the situations for which
\[
(\partial_z + V/c \partial_z) a >> \partial_z a, V/c \partial_z a .
\]

Later on we will see when condition (10) may be valid. Right now, let us perform the transformation \((t, z) \rightarrow (\tau, \xi)\) with \(t = \tau\) and \(\xi = z - (V/c)t\). Using the new variables and the slow modulation assumptions together with the condition (10), we may rewrite Eq. (9) in the form
\[-2i\partial_t + D \partial_x^2]a = \partial^2_p a e^{-i(k + k_w)/\omega}z^*, \tag{11}\]

with

\[D \equiv V^2 / c^2 - 1.\]

A special equilibrium for the field \(a (\equiv \partial e^{i\sigma})\) will be analyzed. This state is the one for which its existence is equal to zero, using in its place a derivative of an homogeneous equilibrium phase \(\sigma\) and \(z^*\) equal to zero. The condition for its existence is

\[-\Delta = \partial^2_p a / \partial \sigma_e (> 0). \tag{12}\]

A similar condition as the above was obtained by Davidson and Wurtele, who considered the factor \(\Delta\) as equal to zero, using in its place a derivative of an homogeneous equilibrium phase \(2V/c \partial \sigma_e / \partial z\), to provide the steady-state balance represented in our case by Eq. (12). Actually, in the Appendix we show that the relevant physical quantity that characterizes either the systems with small equilibrium wave vector represented by \(a_e\), and the condition \(\partial \sigma_e / \partial z = 0\) or \(\Delta = 0\) and \(-2V/c \partial \sigma_e / \partial z = \Delta = 0\) and \(-2V/c \partial \sigma_e / \partial z = \partial^2_p a / \partial \sigma_e \) are equivalent.

Now, let us perturb the equilibrium (12) by functions of the form \((\tilde{a}, \tilde{\sigma}) = e^{i(\tilde{K}z - \tilde{\Omega}t)}(\epsilon << 1)\). In this case the linearization of the complex equation (11) (which produces two real equations involving \(\tilde{a}, \tilde{\sigma}\), and \(z^*\)) about the equilibrium state represented by Eq. (12), together with a “deep trapping” version of Eq. (2)

\[\left[ \partial_t - V \partial_x \right] \theta_1 + \omega^2_B (\theta_1 + \theta_2) = 0, \]

with

\[\theta_1 = \frac{k + k_w}{\omega} e^{z*}, \]
\[\omega^2_B = \left( \frac{\epsilon}{\gamma^2} \right)^2 \frac{k + k_w}{\omega} \Lambda a_w \tilde{\sigma}_e, \]
\[\theta_1 = e^{i(\tilde{K}z - \tilde{\Omega}t)} + c.c., \]

gives the following linear dispersion relation between the frequency \(\Omega\) and the wave vector \(K\):

\[4\Omega^2 = (\Delta + DK^2)[\Delta(1 + \alpha) + DK^2], \tag{13}\]

with

\[\alpha = \frac{\omega^2_B}{[\Omega + (V/c)K]^2 - \omega^2_B} = \frac{\omega^2_B}{\delta + 2\Omega (V/c)K}, \]
\[\delta = \frac{V}{c} K^2 - \omega^2_B, \quad \Omega << (V/c)K. \]

Notice that to arrive at the first equation for \(\theta\), we begin by departing from our set (2). Then we expand the argument of the cosine using the deep trapping hypothesis \((k + k_w)z^* + \sigma << 1\). The next step to observe is that the particles \((z^*)\) basically vibrate within the range of one potential well, to convert the total time derivative into a derivative that does not operate on the slow spatial dependence of the relevant functions. Then we use the transformation \((t, z) \rightarrow (\tau, \xi)\) and proceed.

To satisfy requirement (10) we would have to show that \(\Omega^2 / K^2 \) (adimensional quantities) is much smaller than 1. From Eq. (11), we see that the term \(DK^2\) already gives contributions to the relevant frequency \(\Omega\) in the correct scale because \(|D| \leq O(1)\) and \(K \ll 1\). So, in practical terms, what we still must require upon Eq. (13) is that the contribution to the frequency produced by the nondispersive terms be such that \(\Delta^2(1 + \alpha)/K^2 << 1\). By using the definition of \(\alpha\) and the condition \(\Omega^2 << K^2\), we obtain

\[\left| \frac{\Delta^2}{\omega^2_B} \right| << 1. \tag{14}\]

If \(\alpha\) is very small, condition (14) may be satisfied even for large values of \(\Delta / \omega_B\). However, we will see that this situation does not correspond to any type of instability. So, our next step will be to find out under which conditions even large values of \(\alpha\) \((|\alpha| \approx O(1))\) still render relation (14) as a true one. From this relation, it is already possible to see that \(\Delta / \omega_B\) must be small \((\Delta / \omega_B << 1)\); let us show that it is also a sufficient condition.

Without the terms proportional to \(DK^2\), Eq. (13) is just the dispersion relation derived in Ref. 2, under condition (14) (with the corresponding frequency being represented by \(\Omega^2_{K1}\))

\[4\Omega^2_{K1} = \frac{\Delta^2}{\omega^2_B} K^2. \tag{15}\]

With this relation and the definition of \(\alpha\) [Eq. (13)] one can find a restriction on its maximum permissible value

\[|\alpha| < |\alpha_{max}| = |\omega_B / \Delta|^{2/3}, \]
if we consider (from now on) \(V/c \approx O(1)\).

By replacing \(\alpha\) by \(\alpha_{max}\) in relation (14), we find the sought sufficient (and necessary) validity condition of our theory

\[|\Delta / \omega_B| << 1, \]

which defines a particular regime in the FEL sideband instability, the strong pump regime as introduced by Davidson and Wurtele. 

Now let us analyze relation (13) for a wide range of \(K\) variation.

(i) When \(K^2 << \omega^2_B\), we write \(1 + \alpha \approx K^2 / \omega^2_B\) to conclude that we may have an instability of the MHD type (the sign of the complex frequency \(\Omega^2\) depends on that of

\[\Delta / \omega_B\].
\[DK^2 \text{ if } 1 < H < (\omega_B/K)^2, \text{ with } H \equiv |D\omega_B^2/\Delta|. \text{ If } H < 1, \text{ the instability is a MIK one and if } H > (\omega_B/K)^2, \text{ no instability is present.}\]

(ii) When \( K^2 \rightarrow \omega_B^2 \), one may approximate the dispersion relation by

\[\delta + 2(V/c)\omega_B \Omega [4\Omega^2 - (\Delta + D\omega_B^2)D\omega_B^2] = \Delta\omega_B^2(\Delta + D\omega_B^2). \quad (16)\]

Introducing the frequency \( \Omega_0 \) as

\[2\Omega_0 \equiv (\Delta + D\omega_B^2)D\omega_B^2 \]

from Eq. (16), two important unstable regimes may be analyzed.

(a) \( \Omega^2 \approx \Omega_0^2 \). Here, writing \( \Omega \equiv -\text{sgn}(\Delta/D)\Omega_0 + v \) (\(|v| \ll \Omega_0\)), we obtain the maximum growth rate (\( v_{\text{max}} \equiv \Gamma I, \Gamma > 0 \)) in the form

\[\Gamma \approx |D\omega_B^2 / H|^{1/2}, \quad (17)\]

which is valid when (consider the smallness of \( v \))

\[H \gg |\omega_B/\Delta|^{1/2}(> 1).\]

(b) \( |\Omega| \gg |\Omega_0| \). Here, in order to have MID, we see from (16) that \( H \) should be greater than 1 (one more time). The maximum growth rate may be estimated as (considering \( H > 1 \))

\[\Gamma \approx \omega_B |D\Delta|^{1/3} \approx \omega_B \frac{|D\omega_B^2|^{1/3}}{H^{1/3}}, \quad (18)\]

with the instability appearing in the parameter space range

\[1 \ll H \ll |\omega_B/\Delta|^{1/2}. \quad (19)\]

One should notice that this last instability [Eq. (19)], is to be interpreted as a direct extension of the strong pump regime of the MIK. While still that when \( H > 1 \), the value of \( \Gamma \) given by Eq. (18) is much larger than the one calculated without dispersion; in fact, the growth rate (18) is \( H^{1/3} \) times larger than \( \Gamma_{KL} \) [which can be calculated from Eq. (15) with \( \alpha = \alpha_{\text{max}} \)]

(iii) If \( \alpha \rightarrow 0 \) \((K^2 >> \omega_B^2)\), we finally have from relation (13) \( 4\Omega^2 = (\Delta + DK^2)^2 \), which indicates that in this situation the system is stable.

Now, let us recall the two main points derived in this work. First of all, we note that MID may set in only when \( H > 1 \). With this in mind, we see that when \( g(k^2) \rightarrow k^2 \) (the case where wave-guiding effects are small), the instability, in principle, cannot be a MID one. Indeed, in this situation we would have \(-D \rightarrow \delta_B^2 + \Delta \) (take a look at the definition of \( D, V \), and \( \Delta \)), which would produce a value of \( H \) such that \( H >> \omega_B^2 \) (<< 1) if we consider the fact that, in practice, as the amplitude of the laser field is much smaller than the one corresponding to the wiggler, \( \delta_B^2 << |\Delta| \) [see Eq. (12)]. On the other side, when wave guiding is important, \( g(k^2) \neq k^2 \), \( D \approx O(1)<0 \), and \( H \) may be very large, which indeed allows to the possibility of MID. These reasonings also permit to conclude that the dispersive terms have a stabilizing influence on the small frequency range \((K^2 >> \omega_B^2)\); consider the negativity of \( \Delta \) and \( D \) together with point (i) above.

As for the second point, we have seen that at synchrotron resonance \((K^2 \rightarrow \omega_B^2)\), where the most important instabilities are located, two different behaviors for the unstable mode were found to occur. One of them, given by relation (18), is the direct extension of the one analyzed by Davidson and Wurtele. The other one, given by relation (17), has a different structure, it being absent if \( D \) is a very small quantity. In fact, mode (17) is a further extension of mode (18) to even larger values of \( H \). With this in perspective, we may say that as \( \Delta \) tends to small values, beginning from \(|\Delta/\omega_B| << 1\), MI passes by three regimes: kinetic regime (MIK) when \( H < 1 = |\Delta| > |D/\omega_B^2| \), a dispersive regime [given by (18)] when \( 1 \ll H << |\omega_B/\Delta|^{1/2}(= \Delta/\omega_B << |\Delta| << |D/\omega_B^2|) \), and another dispersive regime [given by Eq. (17)] when \( H >> |\omega_B/\Delta|^{1/2}(= |\Delta| << |D/\omega_B^2|) \) (see Fig. 1). One shall be aware of the importance of small \(|\Delta|\) as it is this range that corresponds to strong amplitude laser waves and/or tenuous beams. The precise structure of the waveguide was not taken into account. Calculations including these (and others) effects should be carried out in the future. We think, however, that the conclusions regarding the importance of the usually discarded wave equation's slow second derivatives will not be invalidated anyhow.

The author acknowledges the benefit of useful discussions with Dr. H. P. Freund and the hospitality of Dr. C. S. Wu's group at the Institute for Physical Science and Technology, University of Maryland. This work was supported by Conselho Nacional de Desenvolvimento Cientifico (CNPq), Brazil.
APPENDIX

Let us demonstrate the equivalence between \( \Delta \) and (a possible) \( \partial \sigma_0 / \partial z \). To do so, let us write the field \( a \) in the form

\[
A = \hat{a} e^{i(\lambda z + \sigma) + c.c.} \tag{A1}
\]

where \( \lambda \) represents the small equilibrium phase inhomogeneity (\( \lambda \equiv - \partial \sigma_{eq} / \partial z \), with \( |\partial \sigma_{eq} / \partial z| \ll |(k + k_w) \sigma_{eq}| \)). Then, on applying the left-hand side operator of Eq. (9) on (A1), we get

\[
e^{ikz} \left[ -\Delta - 2V/c \lambda - 2i(\partial_z + V/c \partial_x) + \partial_r^2 - \partial_x^2 \right] \hat{a} = -\partial_r^2 a \left[ e^{-i \frac{2V_c}{c \lambda} \sigma} - i e^{i(\lambda + k_w)/c \lambda \sigma} \right] \tag{A2}
\]

if we disregard small terms of order \( \lambda^2 \). After simplifying \( e^{-i \lambda z} \) with the corresponding dephasage of the equilibrium positions, we obtain an equation similar to Eq. (9), where

\[
\Delta \rightarrow \Delta + (2V/c) \lambda. \tag{A3}
\]

Davidson and Wurtele\(^2\) analyzed \( \Delta = 0 \) and \( \lambda \neq 0 \), and we have studied the case \( \Delta \neq 0 \) and \( \lambda = 0 \). Both representations are equivalent, because from Eq. (A2) or (A3), we see that the relevant physical quantity, which must be used to get both works in contact, is not \( \Delta \) or \( \lambda \) alone, but the sum \( \Delta + (2V/c) \lambda \).

---