FEL

Bernhard Schmidt, SS 2015, revised SS 2018

This is a very brief write-up of my lecture notes. It does NOT substitute reading a text book.

The basic principle of an FEL is energy exchange between an relativistic electron beam and an electromagnetic field inside an undulator (see section on undulators and undulator radiation). In contrast to the "spontaneous undulator radiation" produced by the electrons moving in the static magnetic undulator field, the FEL process requires an external, time dependent, EM field forcing the electrons either to lose or gain energy. Since the energy lost by the electrons goes into a coherent enhancement of the driving EM field, this process can be understood as *induced emission* of photons. The alternative process, the electrons gain energy from the EM field, is a typical absorption process. In both cases, the field energy goes to or comes from the longitudinal kinetic energy of the relativistic electrons, the coupling between field and electrons is mediated by the periodic undulator motion of the electrons.

In the following we will look at the energy exchange process in a rather "classical" view (neglecting the quantized absorption and emission). The energy lost or gained by the electrons is described by their motion in the "ponderomotive potential" of the EM field. The energy gained or lost by the EM field is described by Maxwell's equations, treating the electron beam as time dependent current density. We will show that both energies are of course balanced.

Ponderomotive Phase and Resonance Condition

Key issue of the FEL process is energy exchange between the electrons and the radiation field inside the undulator. This energy exchange comes "on top" of the omnipresent generation of "spontaneous" undulator radiation.

The energy exchange between the purely transverse radiation field E_x and the electrons is due to the transverse velocity component v_x of the electrons in the undulator magnetic field. The energy exchange rate is given by

$$\frac{\mathrm{dW}}{\mathrm{dt}} = -e\, v_x\, E_x(t) = -e\, \frac{K\, c}{\gamma}\, \mathrm{Cos}(k_u\, z)\, E_0\, \mathrm{Cos}(k\, z\, -\, \omega\, t\, +\, \phi_0) \tag{1.1}$$

with k the wave number of the radiation field and k_u the wave number of the undulator.

The product of cosines in eq(1) can be expanded to

$$\frac{\mathrm{dW}}{\mathrm{dt}} = -e\, \frac{K\, c}{2\, \gamma}\, E_0(\mathrm{Cos}[(k + k_u)\, z\, -\, \omega\, t\, +\, \phi_0] + \mathrm{Cos}[(k - k_u)\, z\, -\, \omega\, t\, +\, \phi_0]) \equiv -e\, \frac{K\, c}{2\, \gamma}\, E_0(\mathrm{Cos}[\psi] + \mathrm{Cos}[\chi]) \tag{1.2}$$

While the second Cos is always rapidly oscillating for positive k, the first Cos can be made constant (by proper choice of k) leading to a continuous unidirectional energy exchange.

The phase ψ is normally called "ponderomotive phase", we define it as

$$\psi\, = (k + k_u)\, z\, -\, \omega\, t\, +\, \phi_0 \tag{1.3}$$

The electrons move along the undulator with an average longitudinal velocity $v_z = \overline{\beta}\mathrm{c}$, so the condition for constant ψ reads

$$\frac{d\psi}{dt} = (k + k_u)\, \overline{\beta}\, c - \omega\, = (k + k_u)\, \overline{\beta}\, c - k\, c == 0 \tag{1.4}$$

This defines the "resonant k" of the light wave to be

$$k_{\mathrm{res}}\, = k_u \left(\frac{\overline{\beta}}{1 - \overline{\beta}} \right) \approx k_u \left(\frac{1}{1 - \overline{\beta}} \right) \tag{1.5}$$

The sinusoidal motion in the (planar) undulator field results in an average longitudinal velocity (see notes on undulator radiation)

$$\overline{\beta} = 1 - \frac{1}{2\,\gamma^2}\left(1 + K^2/2\right) \tag{1.6}$$

and the resonant wavelength, leading to a constant ponderomotive phase according to (5), is identical to the wavelength of spontaneous undulator radiation in forward direction:

$$\lambda_{\text{res}} = \frac{\lambda_u}{2\,\gamma^2}\left(1 + K^2/2\right) \tag{1.7}$$

In both cases, spontaneous radiation and FEL energy exchange, this is caused by the requirement that the electrons "slip" against the EM field by exactly *one* wavelength per undulator period.

If eq. (4) is fulfilled, that is ψ=const, the energy exchange oscillates rapidly according to Cos[χ] around Cos[ψ] as shown in Fig. 1. For $\psi = -\pi/2$, the average energy exchange is zero, for $-\pi < \psi < -\pi/2$ is positive (the electrons gain energy from the field) while for $-\pi/2 < \psi < 0$ it is negative, the electrons lose energy to the field.

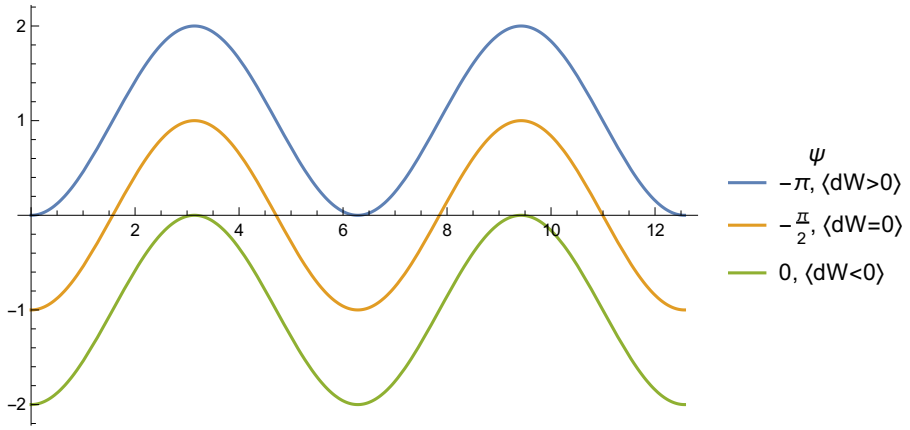


Fig. 1 : Energy exchange between electrons and EM field as function of the rapidly oscillating phase χ for three different ponderomotive phases.

You sometimes read, "the ponderomotive phase describes the longitudinal position of the electron with respect to the phase of the EM wave". This is of course incorrect, it would mean that the electrons move with c. A constant ponderomotive phase means, that the longitudinal position with respect to the wave is *periodic* with the periodlength λ_u, that is the electrons slip backward exactly by one λ_{light} per undulator period. During this slippage, their phase and thus the energy exchanges oscillates as shown in figure 1.

For the following it is more convenient to shift the ponderomotive phase by $\pi/2$ and define $\theta = \psi + \pi/2$ such that for $\theta = 0$ there is no average energy exchange between electrons and EM field.

Since the electron bunches are very long compared to the EM wavelength, all phases are uniformly populated initially. Along the bunch there will be a periodic sequence of areas where the electrons (on average) gain energy and lose energy. On grand average, there is no net energy transfer between bunch and field. But the bunch will develop a periodic *energy modulation* while moving along the undulator. Since the resonance condition for θ = const. (eq. (5)) depends on γ, that is on the energy of the electrons, this energy modulation causes a non-uniform phase shift $\frac{d\theta}{dt}$ along the bunch. The resulting non-uniform population in θ, that is a *density modulation* along z, finally leads to a net energy transfer between electron bunch and EM field. Quantitatively this will be studied in the next section.

Before we continue, a small refinement should be mentioned: Due to the sinusoidal motion of the electrons in the undulator magnetic field there is on top of the average velocity $\overline{\beta}\,c$ a longitudinal oscillation with twice the period k_u. Taking this into account requires some elaborated mathematics (see text books), for the fundamental resonance it leads to a modification of the K parameter by a factor often called "[JJ]", we have to replace

$$K \to K\,[\text{JJ}] = K_{\text{JJ}} = K\left(J_0\left(\frac{K^2}{4 + 2\,K^2}\right) - J_1\left(\frac{K^2}{4 + 2\,K^2}\right)\right) \tag{1.8}$$

with J_0 and J_1 being Bessel functions.

To get a feeling about the strength of the modification as function of K, here a plot:

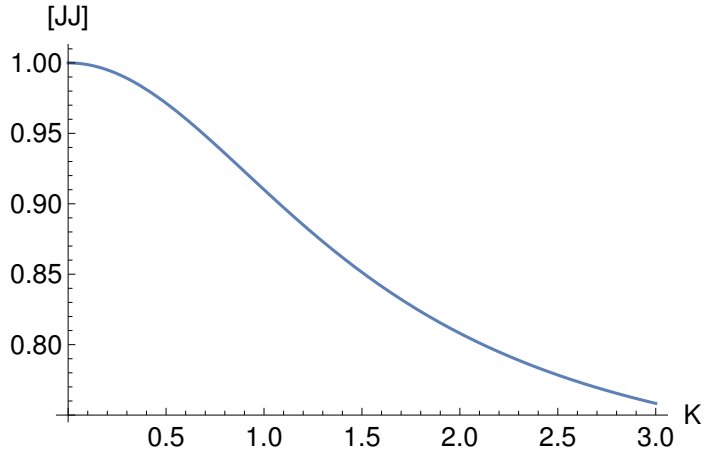


Fig. 2 : Influence of the longitudinal oscillation on the effective K-parameter

For typical $K \approx$ 1values used in an FEL the correction is small.

1-D FEL equations for constant EM field

In the following, we assume that the EM field in the undulator does not change noticeably during the passage of one single electron bunch. This restriction basically limits the length of the undulator. The length scale of this limitation will be derived later. The results obtained in this section apply to a typical "FEL oscillator=FELO", where a short undulator is placed inside a resonator cavity for the EM field. The field in the resonator changes little from the passage of **one** electron bunch but piles up to very high levels after the passage of many bunches. After we got the theoretical model done, we will have a more detailed look into the behavior of such a FELO.

Pendulum Equation and Phase Space

As indicated above, the interaction of the electrons with the EM field induces coupled phase and energy modulations.

So we introduce the relative energy deviation from the "resonant energy"

$$\eta = \frac{\gamma - \gamma_r}{\gamma_r} \tag{1.9}$$

Together with the ponderomotive phase θ, that is the longitudinal position with respect to the phase of the resonant EM field, this pair defines the "longitudinal FEL phase space". The phase space motion with z (the coordinate along the undulator axis) as independent variable is given by the two coupled **pendulum equations**

$$\begin{aligned} \frac{d\eta}{dz} &= -\epsilon \,\mathrm{Sin}(\theta) \\ \frac{d\theta}{dz} &= 2\, k_u\, \eta \end{aligned} \tag{1.10}$$

with k_u the undulator wave number and ϵ

$$\epsilon = \frac{e\, E_0\, K_{\mathrm{JJ}}}{2\, m\, c^2\, \gamma_r^2} \tag{1.11}$$

with E_0 the (assumed constant) EM-field strength and K the undulator parameter.

(The name "pendulum equations" comes from the fact that these are the equations of motion for a rigid pendulum (Schiffschaukel)).

getting the pendulum equations

Phase equation :

We start from eq.(4) switching from $\psi \to \theta$:

$$\frac{d\theta}{dt} = (k + k_u)\,\overline{\beta}\,c - k\,c \tag{1.12}$$

As a short term abbreviation we use

$$X = \frac{1}{2}\left(1 + K^2/2\right) \tag{1.13}$$

and can write from eq (6)

$$\overline{\beta} = 1 - X/\gamma^2 \tag{1.14}$$

and from eq. (7)

$$k = k_u\,\gamma_r^2/X \tag{1.15}$$

Inserting (14) in (12) we get

$$\frac{d\theta}{dt} = c\left(k_u - \frac{(k + k_u)\,X}{\gamma^2}\right) \approx c\left(k_u - \frac{k\,X}{\gamma^2}\right) = c\,k_u\left(1 - \left(\frac{\gamma_r}{\gamma}\right)^2\right) \tag{1.16}$$

where we have made use of $k_u << \mathrm{k}$.

According to (9) we have $\gamma = (\eta + 1)\,\gamma_r$ or

$$\left(\frac{\gamma_r}{\gamma}\right)^2 = (\eta + 1)^{-2} \approx 1 - 2\,\eta \quad \text{for } \eta << 1 \tag{1.17}$$

Which results in

$$\frac{d\theta}{dt} = 2\,c\,k_u\,\eta \tag{1.18}$$

or with dz = c dt finally

$$\frac{d\theta}{dz} = 2\,k_u\,\eta \tag{1.19}$$

Energy equation :

$$\frac{d\eta}{dt} = \frac{1}{\gamma_r}\left(\frac{d\gamma}{dt}\right) = \frac{1}{\gamma_r}\left(\frac{dW}{dt}\right)\frac{1}{m\,c^2} \tag{1.20}$$

The energy exchange rate we get from (2), replace $\theta = \psi + \pi/2$ and neglect the rapidly oscillating term

$$\frac{\mathrm{dW}}{\mathrm{dt}} = -e\,\frac{K\,c}{2\,\gamma_r}\,E_0\,\mathrm{Sin}[\theta] \tag{1.21}$$

and have immediately

$$\frac{d\eta}{dt} = -c\;\frac{e\,K\,E_0}{2\,\gamma_r^2\,\mathrm{mc}^2}\,\mathrm{Sin}[\theta] \tag{1.22}$$

or

$$\frac{d\eta}{dz} = -\,\epsilon\;\mathrm{Sin}[\theta] \tag{1.23}$$

with ϵ from (11) and the refinement for longitudinal oscillations (8).

For later use we write the pendulum equations in Mathematica notation :

```
eqη = D[η[z], z] == - ϵ Sin[θ[z]]
η'[z] == -ϵ Sin[θ[z]]
eqθ = D[θ[z], z] == 2 ku η[z]
θ'[z] == 2 ku η[z]
```

phase space trajectories

The phase space trajectories are determined by a constant Hamiltonian which is given by

$$H = k_u^2 \eta^2 - \epsilon \, \mathrm{Cos}(\theta) \tag{1.24}$$

Check yourself that the Hamilton equations (or "kanonische Gleichungen") lead to the pendulum equations above.

Typical trajectories are shown in Fig. (3) :

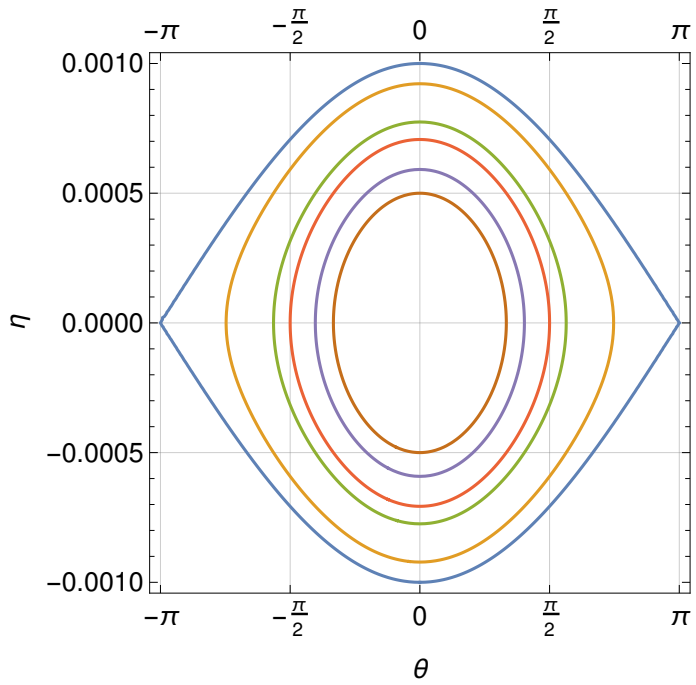


Fig. 3 : Contours of constant H for particle motion in the FEL phase space. The contours describe possible trajectories of the electrons.

The trajectory going through $\theta = \pm\pi$ defines the "separatrix" between trapped and untrapped motion. The maximum H for trapped motion is therefore from $\{\theta{=}\pi,\eta{=}0\}$

$$H_{\mathrm{max}} = \epsilon \tag{1.25}$$

the maximum η allowed for a trapped motion is than

$$\eta_{\mathrm{max}} = \sqrt{\frac{2\,\epsilon}{k_u}} \tag{1.26}$$

The curve of the separatrix is given by

$$\eta_{\mathrm{sep}} = \pm\sqrt{\epsilon\,(1 + \mathrm{Cos}[\theta])\,/\,k_u} \tag{1.27}$$

Note: the separation between bound and unbound (untrapped) motion is determined by ϵ, that is for fixed beam and magnet parameters by the strength of the EM field E_0.

Perturbation approximation

If ϵ is "sufficiently small", the electrons move basically *unbound, way out of the separatrix,* with little change of η before the bunch leaves the undulator. In this case, it makes sense to expand η and θ as power series of ϵ.

Since ϵ is not a dimensionless number (it has the dimension of an inverse length), this is a somewhat fishy approach since we have to define the scale defining "small". This will be done but we postpone it for a moment.

So we write up to second order :

$$\begin{aligned}\theta[z] &\to \theta_0[z] + \epsilon\,\theta_1[z] + \epsilon^2\,\theta_2[z]\\ \eta[z] &\to \eta_0[z] + \epsilon\,\eta_1[z] + \epsilon^2\,\eta_2[z]\end{aligned} \tag{1.28}$$

and similarly for the derivatives..

The equation for η then reads :

$$\eta_0'[z] + \epsilon\,\eta_1'[z] + \epsilon^2\,\eta_2'[z] == -\epsilon\,\text{Sin}\big[\theta_0[z] + \epsilon\,\theta_1[z] + \epsilon^2\,\theta_2[z]\big] \tag{1.29}$$

and for the phase θ :

$$\theta_0'[z] + \epsilon\,\theta_1'[z] + \epsilon^2\,\theta_2'[z] == 2\,k_u\,\big(\eta_0[z] + \epsilon\,\eta_1[z] + \epsilon^2\,\eta_2[z]\big) \tag{1.30}$$

Since ϵ is assumed to be "small", we expand the Sin in the η equation up to second order :

$$\eta_0'[z] + \epsilon\,\eta_1'[z] + \epsilon^2\,\eta_2'[z] == -\epsilon\,\text{Sin}[\theta_0[z]] - \epsilon^2\,\text{Cos}[\theta_0[z]]\,\theta_1[z] \tag{1.31}$$

Since both equations have to hold for all z, they have to hold **for the powers of ϵ individually**. So we solve them with increasing order.

Zero order expansion

This is quite simple and could be seen directly, but we do it formally as for the higher orders.

$$\eta_0'[z] = 0 \tag{1.32}$$

gives immediately

$$\eta_0[z] = \eta_0 \tag{1.33}$$

In zero order, the energy deviation is constant, the initial energy deviation η_0.

$$\theta_0'[z] == 2\,k_u\,\eta_0[z] = 2\,k_u\,\eta_0 \tag{1.34}$$

leads to

$$\theta_0[z] = 2\,k_u\,z\,\eta_0 + \theta_0 \tag{1.35}$$

The phase increases linearly with z, the "phase velocity (difference to c)" is proportional to the initial energy deviation from the resonance energy. Only particles with the resonant γ_r have a constant ponderomotive phase, that is the correct slippage of one λ_{light} per undulator period.
These "zero order terms" describe the behavior without interaction with the EM field (ϵ=0).

First order expansion

Same procedure, we use the result of the zero order :

$$\eta_1'[z] == -\,\text{Sin}[2\,k_u\,z\,\eta_0 + \theta_0] \tag{1.36}$$

what can be integrated to

$$\eta_1[z] \to -\frac{\text{Sin}[k_u\,z\,\eta_0]\,\text{Sin}[k_u\,z\,\eta_0 + \theta_0]}{k_u\,\eta_0} \tag{1.37}$$

Now we define a new variable, scaling the energy deviation with the normalized (to λ_u) length of the undulator L_u

$$\xi_0 = \eta_0\,k_u\,z_{\max} = k_u\,\eta_0\,L_u \tag{1.38}$$

and have

$$\eta_1[L_u] = -\frac{L_u \,\text{Sin}[\xi_o]\,\text{Sin}[\xi_o + \theta_0]}{\xi_o} \tag{1.39}$$

For $\xi_0 = 0$, that is an on-resonance beam, we have $\eta_1[L_u] = -L_u \,\text{Sin}[\theta_0]$. In first order, the bunch develops an sinusoidal **energy modulation** in θ_0.

But, important, on average there is **no net energy loss or gain in first order**, since all initial phases θ_0 are equally populated,

$$\langle \eta_1 \rangle_{\theta_0} = \frac{1}{2\,\pi} \int_{-\pi}^{\pi} \eta_1[z]\, d\theta_0 \;=\; 0 \tag{1.40}$$

Any "gain" (that is net energy exchange with the field), has to be at least second order in ϵ.

The phase development is not so descriptive but we need it for the next step:

$$\theta_1[L_u] = -\frac{1}{2\,\xi_0^2} k_u \,{L_u}^2 \,(\text{Sin}[\theta_0] - \text{Sin}[\theta_0 + 2\,\xi_0] + 2\,\text{Cos}[\theta_0]\,\xi_0) \tag{1.41}$$

and

$$\theta_1'[z] = -\frac{2\,k_u\,L_u\,\text{Sin}[\xi_0]\,\text{Sin}[\theta_0 + \xi_0]}{\xi_0} = \,2\,k_u\,\eta_1[z] \tag{1.42}$$

Second order expansion

Same procedure, we use the result of the zero and first order and get :

$$\eta_2'[z] == \frac{1}{2\,k_u\,\eta_0^2}\,\text{Cos}[2\,k_u\,z\,\eta_0 + \theta_0]\,(\text{Sin}[\theta_0] - \text{Sin}[2\,k_u\,z\,\eta_0 + \theta_0] + 2\,k_u\,z\,\text{Cos}[\theta_0]\,\eta_0) \tag{1.43}$$

This is not so obvious to integrate, finally *Mathematica* pays off...

$$\eta_2[z] = \frac{1}{16\,{k_u}^2\,\eta_0^3}\,(-4 + 4\,\text{Cos}[2\,k_u\,z\,\eta_0] - \text{Cos}[2\,\theta_0] + \text{Cos}[2\,(2\,k_u\,z\,\eta_0 + \theta_0)] + 8\,k_u\,z\,\text{Cos}[\theta_0]\,\text{Sin}[2\,k_u\,z\,\eta_0 + \theta_0]\,\eta_0) \tag{1.44}$$

what we can rewrite for $z = L_u$ using the ξ_o parameter as

$$\eta_2[L_u] = \frac{1}{16\,\xi_0^3} k_u \,{L_u}^3\,(-4 - \text{Cos}[2\,\theta_0] + 4\,\text{Cos}[2\,\xi_0] + \text{Cos}[2\,(\theta_0 + 2\,\xi_0)] + 8\,\text{Cos}[\theta_0]\,\text{Sin}[\theta_0 + 2\,\xi_0]\,\xi_0) \tag{1.45}$$

The average energy exchange does *not* vanish if $\xi_o \neq 0$...

$$\langle \eta_2 \rangle_{\theta_0} = \frac{1}{2\,\pi} \int_{-\pi}^{\pi} \eta_2[z]\, d\theta_0 = \frac{k_u\,{L_u}^3\,\epsilon^2\,(\xi_o\,\text{Cos}[\xi_o] - \text{Sin}[\xi_o])\,\text{Sin}[\xi_o]}{2\,{\xi_o}^3} \tag{1.46}$$

Thus, **FEL - energy exchange is a second order effect** with respect to the strength of the external EM field, it scales with ϵ^2.

Since the energy density of the initial EM field is ~ ${E_0}^2$ and thus to ϵ^2, the energy exchange is proportional to the present EM energy density and it makes sense to talk about "FEL Gain". Furthermore, since ${E_0}^2$ is proportional to the photon density in the field, the understanding as *induced emission* makes sense: the number of produced photons is proportional to the number of already existing photons.

Notice that for $\xi_0 = 0$, there is no net energy exchange in this model, FEL gain needs a finite energy offset from resonance. We will later, in a more refined model, learn that this is an approximation which has to be revised in the case of an long-undulator aka high-gain FEL.

The ξ_o dependence of eq. (46) can be written in a more elegant form as the derivative of a Sinc function :

$$\partial_{\xi_o}\left(\frac{\text{Sin}[\xi_o]}{\xi_o}\right)^2 = \frac{2\,(\xi_o\,\text{Cos}[\xi_o] - \text{Sin}[\xi_o])\,\text{Sin}[\xi_o]}{{\xi_o}^3} \tag{1.47}$$

so we have finally the net energy exchange (the "gain curve") in the short undulator as function of the initial energy deviation ξ_0

$$\langle \eta \rangle_{\theta_0} = -\frac{k_u}{4}\,\epsilon^2\,L_u^3\,\partial_{\xi_o}\left(\frac{\text{Sin}[\xi_o]}{\xi_o}\right)^2 \tag{1.48}$$

In fact this result is to some extend inconsistent with the initial assumption of the model, namely that the electric field does not change during the passage of one bunch. Now we see that the bunch has an net energy change and this inevitably leads as well to a change of the field. We will see later how this inconsistency vanishes if E is allowed to vary and that the energy is actually balanced.

Madey theorem, the Small Signal Gain (SSG) curve

In an undulator, the frequency of the radiation is $\sim 1/\gamma^2$, therefore the relative energy deviation η_0 is related to the relative frequency deviation (small numbers) by

$$\Delta\omega/\omega_r = -2\,\Delta\gamma/\gamma_r = -2\,\eta_0 \tag{1.49}$$

thus

$$\Delta\omega/\omega_r = \xi_o/(\pi\,N_u) \tag{1.50}$$

From the lecture on undulator radiation, we learned that the spectral line shape of undulator radiation given by a Sinc^2 function in ξ_o

$$I[\omega] \sim \left(\frac{\mathrm{Sin}\,(\xi_o)}{\xi_o}\right)^2 \tag{1.51}$$

So we see the **Madey theorem**: the shape of the gain curve of the low gain FEL is the derivative of the spectral shape of the spontaneous undulator radiation.

How does this gain depend on the initial energy deviation, now measured by ξ_0 ? As mentioned above, there is no energy exchange for $\xi = 0$. Furthermore the gain curve is anti-symmetric in ξ and looks like shown in Fig. 4.

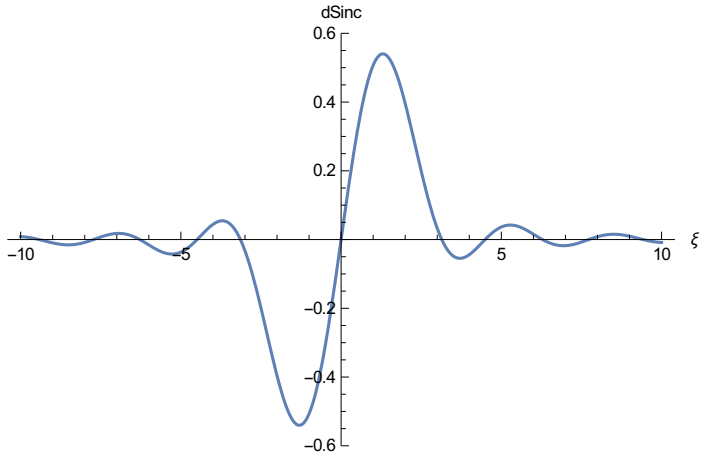


Fig. 4 : The universal shape of the Madey gain curve describing the energy exchange in a short undulator (low gain) FEL.

Notice that this function is universal in ξ, the relative energy deviation η_0 times 2π the number of undulator periods. The maximum gain is achieved at about $\xi \sim 1.3$, corresponding to a relative energy deviation of $\eta_0 \approx \frac{1}{5\,N_u}$. The sign of the curve is such that positive values indicate energy transferred to the field, that is "positive gain". To achieve this, the energy of the incoming beam has to be slightly above resonance energy. The reason for that will become obvious when we discuss the phase space motion of the electrons in more detail. But first we write down the total energy transferred from the beam, that is the power gain of the FEL amplifier.

Total energy loss or gain

To get the net change of beam energy we have to remember that η is the *relative* energy change for *one* electron. To get the total change of the **energy density** we have to multiply $\langle\Delta\eta\rangle$ with the beam energy $(\gamma_r\,m\,c^2)$ and the electron number density n_e using the definition of ϵ from above :

$$\Delta \mathrm{W}_{\mathrm{beam}} = \frac{e^2\,k_u\,L_u{}^3\,E_0^2\,K_{\mathrm{JJ}}^2\,n_e}{16\,m\,c^2\,\gamma_r^3}\,\partial_\xi\left(\frac{\mathrm{Sin}[\xi]}{\xi}\right)^2 \tag{1.52}$$

Now we normalize to the energy density of the EM field $\left(W_{\mathrm{EM}} = \frac{1}{2}\,\epsilon_0\,E_o{}^2\right)$and define the "gain" as ratio of both

$$G = \frac{\Delta W_{\text{beam}}}{W_{\text{EM}}} = \frac{e^2\, {L_u}^3\, \pi\ K_{\text{JJ}}^2\, n_e}{4\, m\, c^2\, \gamma_r^3\, \epsilon_0\, \lambda_u}\, \partial_\xi \left(\frac{\text{Sin}[\xi]}{\xi}\right)^2 \tag{1.53}$$

Another way of expressing the result is to express the electron density n_e by the total beam current I_b. The electron density is related to the total beam current I_b and the cross section area of the beam $A_b = 2\,\pi\,{\sigma_b}^2$ by $n_e = I_b/(e\,c\,A_b)$, so we have

$$G = \frac{e\, {L_u}^3\, K_{\text{JJ}}^2\, I_b}{8\, c^3\, m\, \gamma_r^3\, \epsilon_0\, \lambda_u\, \sigma_b^2}\, \partial_\xi \left(\frac{\text{Sin}[\xi]}{\xi}\right)^2 \tag{1.54}$$

This formula shows the importance of a high beam current and small transverse cross section of the beam. (Footnote: the relation of σ_b and A_b is made such that for the Gaussian beam of $\sigma_x = \sigma_y = \sigma_b$ the integrated charge is peak value times A_b).

Another abbreviation which is used quite often is the Alfvén current $I_A = 4\,\pi\,\epsilon_0\, m\, c^3 \,/\, e \approx 17$ kA , so we rewrite

$$G = \frac{{L_u}^3\, \pi\ K_{\text{JJ}}^2\, I_b}{2\, \gamma_r^3\, I_A\, \lambda_u\, \sigma_b^2}\, \partial_\xi \left(\frac{\text{Sin}[\xi]}{\xi}\right)^2 \tag{1.55}$$

or as product of dimensionless numbers, replacing the undulator length $L_u \to N_u\, \lambda_u$:

$$G = \frac{\pi}{2}\, \frac{K_{\text{JJ}}^2\, N_u^3}{{\gamma_r}^3} \left(\frac{I_b}{I_A}\right) \left(\frac{\lambda_u}{\sigma_b}\right)^2 \partial_\xi \left(\frac{\text{Sin}[\xi]}{\xi}\right)^2 \tag{1.56}$$

Remark: this gain definition assumes a perfect lateral overlap of the EM field and the electron beam. Any "mismatch" results in a gain degradation. To achieve this perfect overlap, the beam profile (defined by the β-functions and the emittance of the beam) and the radiation profile (defined by the Rayleigh length of the field in the resonator) have to be matched properly.

Please notice that the "Alfvén current" does not play the rôle of a "current limit" or otherwise defining a "scale" in FEL physics. It is a useful short form for a combination of natural constants, not more. (The Alfvén current plays an important role in low energy plasma physics where it defines the scale of an instability for electron currents in a neutral background plasma. Under these conditions, the repulsive electric force of the electrons is perfectly balanced by the ion background. The remaining magnetic force, due to the current, leads to an instability for currents above I_A. For a relativistic electron beam in vacuum (as we have in the FEL), the electric and magnetic forces are perfectly balanced up to a level of $\frac{1}{\gamma^2}$ and such an effect does not exist).

Limits of the "small signal gain" (SSG) curve

This "Madey" gain curve, or SSG curve, has been derived using a power series expansion in ϵ, it is valid only **as long as the EM field is sufficiently "small"**.

Now finally we have to define what small means. To get rid of the dimensional quantity ϵ, we normalize z to the length of the undulator L_u, setting $\tilde{z} = z/L_u$. The first pendulum equation now reads

$$\frac{d\theta}{d\tilde{z}} = L_u\, \frac{d\theta}{dz} = 2\, k_u\, L_u\, \eta = 2\,\xi \tag{1.57}$$

The equation for ξ gets

$$\frac{d\xi}{d\tilde{z}} = k_u\, L_u\, \frac{dz}{d\tilde{z}}\, \frac{d\eta}{d\tilde{z}} = k_u\, {L_u}^2\, \frac{d\eta}{d\tilde{z}} = -\left(k_u\, {L_u}^2\, \epsilon\right) \text{Sin}(\theta) \equiv -\tilde{\epsilon}\, \text{Sin}(\theta) \tag{1.58}$$

The "expansion parameter" is now $\tilde{\epsilon} = \left(k_u\, {L_u}^2\, \epsilon\right)$, a dimensionless number. Thus, for the Madey curve to be applicable, we request that $\tilde{\epsilon} < 1$ or

$$\epsilon < \frac{1}{k_u\, {L_u}^2} = \frac{\lambda_u}{2\,\pi\, {L_u}^2} \tag{1.59}$$

holds.

If we put this to the equation of the separatrix given in eq. (26) we have

$$\eta_{\text{sep}} << \frac{\sqrt{2}}{k_u\, L_u} \tag{1.60}$$

while the optimum gain is achieved for $\eta_0 \approx 1.3\,/(k_u\, L_u)$.

We see that the Madey curve assumes for the initial energy offset η_0 typical values well above the separatrix limit, that is *electrons which are not trapped* to the FEL bucket.

Modelling the "low gain FEL"

The typical application for a "low gain" oscillator FEL is in the near and far IR regime (say "few" to a "few hundred" micrometer). The advantages (compared to other laser techniques) are the wide and continuously tunable wavelength and the high average power achievable. Some examples compared to other light sources are compiled in figure 5.

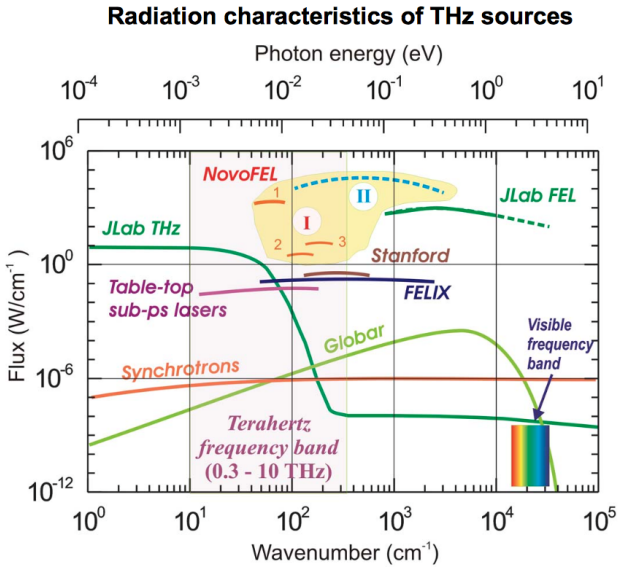


Figure 5: Overview on radiation sources in the IR and THz regime (from http://epaper.kek.jp/f06/TALKS/TUCAU03_TALK.PDF).
FELIX is an FEL facility in the Netherlands, novoFEL at Novosibirsk (Russia), JLAB at Jefferson Lab (USA). A large installation in Germany is the ELBE facility at HZDR, Dresden-Rossendorf.

A typical FELO set-up is shown in Figure 6: a linear accelerator generates a high current electron beam which drives the FEL process in an undulator placed inside the optical cavity. The spent beam is either dumped or (at high average power applications) re-circulated through the RF resonators. The re-circulated phase is adjusted such that the beam is de-celerated and the beam power "recycled" into the RF resonators. Such "energy recovery linacs" are able to produce average electron beam powers of several hundred kilowatts.

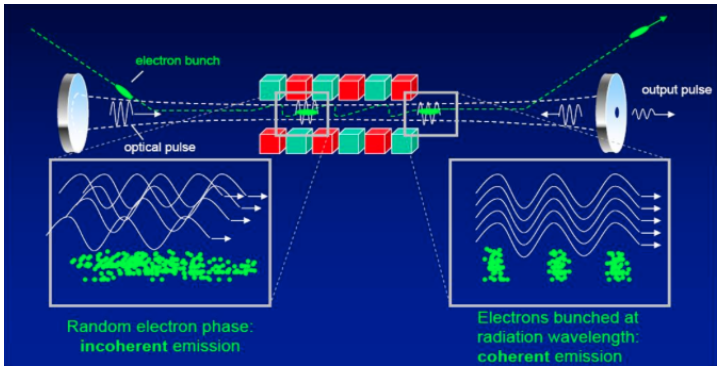


Figure 6 : Schematic set-up of an resonator FEL (from http://www.stfc.ac.uk/astec/17452.aspx).

The start-up of the FEL process is typically from "noise", that is from the initial spontaneous undulator radiation produced by the first bunches entering the undulator. This field is amplified by the subsequent bunches, the field-increase per bunch is proportional to the already present field. This process continues until the energy transferred from the electrons just compensates the resonator losses, that is mainly the out-coupled user beam. In this state, the FEL oscillator can run in a continuous mode. The time structure of the EM radiation is that of the electron beam, the length of the "wave packets" is determined by the length of the electron bunches and the repetition rate by their distance. To optimize the overlap between the bunches and the

field, the Rayleigh length of the optical resonator is matched to the transverse size of the electron beam which is determined by the emittance of the beam and the β-function of the beam optics.

The fundamental assumption made in the description so far is that the field inside the resonator does not change substantially during one passage of a bunch both in amplitude and phase. Under these assumptions, we can **model the FEL process in an 1-D semi analytical model** based on the pendulum equations (10) as follows:

- we start with a decent number of electrons entering the undulator with fixed energy offset η_0 and equally distributed phases
- For these initial conditions, the pendulum equations are solved numerically and evaluated at the desired length of the undulator
- By averaging the resulting energy offsets $\langle \eta_i \rangle$ and subtracting $\eta_0$we get the energy loss or gain of the electron beam which we claim to be transferred to the EM field

(In fact this approach is a bit fishy since on the other hand we claim the electric field not to change. The key word is "substantially", the change has to be small enough. The model is "1-D" since we do not consider the transverse structure of the beam and the field. It is semi-analytical since it is based on numerical solving the differential equations).

For the following examples, we use the parameters of the IR- FEL "FELIX", summarized in the following table :

λ_L	λ_u	K	N_u	$\sigma_{x,y}$	σ_z	L_u	K_{JJ}	γ_r	η_{opt}	E_b	I_b	Q_b
20 μm	65 mm	0.5	38	2 mm	0.9 mm	2.47 m	0.4857	42.76	5.4×10^{-3}	22 MeV	57.3 A	172 pC

Table 1: Some typical parameters of the infrared oscillator FELIX at Radboud University Nijmegen. The wavelength of 20 μm is an example, the FEL can operate between 3 μm and about 130 μm.

Modelling the FEL gain as described above leads to the "gain curves" shown in figure (8) as dots and compared to the "SSG" curve of the Madey formula. For low enough initial field (here 1 MV/m), the SSG curve is a very good approximation while for higher initial fields, deviations become clearly visible. The optimum ξ=1.3 for low fields corresponds to a relative energy deviation η_0 of about 0.54 % in this case (equation 38).

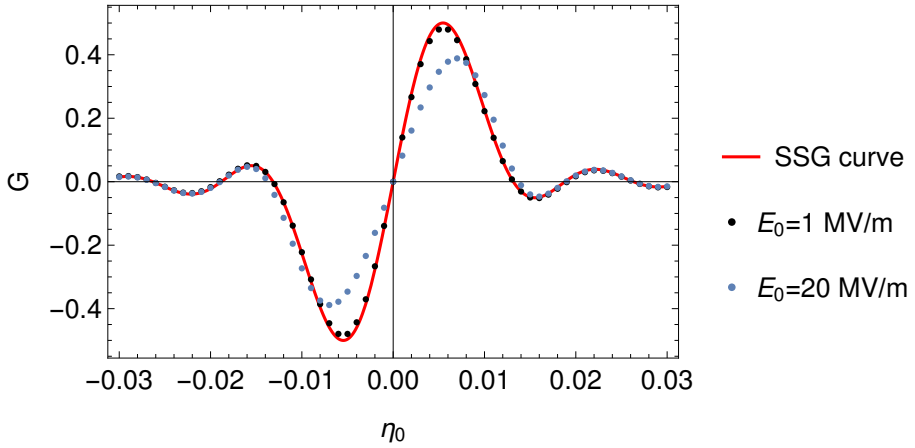


Fig. 8: FEL gain from a 1-D simulation using the pendulum equations for two different initial field strengths. The solid line represents the Madey formula.

To understand why there is no "gain" for $\eta_0 = 0$ (on-resonance beam) we look a bit more in the evolution of energy and density modulations, figure (9). We see the energy modulation increasing linearly with the undulator length and a density modulation growing quadratically. The zero crossing of the EM field is at θ=0, that is the red and blue marked segments indicate electrons either loosing (blue) or gaining (red) energy. Notice that the density modulation and the energy modulation are phase shifted by $\pi/2$. Since the electrons are exactly on resonance, their ponderomotive phase is constant and the number of loosing and gaining electrons is always balanced. There is no net energy transfer.

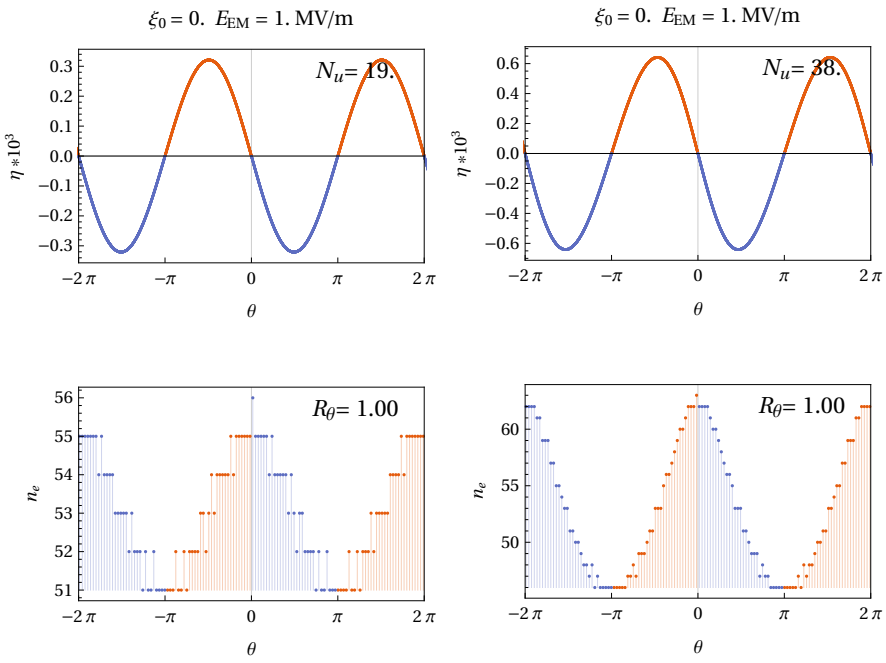


Figure 9: Energy and density distribution for a bunch with resonant energy after 19 and 38 undulator periods and an initial field of 1 MV/m. The red and blue areas mark electrons gaining or loosing energy from (to) the field. The number R_θ is the ratio of both. R_θ = 1 indicates perfect balance.

In Figure (10), the initial energy offset is at the optimum ξ=1.3 according to the Madey curve. Since the electrons are off-resonance, their ponderomotive phase evolves linearly with time. The density modulation gets phase shifted such that the electrons loosing energy outbalance the electrons gaining energy. There is a net energy transfer from the beam to the EM field.

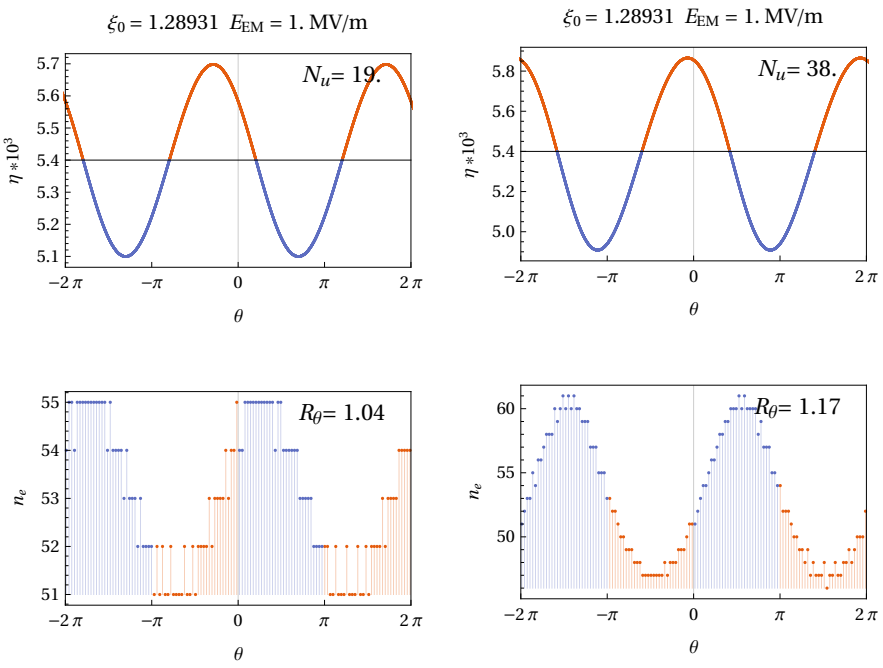


Figure 10: Energy and density distribution for a bunch with optimum initial η after 19 and 38 undulator periods and an initial field of 1 MV/m. The red and blue areas mark electrons gaining or loosing energy from (to) the field. The number R_θ is the ratio of both. Due to the off-resonance condition, the phase is not constant but shifts such that the electrons loosing energy (blue) overbalance the gaining ones (red).

From what we see in Figure (10), it is clear that there is an optimal undulator length for this process. If the phase shift increases further (as it does), the density distribution first gets balanced again (when a phase shift of π is reached) and finally "inverted". The net energy now goes from the beam to the electrons, the gain drops again. (Figure (11)). In this specific case (parameters), the gain would rise up to about 65 undulator periods before it starts degrading again.

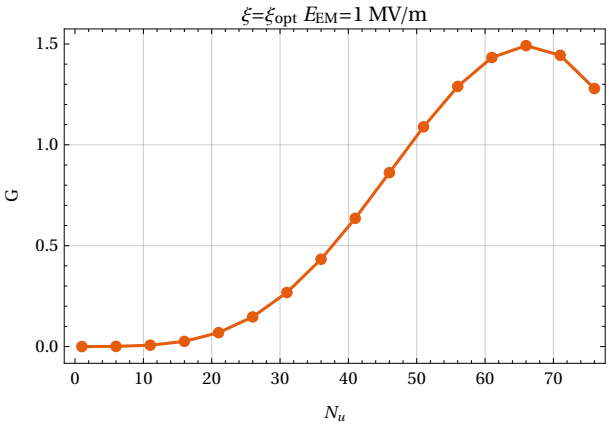


Figure 11: Gain as function of depth in the undulator for an initial field of 1 MV/m. The gain (the total amplification of the energy EM energy density) reaches a maximum after about 65 undulator periods in this case. For longer undulators, the linear phase shift due to the initial energy deviation "out-phases" the density modulation, the beam re-gains the energy from the field.

(Please notice the FEL people use a bit peculiar defintion of "gain". It is defined as G = $\frac{P_{out}-Pin}{P_{in}}$).

Up to here, we looked at just one bunch passing the undulator. In a FEL oscillator, many subsequent bunches enter the undulator and see the EM field produced by their predecessors. Since the energy transfer is proportional to the existing field, the field in the resonator will initially grow exponentially with the number of bunches having passed the undulator. The resonator has a well defined loss factor, normally dominated by the fraction F_u out-coupled to the "user beam". Thus the field intensity inside

the resonator reaches its saturation level when the gain just balances the losses, that is $G = F_u$. For our example, we use an out-coupling fraction of $F_u = 0.1$ (see figure 12).

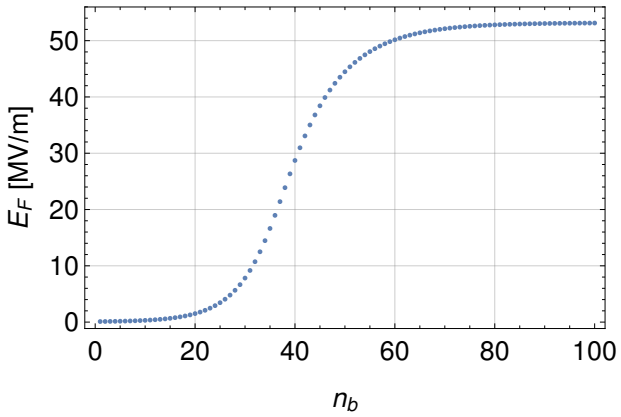


Figure 12: Electric field strength in the resonator as function of the number of bunches passed. The out-coupled power fraction was set to 10%, the undulator has 38 undulator periods. After about 70 bunches, the saturation level of 53 MV/m is reached, gain and out-coupling are balanced.

Figure 12 shows the "field build up" in the resonator as function of the number of bunches passed. After about 70 bunches, the saturation level of 53 MV/m is reached and gain and out-coupling are balanced.

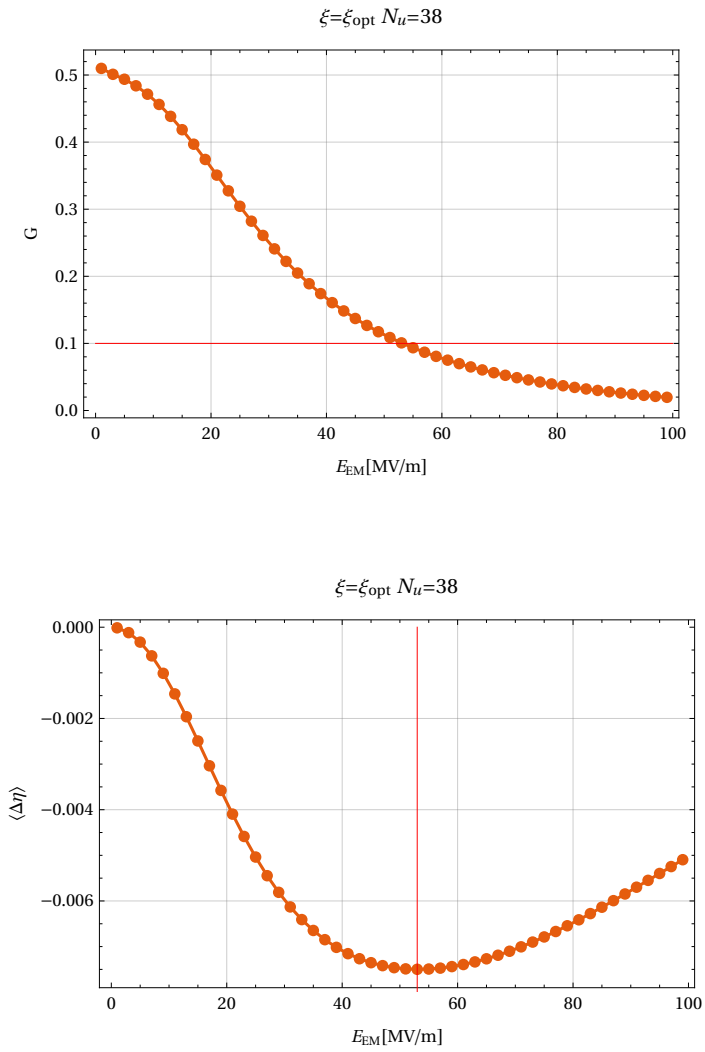


Figure 13: Electric field strength in the resonator as function of the number of bunches passed. The out-coupled power fraction was set to 10%, the undulator has 38 undulator periods. After about 70 bunches, the saturation level of 53 MV/m is reached, gain and out-coupling are balanced (red line).

In Figure 13 we look at the gain (relative increase of EM power) and the energy loss from the beam as function of the electric field inside the resonator. The gain drops from its initial value of 0.5 (defined by the undulator and beam parameters) with increasing field, at 53 MV/m the out-couple balance of 0.1 is reached. Looking at the relative energy transfer $\langle \Delta\eta \rangle$, we see that the undulator length is chosen such that the maximum energy transfer is reached close to this saturation field. This maximum energy loss of the beam is of the order 0.8 %. We will learn later that this is a typical number for both, "low gain" and "high gain" FELs.

Notice that these plots are based on a model using the pendulum equations and ignoring the evolution of the electric field during the passage of one bunch. If this is validate and how to overcome these limitations will be discussed in the next section.

To complete the discussion we look again at the "phase space evolution" of the bunches in our FEL oscillator. In Figure (10) we already saw the energy and density modulation for one of the "first" bunches, that is for a low initial field of 1 MV/m. For bunches at stable saturation level (the is the operation point of the FELO), the interaction between beam and field is much more pronounced (see Figure (14)). After about half the undulator depth ($N_u = 19$), the bunch has developed an extremely pronounced density modulation where a large fraction of the electrons is concentrated in narrow "slices" at phases where they transfer energy to the EM field. These "slices" are called "micro-bunches". The beam is structured now in a way that the number of electrons doing "induced emission" is substantially larger than those doing "absorption". Please notice that the same slices at the wrong phase (shifted by π) would produce the opposite effect: a strong dis-balance in favor of "absorption", the beam would gain energy from the field.

At the end of the undulator ($N_u = 38$), the electrons have "over-rotated", that is moved backward in phase. The "micro-bunching" has been destroyed and the phase distribution is almost balanced again.

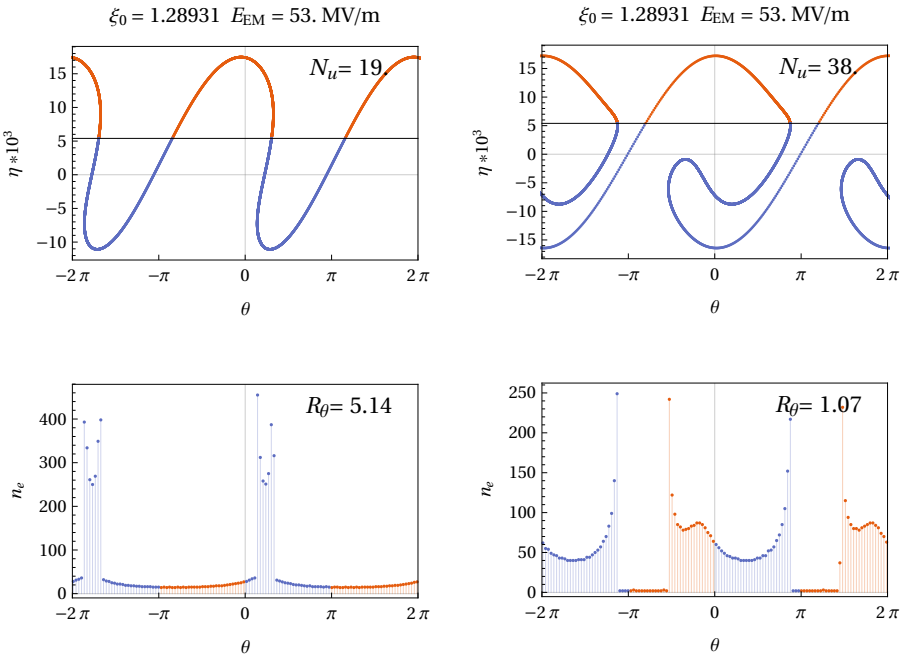


Figure 14: Electric field strength in the resonator as function of the number of bunches passed. The out-coupled power fraction was set to 10%, the undulator has 38 undulator periods. After about 70 bunches, the saturation level of 53 MV/m is reached, gain and out-coupling are balanced (red line).

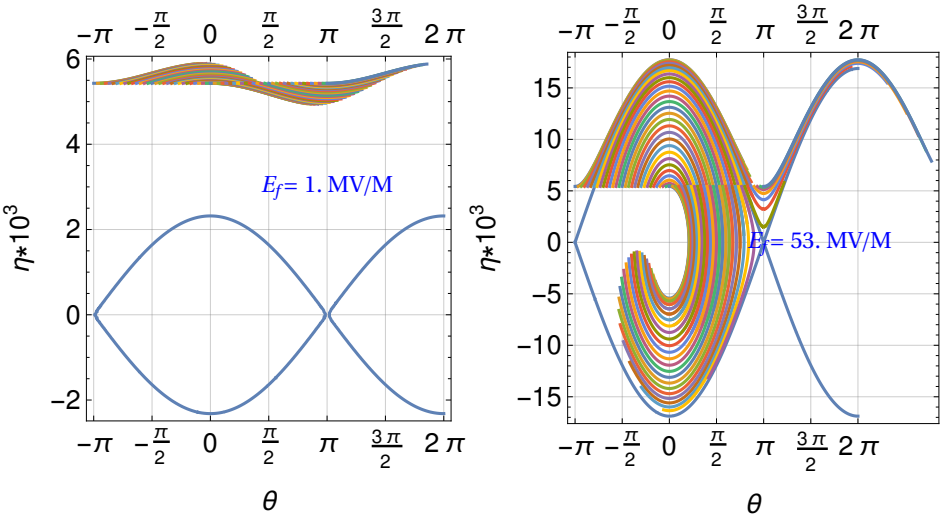


Figure 15: Phase space trajectories of the electrons in our model FELO.

Figure (15) finally shows the phase space trajectories of the electrons for a low initial field and at saturation (operation) level of the FEL, in both cases starting from the optimal η_0 according to the Madey curve and equally distributed along the ponderomotive phase θ. At low initial fields, the electrons are well outside the separatrix and basically move freely. In fact the Madey curve is an approximation for $E_f = 0$. At the saturation field, the electrons are almost completely trapped inside the separatrix. The total energy loss is much higher than at low fields, but the "gain" (the relativ energy loss compared to the exisiting field) is substantially reduced.

Notice for later: the "gain" obviously depends both, the energy loss of the electrons and present energy in the EM wave. The "high gain FEL", discussed below, has about the same total energy loss of the electrons but the *initial* EM field energy (when the bunches enter the undulator) is very small.

1-D FEL equations including field change

The following way to the "coupled equations" of a 1-D FEL theory follows exactly the lecture notes by Kwang-Je Kim (ANL), Zhirong Huang (SLAC) and Ryan Lindberg (ANL).

Up to now, the interaction between the field and the electrons has been reduced to the electron energy changing, as described in equation (2), while the electrons pass the undulator. This is insufficient for long undulators, here we have to consider that the field is influenced by the interaction with the electrons as well. Not only its amplitude increases, also its phase shifts with respect to the beam modulation. This phase slippage is an important asset to understand why and how a "long undulator FEL" works and a sustained energy transfer from the beam can be achieved.

As above, we neglect any transverse dependence of the beam and field, that is we reduce the problem to 1 dimension, the longitudinal coordinate z along the undulator. Beam and field are treated as infinite and uniform radially. The influence of 3-D effects will be briefly mentioned at the end of the discussion, for more details you a referred to the well known blue text book.

As further simplification, we restrict the EM field to only one fixed wavelength (wave-number k) which is present as "initial field" when the bunch enters the undulator. The interaction of the field with the electrons is described by Maxwell's wave equation:

$$\left(\frac{1}{c^2}\left(\frac{\partial^2}{\partial t^2}\right) - \frac{\partial^2}{\partial z^2}\right) E_x(z,\ t) = -\left(\frac{1}{\epsilon_0\, c^2}\right)\frac{\partial\, j_x(z,\ t)}{\partial\, t} \tag{1.61}$$

with $j_x(z,\ t)$ the *transverse* current density caused by the undulator motion of the electrons. We further assume the problem to be (locally) periodic in the ponderomotive phase, that is with a period length of $\lambda = 2\pi/k$.

The (transverse) electric field will oscillate rapidly while amplitude and phase vary "slowly" due to the interaction with the electrons (index "s" means "slowly varying, SVA = slowly varying amplitude approximation). So we write :

$$E_x\,(z,\ t)\ =\ E_s\,(z,\ t)\,\text{Cos}[k\,z\ -\ \omega\ t\ +\ \phi_s\,(z,\ t)] \tag{1.62}$$

We define a (SV) complex amplitude function

$$\tilde{E}(z,\ t)\ =\ \frac{1}{2}\,E_s(z,\ t)\,e^{i\,\phi_s(z,t)} \tag{1.63}$$

and rewrite

$$E_x\,(z,\ t)\ =\ \tilde{E}(z,\ t)\,e^{i(k\,z-\omega\,t)}\ +\ \tilde{E}^*(z,\ t)\,e^{-i(k\,z-\omega\,t)} \tag{1.64}$$

Furthermore, we decompose the wave operator

$$\frac{1}{c^2}\left(\frac{\partial^2}{\partial\, t^2}\right) - \frac{\partial^2}{\partial\, z^2} =\ D_+\ D_- \qquad \text{with}\quad D_\pm =\ \frac{1}{c}\left(\frac{\partial}{\partial\, t}\right) \pm \frac{\partial}{\partial\, z} \tag{1.65}$$

"Slowly varying amplitude" means that it does not change substantially on the scale of one wavelength, that is

$$\left|D_\pm\,\tilde{E}\right|\,\lambda << |E_s| \tag{1.66}$$

or

$$\left|D_\pm\,\tilde{E}\right| << |E_s|\ k \tag{1.67}$$

and so on..

We consider that

$$D_-\ e^{\pm i(k\,z-\omega\,t)} = \mp 2\ i\ k\ e^{\pm i(k\,z-\omega\,t)} \tag{1.68}$$

while

$$D_+\ e^{\pm i(k\,z-\omega\,t)} = 0 \tag{1.69}$$

(The reason for the asymmetry is, that we only consider a wave traveling in forward $+z$ - direction. For the general Ansatz, a

sum of two opposite waves, the wave operator would be symmetric. The backward wave can be neglected since it is not amplified).

Now we get

$$D_- E_x(z,\ t) = \left(-2\ i\ k\ \tilde{E} + D_-\ \tilde{E}\right) e^{i(k\,z-\omega\,t)} + \left(2\ i\ k\ \tilde{E}^* + D_-\ \tilde{E}^*\right) e^{-i(k\,z-\omega\,t)} \tag{1.70}$$

Since we assume that $\tilde{E}$ is slowly varying, we use this to disregard the derivatives getting

$$D_- E_x(z,\ t) = -2\ i\ k\ \tilde{E}\, e^{i(k\,z-\omega\,t)} + 2\ i\ k\ \tilde{E}^*\ e^{-i(k\,z-\omega\,t)} \tag{1.71}$$

Now we apply the D_+ operator and make use of (69) to get for the wave equation

$$-2\ i\ k\ \left(D_+\ \tilde{E}\right) e^{i(k\,z-\omega\,t)} + 2\ i\ k\left(D_+\ \tilde{E}^*\right) e^{-i(k\,z-\omega\,t)} = -\left(\frac{1}{\epsilon_0\ c^2}\right)\frac{\partial\, j_x}{\partial\, t} \tag{1.72}$$

We multiply both sides by $e^{-i(k\,z-\omega\,t)}$

$$-2\ i\ k\ \left(D_+\ \tilde{E}\right) + 2\ i\ k\left(D_+\ \tilde{E}^*\right) e^{-2\ i(k\,z-\omega\,t)} = -\left(\frac{1}{\epsilon_0\ c^2}\right)\frac{\partial\, j_x}{\partial\, t}\ e^{-i(k\,z-\omega\,t)} \tag{1.73}$$

Now we again consider that the amplitudes are slowly varying, that is we average on both sides over a (sufficiently small) number of the rapidly oscillating periods, that is we average over $\Delta t = 2n\pi/\omega$. By this, the $\tilde{E}^*$ term on the left hand side vanishes, the first term is constant and we have

$$-2\ i\ k\ \left(D_+\ \tilde{E}\right) = -\left(\frac{1}{\epsilon_0\ c^2}\right)\left(\frac{1}{\Delta t}\right)\int_{-\Delta t/2}^{+\Delta t/2}\left(\frac{\partial\, j_x}{\partial\, t}\right) e^{-i(k\,z-\omega\,t')}\ dt' \tag{1.74}$$

Next trick : we integrate by parts and make use of the fact that j_x is assumed to be periodic in λ and get

$$2\ i\ k\ \left(D_+\ \tilde{E}\right) = \left(\frac{i\,\omega}{\epsilon_0\ c^2}\right)\left(\frac{1}{\Delta t}\right)\int_{-\Delta t/2}^{+\Delta t/2} j_x\ e^{-i(k\,z-\omega\,t')}\ dt' \tag{1.75}$$

What about the current density j_x ? The current is made from individual electrons at locations $z_j(t)$ having a transverse velocity from the undulator motion. If we take $-e/\left(2\,\pi\,{\sigma_x}^2\right)$ as charge per unit area for one electron, we have

$$j_x = -\left(\frac{e}{2\,\pi\,{\sigma_x}^2}\right)\left(\frac{c\,K}{\gamma_r}\right)\mathrm{Cos}[k_u\ z]\ \ \sum_j \delta[z - z_j[t]] \tag{1.76}$$

the sum counting over all contributing electrons. (We neglect that the electrons might have different γ). Using this, the integral over the δ-functions picks out all electrons in our $\pm\Delta t/2$ slice of the bunch

$$2\ i\ k\ \left(D_+\ \tilde{E}\right) = -\left(\frac{i\,\omega}{\epsilon_0\ c^2}\right)\left(\frac{e\,c\,K}{2\,\pi\,{\sigma_x}^2}\right)\left(\frac{1}{\gamma_r}\right)\left(\frac{1}{c\,\Delta t}\right)\sum_{j=1}^{N_\Delta}\mathrm{Cos}[k_u\ z]\ e^{-i(k\,z-\omega\,t_j)} \tag{1.77}$$

We replace now the sum by the average of all electrons in the slice N_Δ and use the volume number density of the electrons n_e to express $N_\Delta = n_e\left(2\,\pi\,{\sigma_x}^2\right)(\Delta t\ c)$ and have

$$2\ i\ k\ \left(D_+\ \tilde{E}\right) = -(i\ k)\left(\frac{e\ K\ n_e}{\epsilon_0\ \gamma_r}\right)\left\langle\mathrm{Cos}[k_u\ z]\ e^{-i(k\,z-\omega\,t_j)}\right\rangle \tag{1.78}$$

or

$$\left(D_+\ \tilde{E}\right) = -\left(\frac{e\ K\ n_e}{2\ \epsilon_0\ \gamma_r}\right)\left\langle\mathrm{Cos}[k_u\ z]\ e^{-i(k\,z-\omega\,t_j)}\right\rangle \tag{1.79}$$

Now we want to express the term to average by the slowly varying phase θ_j of the electrons

$$\theta_j(z) = (k + k_u)\ z - \omega\ \overline{t_j}(z) \tag{1.80}$$

The average time of the electrons in the slice we get by subtracting off the longitudinal motion of the electrons (see chapter on undulator radiation)

$$t_j(z) = \overline{t_j}(z) + \left(\frac{K^2}{\omega\left(4 + 2\,K^2\right)}\right)\text{Sin}[2\,k_u\,z] \tag{1.81}$$

and get for the function to be averaged

$$\text{Cos}[k_u\,z]\;e^{-i(k\,z-\omega\,t_j)} = e^{-i\,\theta_j}\;e^{i\,k_u\,z}\;e^{i\left(\frac{i\,K^2}{(4+2\,K^2)}\right)\text{Sin}[2\,k_u\,z]}\;\text{Cos}[k_u\,z] \tag{1.82}$$

Very clever people know that this can be expressed as an infinite "Fourier-Bessel" series as

$$= e^{-i\,\theta_j}\sum_n J_n\Big[\frac{K^2}{\left(4 + 2\,K^2\right)}\Big]\left(\frac{1}{2}\right)\left(e^{2\,i(n+1)\,k_u\,z} + e^{2\,i\,n\,k_u\,z}\right) \tag{1.83}$$

There are only two contributions to the sum which are *not* oscillating rapidly, $n = 0$ and n= -1 which we keep and have

$$= e^{-i\,\theta_j}\left(\frac{1}{2}\right)\left(J_0\Big[\frac{K^2}{\left(4 + 2\,K^2\right)}\Big] - J_1\Big[\frac{K^2}{\left(4 + 2\,K^2\right)}\Big]\right) = \frac{1}{2}\,e^{-i\,\theta_j}\;[\text{JJ}] \tag{1.84}$$

(Note : so we finally see where and how the K-modification factor [JJ] of equation (8) comes up).

Finally have

$$\left(D_+\,\tilde{E}\right) = -\kappa_2\,n_e\left\langle\,e^{-i\,\theta_j}\right\rangle \tag{1.85}$$

with

$$\kappa_2 = \frac{e\,K[\text{JJ}]}{4\,\epsilon_0\,\gamma_r} \tag{1.86}$$

We rewrite the operator D_+ in the coordinates z and θ

$$D_+ = \frac{1}{c}\left(\frac{\partial}{\partial\,t}\right)\Big|_z + \frac{\partial}{\partial\,z}\Big|_t = \frac{\partial}{\partial\,z}\Big|_\theta + k_u\,\frac{\partial}{\partial\,\theta}\Big|_z \tag{1.87}$$

and get

$$\Big[\frac{\partial}{\partial\,z} + k_u\,\frac{\partial}{\partial\,\theta}\Big]\,\tilde{E} = -\kappa_2\,n_e\left\langle\,e^{-i\,\theta_j}\right\rangle \tag{1.88}$$

The derivative of $\tilde{E}$ with θ is negligible as long as only the fundamental mode is considered so we have at the very end for the derivative of the complex, slowly varying field amplitude:

$$\frac{d\,\tilde{E}}{d\,z} = -\kappa_2\,n_e\left\langle\,e^{-i\,\theta_j}\right\rangle \tag{1.89}$$

Together with the pendulum equations, this forms the **coupled equations for the 1-dim FEL model:**

$$\frac{d\,\theta_j}{d\,z} = 2\,k_u\,\eta_j \tag{1.90}$$

$$\frac{d\,\eta_j}{d\,z} = \kappa_1\left(2\,\text{Re}\Big[\tilde{E}\,e^{i\,\theta_j}\Big]\right) = \kappa_1\left(\tilde{E}\,e^{i\,\theta_j} + \tilde{E}^*\,e^{-i\,\theta_j}\right) \tag{1.91}$$

with

$$\kappa_1 = \frac{e\,K\,[\text{JJ}]}{2\,\gamma_r{}^2\,m\,c^2} \tag{1.92}$$

Notice that this are as many equations as we have electrons in the slice which are coupled by the averaging in the field development.

The phase average $b = \langle e^{-i\theta_j} \rangle$ is called **bunching factor**, it actually measures the Fourier component of the density modulation with the wavelength of the resonant EM field. Notice that b is a complex number. If the θ_j are uniformly distributed, the bunching factor is zero. (Note: this would be true if the electrons are arranged on a regular grid which they are of course not. If the electrons are distributed randomly, the bunching factor will have a finite (random) value. This is important for the so called "SASE" process to be discussed later). If all electrons have the same phase θ_f, that is the bunch consists of periodic sheets of electrons separated by one wavelength λ_r, the bunching factor would be $b_f = e^{-i\theta_f}$ and have modulus $|b| = 1$.

Another point of attention : according to the definition in eq. (63), we have

$$\left|\tilde{E}(z,\ t)\right| = \frac{1}{2}|E_s(z,\ t)| \tag{1.93}$$

Thus for the energy density of the field we have

$$U_{\text{field}} = \frac{1}{2}\epsilon_0 |E_s(z,\ t)|^2 = 2\,\epsilon_0 \left|\tilde{E}(z,\ t)\right|^2. \tag{1.94}$$

The "fix field pendulum equations model", which we used in the low-gain case, is intrinsically not energy conserving since the electrons lose energy but the field does not change. So it is interesting to check if this has been cured with inclusion of the field variation.

According to equation (94), the change of field energy with z is

$$\frac{d\,U_{\text{field}}}{d\,z} = 2\,\epsilon_0 \frac{d}{d\,z}\left|\tilde{E}(z,\ t)\right|^2 = 2\,\epsilon_0 \frac{d}{d\,z}\left(\frac{1}{2}\left(\tilde{E}^2 + \tilde{E}^{*2}\right)\right) = 2\,\epsilon_0\left(\tilde{E}\frac{\partial\,\tilde{E}}{\partial\,z} + \tilde{E}^*\frac{\partial\,\tilde{E}^*}{\partial\,z}\right) = -2\,\epsilon_0\,\kappa_2\,n_e\left(\tilde{E}\left\langle e^{-i\theta_j}\right\rangle + \tilde{E}^*\left\langle e^{i\theta_j}\right\rangle\right) \tag{1.95}$$

Since the field amplitude is slowly varying, we can treat it as being constant over the slice to be averaged, that is we average the energy change of the electrons to

$$\frac{d\langle\eta_j\rangle}{d\,z} = \kappa_1\left(\tilde{E}\left\langle e^{i\theta_j}\right\rangle + \tilde{E}^*\left\langle e^{-i\theta_j}\right\rangle\right) \tag{1.96}$$

Inserting this into eq. (95) and using (92) and (86) yields

$$\frac{d\,U_{\text{field}}}{d\,z} = -2\,\epsilon_0\left(\frac{\kappa_2}{\kappa_1}\right)n_e\frac{d\langle\eta_j\rangle}{d\,z} = -\gamma\,m\,c^2\,n_e\frac{d\langle\eta_j\rangle}{d\,z} = -\frac{d\,U_{\text{beam}}}{d\,z} \tag{1.97}$$

The coupled equations are indeed intrinsically energy conserving.

1-D modelling by solving the coupled equations

After a lot of mathematical gymnastics, we are now ready to see how the FEL description changed by including the back-action of the electrons on the field. For direct comparison we use the same parameters as above. To follow conventions, we now use the modified definition of "gain". While above we named the energy *exchange* normalized to the initial field energy as gain (that is $G_{\text{old}} = \frac{\Delta U}{U_0}$), we now use the "power amplification factor" $G_{\text{new}} = \frac{U_{\text{out}}}{U_0} = G_{\text{old}} + 1$.

As another step, we now measure the depth in the undulator in units of the **gain length** L_g, definition and reasons becoming obvious in the next section. The 38 periods of the short undulator from above correspond to 2.1 gain lengths.

Again we start from an unmodulated beam, that is electrons equally distributed in θ, and an initial field $E_x(0)$. As a first test, we re-calculate the gain in the short undulator starting with a low initial field of our previous example, comparing the energy loss of the electrons and the increase in field energy, (figure 16 top). As expected, they are perfectly balanced.

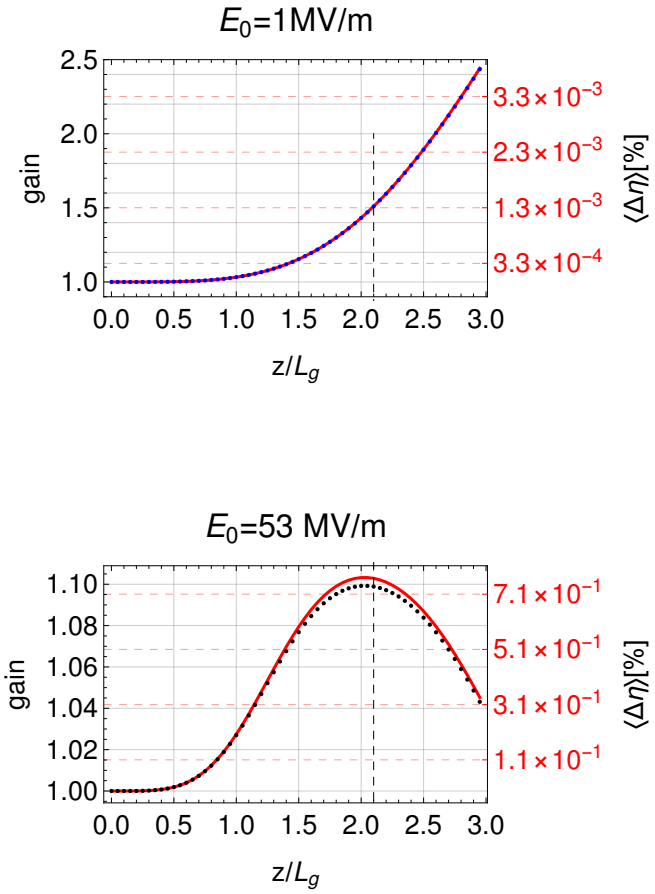


Figure 16: Top plot : Comparison of electron energy loss (blue dots) and field energy gain using the coupled equations. Bottom plot : The gain in the FELO undulator at saturated field strength. The model using a constant electric field (black dots) slightly underestimates the gain as compared to the coupled 1-D equations taking the field change into account (red line). The dashed vertical line indicates the length of the undulator used at FELIX (38 periods).

Figure (16 bottom) shows a comparison at FELO saturation field of the approximation with constant field with the refined "coupled equation" model, taking the field change during the passage of the bunch into account. The constant field approximation is not perfect but also not too bad in this case.

This changes dramatically if we make the undulator longer and start from a very low field as shown in figure (17).

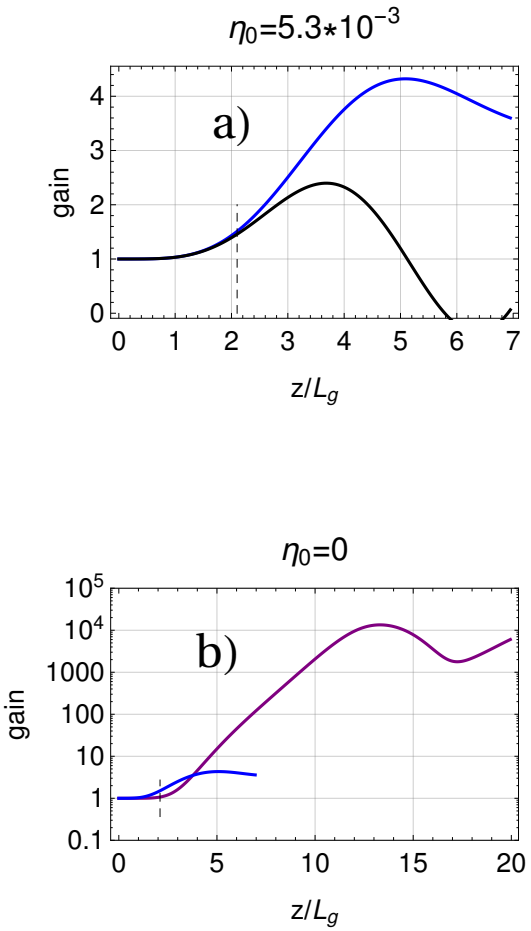


Figure 17: Gain evolution starting from small initial field (0.1 MV/m) as function of depth in the undulator. a) : Initial mean energy deviation at "Madey optimum" $\eta_0 = \eta_{LG}$ comparing the "constant E model" (black line) and the "SVA model" (blue line). b) : Initial energy deviation $\eta_0 = 0$ in the "SVA model" or "1D high gain model". Here the "constant E" model predicts zero gain for all z. The short blue line is the gain evolution with $\eta_0 = \eta_{LG}$ as shown in a). The dashed vertical line indicates the length of the undulator used at the FELIX FELO (38 periods). Note: this is a log scale now !

The refined model including the field evolution along the undulator ("SVA model") shows that using a long undulator, really large amplifications of small initial fields can be achieved, **if** the initial energy deviation η_0 is "at resonance", $\eta_0 = 0$, a condition where the "constant E model" predicts zero gains.

This finding is the basis for the operation of a "high gain FEL", that is for producing a high output power starting from very low initial power in a single pass through the undulator. It allows to build mirror free FEL operating down to wavelengths in the sub-nm regime.

Before we try to understand the physics behind this, we go one step further and normalize our parameters to find the universal scaling laws.

Normalized parameters

Making the coupled differential equations dimensionless greatly simplifies the understanding of some universal scaling parameters and scaling rules. The goal is to find the proper "scaling parameter" to end up with dimensionless equations with unity coefficients.

First we introduce a scaled, normalized longitudinal coordinate as

$$\hat{z} = 2\, k_u \rho\, z \tag{1.98}$$

The **parameter "ρ" is kept free** to later fulfill the second requirement of unity coefficients. Next we introduce a re-scaled energy deviation by

$$\hat{\eta} = \eta / \rho \tag{1.99}$$

Finally we introduce for the EM field a dimensionless complex field amplitude

$$a = \frac{\kappa_1}{2\, k_u \rho^2}\, \tilde{E} \tag{1.100}$$

The coupled equations are now dimensionless and read

$$\frac{d\theta_j}{d\hat{z}} = \hat{\eta}_j \tag{1.101}$$

$$\frac{d\hat{\eta}_j}{d\hat{z}} = a\, e^{i\theta_j} + a^*\, e^{-i\theta_j} \tag{1.102}$$

$$\frac{da}{d\hat{z}} = -\left(\frac{\kappa_1\, \kappa_2\, n_e}{4\, k_u^2\, \rho^3}\right)\left\langle e^{-i\theta_j}\right\rangle \equiv -\left\langle e^{-i\theta_j}\right\rangle \tag{1.103}$$

To make the last equation (as the two others) of **unity coefficient**, we choose the free parameter ρ to be

$$\rho = \left(\frac{\kappa_1\, \kappa_2\, n_e}{4\, k_u^2}\right)^{1/3} = \left(\left(\frac{1}{8\,\pi}\right)\left(\frac{I_b}{I_a}\right)\left(\frac{K\,[\mathrm{JJ}]}{1+\frac{K^2}{2}}\right)^2\left(\frac{\gamma\,\lambda^2}{2\,\pi\,\sigma_x^2}\right)\right)^{1/3}$$

With this dimensionless parameter, called either **"ρ-parameter" or "Pierce parameter"**, the "gain evolution" in the high gain FEL regime is reduced to a universal, scaled behavior as we will see in the next section. The order of magnitude of ρ can be easily estimated from the "dimensionless fractions" notation shown above. If we assume an X-ray FEL ($\lambda = 10^{-9}$ m, $\gamma = 10^4$) and a beam diameter of $\sigma_x = 10^{-4}$m, the last bracket would be about 10^{-7} . The beam current is typically 0.3 I_a (I_a the Alfvén current), K of order unity so that $\rho \simeq \left(\left(10^{-9}\right)^{1/3}\right) \simeq 10^{-3}$.

The "Cubic equation"

On resonance : $\eta_0 = 0$

The coupled equations derived in the previous section can be solved using numerical techniques on a sufficiently large ensemble of electrons. They are the basic equations of a 1-D FEL model for a transversely uniform, periodically modulated infinite electron bunch.

While the numerical solution gives insight to the particle dynamics including saturation effects by "trapping", very general scaling rules can be obtained from looking at ensemble averages. Two such ensemble averages have already been used: the bunching factor $b = \left\langle e^{-i\theta_j}\right\rangle$ and the normalized complex field amplitude $a(\hat{z})$. Now we also average the normalized $\hat{\eta}$ over the ensemble of electrons in the slice and define the "collective energy modulation" as

$$P(\hat{z}) = \left\langle \hat{\eta}_j\, e^{-i\theta_j}\right\rangle$$

P($\hat{z}$) describes the averaged complex amplitude of the Fourier component of the energy modulation with periodicity one, that is, according to the definition of the phase θ, at the resonance wavelength. Similarly, the bunching factor b measures the strength of the density modulation at this fundamental wavelength. Notice that the average energy deviation is set to be zero, the beam is "on resonance", $\eta_0 = 0$.

The equation for the field amplitude is as before :

$$\frac{da}{d\hat{z}} = -\left\langle e^{-i\theta_j}\right\rangle = -b$$

The equation for the bunching factor reads

$$\frac{d\,b}{d\,\hat{z}} = \frac{d}{d\,\hat{z}}\left\langle e^{-i\,\theta_j}\right\rangle = -i\left\langle\left(\frac{d\,\theta_j}{d\,\hat{z}}\right)e^{-i\,\theta_j}\right\rangle = -i\left\langle\,\eta_j\,e^{-i\,\theta_j}\right\rangle = -i\,P(\hat{z})$$

For $P(\hat{z})$ finally we find

$$\frac{d\,P}{d\,\hat{z}} = \left\langle\frac{d\,\eta_j}{d\,\hat{z}}\,e^{-i\,\theta_j}\right\rangle - i\left\langle\eta_j\,\frac{d\,\theta_j}{d\,\hat{z}}\,e^{-i\,\theta_j}\right\rangle = \left\langle\frac{d\,\eta_j}{d\,\hat{z}}\,e^{-i\,\theta_j}\right\rangle - i\left\langle\eta_j^{\,2}\,e^{-i\,\theta_j}\right\rangle$$

using the DG for $\hat{\eta}$ ( ) we get

$$\frac{d\,P}{d\,\hat{z}} = a + a^*\left\langle e^{-2\,i\,\theta_j}\right\rangle - i\left\langle\eta_j^{\,2}\,e^{-i\,\theta_j}\right\rangle$$

The second term implies the "second order bunching factor", that is bunching at twice the fundamental frequency, the third term is of higher order in the small number η, both terms will be neglected here. This "linearization" assumes that both the bunching factor b as well as the amplitude of the relative energy modulation are << 1. Doing so, the following formulas, especially the "third order equation", is more restricted than the coupled equations (101).

With this restriction, we have finally a quite simple system of three linear equations

$$\frac{d\,a}{d\,\hat{z}} = -b \tag{1.104}$$

$$\frac{d\,b}{d\,\hat{z}} = -i\,P \tag{1.105}$$

$$\frac{d\,P}{d\,\hat{z}} = a \tag{1.106}$$

These three equations reflect the "feedback system" driving the FEL amplification:

- the bunching drives the field evolution (amplitude and phase)
- the energy modulation drives the bunching
- the complex field amplitude drives the energy modulation

The three equations can be simply reduced to one **cubic differential equation** for the complex field amplitude a:

$$\frac{d^3\,a}{d\,\hat{z}^3} = i\,a \tag{1.107}$$

If we make the Ansatz $a(\hat{z}) \sim e^{-i\mu\,\hat{z}}$ we find for μ the algebraic equation

$$\mu^3 = 1 \tag{1.108}$$

This equation has three roots (we now use a bit of *Mathematica* coding)

$$\mu_1 = 1 \tag{1.109}$$

$$\mu_2 = -(-1)^{1/3} = -\frac{1}{2}\left(1 + i\,\sqrt{3}\right) \tag{1.110}$$

$$\mu_3 = (-1)^{2/3} = -\frac{1}{2}\left(1 - i\,\sqrt{3}\right) \tag{1.111}$$

The general solution is a linear combination of all three solutions with coefficients C_l:

$$a(\hat{z}) = \sum_{l=1}^{3} C_l\,e^{-i\,\mu_l\,\hat{z}} = C_1\,e^{-i\,\hat{z}} + C_2\,e^{\frac{1}{2}\left(-i+\sqrt{3}\right)\hat{z}} + C_3\,e^{\frac{1}{2}\left(-i-\sqrt{3}\right)\hat{z}} \tag{1.112}$$

All three terms have an oscillating component, the second term has an exponentially decaying part while the third term offers a exponential growing contribution. For sufficiently large $\hat{z}$, this term will dominate and lead to an exponential growth of the field

amplitude.

The yet unknown coefficients C_l are derived from the initial conditions, that is from $a(0)$, $b(0)$ and $P(0)$. The conditions are

$$a(0) = \sum_{l=1}^{3} C_l \tag{1.113}$$

$$b(0) = -\left.\frac{d\,a}{d\,\hat{z}}\right|_0 = i\sum_{l=1}^{3}\mu_l\ C_l \tag{1.114}$$

$$P(0) = i\left.\frac{d\,b}{d\,\hat{z}}\right|_0 = i\sum_{l=1}^{3}{\mu_l}^2\ C_l \tag{1.115}$$

These equations can be written in matrix form

$$\begin{pmatrix} a_0 \\ b_0 \\ P_0 \end{pmatrix} = M_\mu \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ i\,\mu_1 & i\,\mu_2 & i\,\mu_3 \\ i\,{\mu_1}^2 & i\,{\mu_2}^2 & i\,{\mu_3}^2 \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ i & \frac{1}{2}\left(-i+\sqrt{3}\right) & \frac{1}{2}\left(-i-\sqrt{3}\right) \\ i & \frac{1}{2}\left(-i-\sqrt{3}\right) & \frac{1}{2}\left(-i+\sqrt{3}\right) \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} \tag{1.116}$$

The unknown coefficients C_l are then given by inverting the matrix M_μ, that is

$$\begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = {M_\mu}^{-1} \begin{pmatrix} a_0 \\ b_0 \\ P_0 \end{pmatrix} = \begin{pmatrix} \frac{1}{3} & -\frac{i}{3} & -\frac{i}{3} \\ \frac{1}{3} & \frac{1}{6}\left(i+\sqrt{3}\right) & \frac{1}{3}(-1)^{5/6} \\ \frac{1}{3} & \frac{1}{6}\left(i-\sqrt{3}\right) & \frac{1}{3}(-1)^{1/6} \end{pmatrix} \begin{pmatrix} a_0 \\ b_0 \\ P_0 \end{pmatrix} \tag{1.117}$$

We look now at the case that the beam is initially unmodulated in density and energy, that is $b(0) = 0$ and $P(0) = 0$ and we amplify an initial field amplitude a_0.

In this case, all coefficients C_l identical, $C_l = 1/3\ a_0$, and we can write the field amplitude as

$$a(\hat{z}) = \frac{1}{3}\,a_0\left(e^{-i\,\hat{z}} + e^{-\frac{1}{2}\left(-i+\sqrt{3}\right)\hat{z}} + e^{\frac{1}{2}\left(i+\sqrt{3}\right)\hat{z}}\right) \tag{1.118}$$

The "gain" (defined as "power gain") is then $G = \left(\frac{|a|}{a_0}\right)^2$:

```
gain =
 Simplify[1 / a0^2 (again[zd] * (again[zd] /. {Complex[0, -1] → Complex[0, 1], Complex[
              0, 1] → Complex[0, -1]})) // ComplexExpand] // Expand
```

$$\frac{1}{3} + \frac{1}{9}\,e^{-\sqrt{3}\ \mathrm{zd}} + \frac{e^{\sqrt{3}\ \mathrm{zd}}}{9} + \frac{2}{9}\,e^{-\frac{\sqrt{3}\ \mathrm{zd}}{2}}\,\mathrm{Cos}\left[\frac{3\ \mathrm{zd}}{2}\right] + \frac{2}{9}\,e^{\frac{\sqrt{3}\ \mathrm{zd}}{2}}\,\mathrm{Cos}\left[\frac{3\ \mathrm{zd}}{2}\right]$$

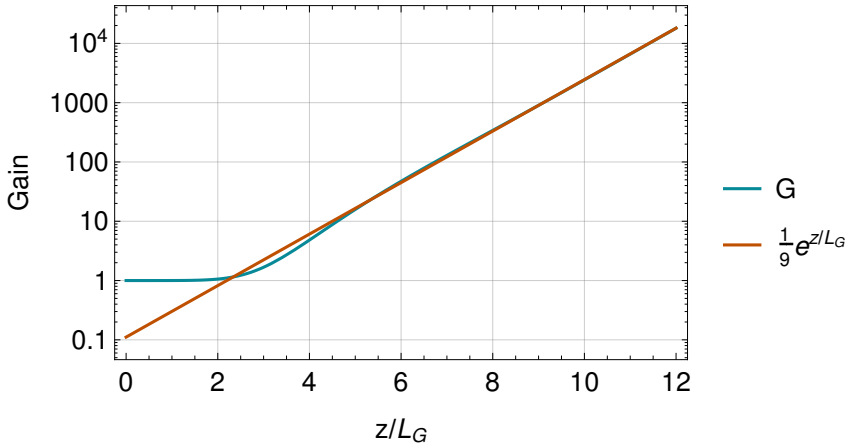
$$G = \frac{1}{9}\left(3 + e^{-\sqrt{3}\ \hat{z}} + e^{\sqrt{3}\ \hat{z}} + 2\,\mathrm{Cos}\left[\frac{3\,\hat{z}}{2}\right]\left(e^{-\frac{\sqrt{3}}{2}\hat{z}} + e^{\frac{\sqrt{3}}{2}\hat{z}}\right)\right) \tag{1.119}$$

This is a **universal curve in the normalized coordinate $\hat{z}$**. For large $\hat{z}$, the gain increases as $G \sim \frac{1}{9}e^{\sqrt{3}\ \hat{z}} = \frac{1}{9}\,e^{\left(2\sqrt{3}\ k_u\rho\right)z}$ and we define the **1D power gain length** as

$$L_G = \frac{1}{2\sqrt{3}\ k_u\,\rho} = \frac{\lambda_u}{4\,\pi\,\sqrt{3}\ \rho} \tag{1.120}$$

The universal gain curve as function of z/L_G looks as shown in the following plot.

```
punivg = LogPlot[Evaluate[{gain, 1 / 9 E^x} /. zd → x / Sqrt[3]], {x, 0, 12},
  Frame -> True, PlotTheme → "Presentation", FrameLabel → {"z/L_G", "Gain"},
  GridLines → Automatic, PlotLegends → {"G", "1/9 e^(z/L_G)"}]
```



While for $z/L_G >> 4$, the exponential increase is dominating, for $z/L_G < 3$ the gain is basically constant at unity level, that is negligible amplification of the incoming wave. This regime is called "lethargy regime". The physics behind it will be become more clear in a moment.

If we do a series expansion for small z/L_g we find that the gain for short distances is given by

$$G_{\text{leth}} = 1 + \frac{1}{1080}\,(z/L_G)^6 = 1 + (z/(3.2\,L_G))^6 \tag{1.121}$$

verifying that the length of the lethargy regime is about 3 gain lengths.

Saturation

In the exponential regime, the normalized field amplitude $|a|$ and the bunching factor $|b|$ are identical

$$|a|^2 = |b|^2 = \frac{1}{9}\,a_0{}^2\,e^{z/L_g} \tag{1.122}$$

In the simplified "collective model", both would increase without limits. On the other hand, the bunching factor $|b| = \langle e^{-i\theta_j}\rangle$ cannot exceed 1, thus the same holds for the normalized field amplitude a. This effect of "saturation" is intrinsically not included in the cubic equation since it is derived under the assuption that |b| << 1. From the maximum allowed bunching factor one can immediately derive the maximum power level at saturation:

From

$$|a| = \frac{\kappa_1}{2\,k_u\,\rho^2}\,|\tilde{E}| \le 1 \tag{1.123}$$

we get

$$|\tilde{E}_{\text{sat}}| \simeq \frac{2\,k_u\,\rho^2}{\kappa_1} \tag{1.124}$$

or for the radiation energy density (see (94))

$$U_{\text{sat}} = 2\,\epsilon_0\,|\tilde{E}_{\text{sat}}|^2 = 2\,\epsilon_0\,\rho\left(\frac{2\,k_u}{\kappa_1}\right)^2\rho^3 = 2\,\epsilon_0\,\rho\,n_e\left(\frac{\kappa_2}{\kappa_1}\right) = \rho\,n_e\,\gamma_r\,m\,c^2 = \rho\,U_{\text{beam}} \tag{1.125}$$

The total radiation power P is given by energy density times beam area times c :

$$P = U\,c\,A_b = 2\,\pi\,\sigma_b{}^2\,c\,U \tag{1.126}$$

The saturation power of the FEL is

$$P_{\text{sat}} = \rho \; U_{\text{beam}} = \rho \; I_b \, E_b \tag{1.127}$$

This is a noteworthy result: the maximum achievable radiation power density at saturation is the total beam power density times the ρ - parameter. Notice that this holds as well for a "high gain", single pass FEL as for the "low gain", oscillator FEL.

The saturation power is independent of the initial power density U_0. Since the gain length as well does not depend on U_0, the undulator length to reach saturation and the maximum gain depend on U_0.

The **Pierce (or rho)-parameter is the only** (!), dimensionless number describing the properties of an individual FEL in the universal 1-D model we presented here.

The gain length ist given by

$$L_G = \frac{\lambda_u}{4\,\pi\,\sqrt{3}\,\rho} \approx \frac{\lambda_u}{22\,\rho} \tag{1.128}$$

and the saturation power is ρ times the beam power. For a typical ρ of 10^{-3} (nm regime) , we would have $L_G \approx 50\,\lambda_u$. That is you need about 150 undulator periods to overcome the lethargy regime and about 500 periods reach saturation. The energy transfer from the e-beam to the EM field is about 0.1%.

Studies for a typical FEL

The following parameters roughly reflect typical operational parameters of FLASH, the machine who pioneered the "nm-FEL" regime.

λ_L	λ_u	K	$\sigma_{x,y}$	σ_z	K_{JJ}	γ_r	E_b	I_b	n_e
10 nm	27.3 mm	0.6	0.2 mm	0.9 mm	0.576	1269	648 MeV	2 kA	$1.6 * 10^{20}$ m^{-3}

From these machine parameters, we get in our 1-D model the follwing FEL parameters :

ρ	L_g	P_{sat}
$8.25 * 10^{-4}$	1.52 m	1 GW

First we compare the result from a numerical simulation using the coupled equations with the universal gain curve, eq. (119) resulting from the cubic quation (figure 18).

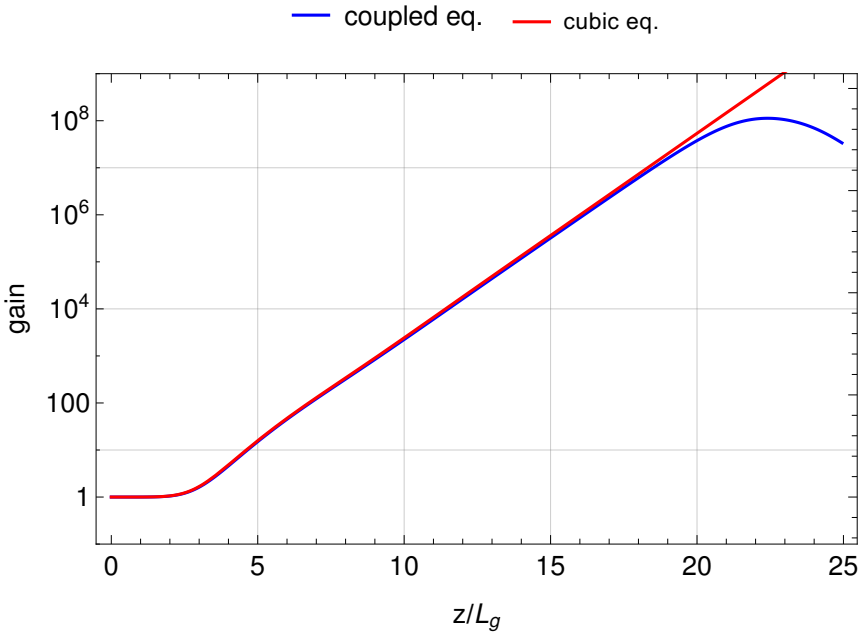


Figure 18 : Comparison of the universal gain curve resulting from the cubic equation to the numerical solution of the 1-D coupled equations. The regime where both curves are in good agreement is called the "linear regime" of the FEL.

Notice that the universal curve (red) does not depend on any of the FEL parameters if the distance is normalized to the gain length. The universal curve agrees extremly well with the numerical solution except for the effect of saturation.

If we plot instead of the gain the real FEL power for two initial field amplitudes (figure 19), we see that what we found above: the slope of the exponential rise, that is the gain length, and the saturation level do not depend on E_0. The saturation power is roughly what we expect, for these parameters about 1 GW (see table above), independent of the initial field.

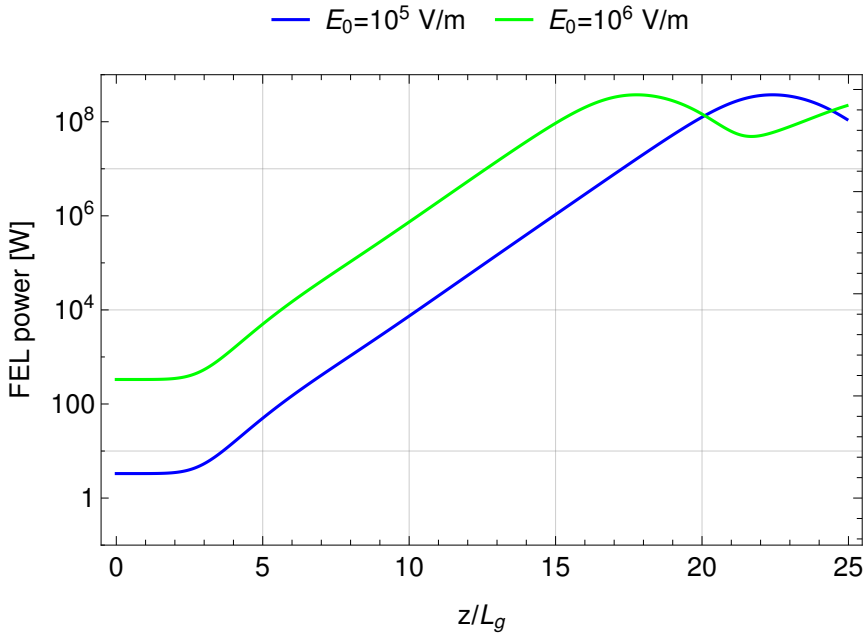


Figure 19 : Power output of the FEL for two different initial field strengths.

If the beam current would be 1 kA instead of 2 kA, the ρ-parameter drops to $6.55*10^{-4}$, the gain length increases to 1.94 m. Figure (20) compares the two cases, the distance now *not* normalized to the gain length. The lower beam current results in a larger gain length and reduced saturation level.

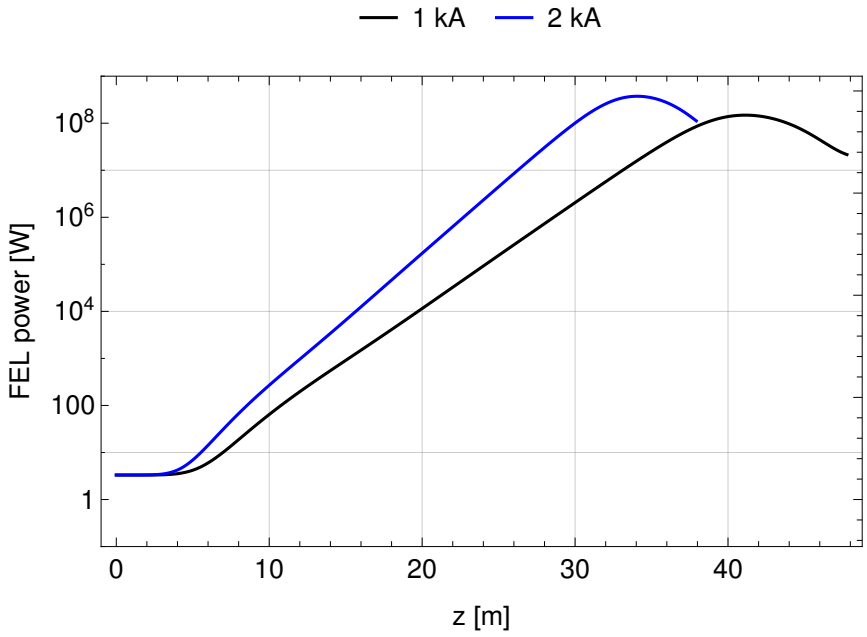


Figure 20 : Power output of the FEL for two different peak currents and otherwise identical conditions.

The magic phase

For the short undulator case, we learned that there is no net energy transfer between the beam the EM field since the density modulation and the field are phase shifted by $\pi/2$. There are as many electrons transferring energy to the field (induced emis-

sion) as getting energy from the field (absorption). In conventional laser physics this corresponds to the case where there is no difference in population between the upper and lower state.

The fundamental process which initiates a net energy transfer even in case of η_0=0 is the fact the EM field has to propagate **inside the electron beam** and exchanges energy with the beam in a periodically modulated pattern. This results in a very slight reduction of the phase velocity of the field, it falls back against the modulated beam resulting in a small net energy transfer to the field. This "coupling" causes also the beam modulation phase to slip back, the process stabilizes itself and field and density modulation propagate "phase locked", such that optimal energy transfer to the field is maintained. This self-sustaining process is a typical "collective instability" which, at other places in accelerator physics, would be highly undiserable. Here, it destroys the beam but produces a lot of EM power, even if the efficiency is of the order ρ, that is about 1 per mille.

What I described here can be seen looking at the phases of the beam modulation and the EM-field as shown in figure (21).

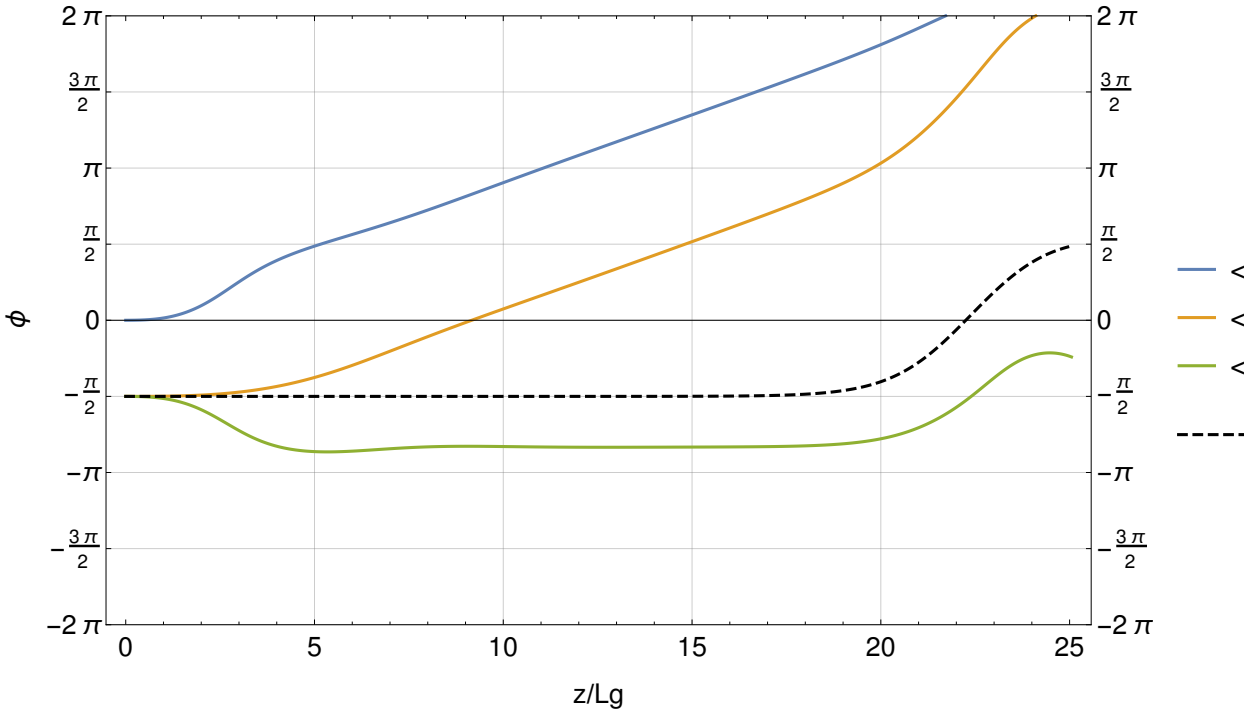


Figure 21: Power output of the FEL for two different peak currents and otherwise identical conditions.

In the lethargy regime, the phase of the beam is more or less constant at -π/2, the phase of the EM field increases. From about 4 L_g, the phase difference between modulation and field stays constant close to -$\frac{3}{4}\pi$ up to the saturation regime. This "phase locking" between beam phase and field phase is the key effect which allows to transfer energy efficiently from the beam to the field over a long distance. The dashed line in fig. (*21*) shows the phase slippage due to the mean energy loss of the electrons. At about 20 L_g, the electrons have lost that much energy that this "detuning phase" becomes dominant, the phase locking between field and beam breaks and the energy transfer comes to an end (saturation).

With energy offset η_0

In the previous section we discussed the gain evolution for a beam entering the undulator "on resonance", the initial scaled energy deviation was η_0= 0. The case with $\eta_0 \neq 0$ is only slightly more complicated. In this case, the "collective energy deviation" has to be described as

$$P(\hat{z}) = \hat{\eta}_0 + \left\langle \hat{\eta}_j \, e^{-i\theta_j} \right\rangle \tag{1.129}$$

While two equations stay unchanged

$$\frac{da}{d\hat{z}} = -b \tag{1.130}$$

$$\frac{d\hat{\eta}_j}{d\hat{z}} = a\, e^{i\theta_j} + a^*\, e^{-i\theta_j} \tag{1.131}$$

the phase development now includes the linear phase increase due to the initial "de-tuning" of the beam energy

$$\frac{d\theta_j}{d\hat{z}} = \hat{\eta}_j + \hat{\eta}_0 \tag{1.132}$$

(We already discussed the linear phase increase in the context of the low gain FEL. Using not-normalized coordinates, we had $\theta'[z] = 2\, k_u\, \eta_0$ + higher orders).

As we did above, we calculate the derivative of P and get

$$\frac{dP}{d\hat{z}} = \left\langle \frac{d\hat{\eta}_j}{d\hat{z}}\, e^{-i\theta_j} \right\rangle - i \left\langle \hat{\eta}_j \frac{d\theta_j}{d\hat{z}}\, e^{-i\theta_j} \right\rangle = \left\langle \frac{d\hat{\eta}_j}{d\hat{z}}\, e^{-i\theta_j} \right\rangle - i \left\langle {\eta_j}^2\, e^{-i\theta_j} \right\rangle - i\,\hat{\eta}_0 \left\langle \hat{\eta}_j\, e^{-i\theta_j} \right\rangle \simeq \left\langle \frac{d\hat{\eta}_j}{d\hat{z}}\, e^{-i\theta_j} \right\rangle - i\,\hat{\eta}_0\, P + i\,{\hat{\eta}_0}^2 \simeq a - i\,\hat{\eta}_0\, P + i\,{\hat{\eta}_0}^2 \tag{1.133}$$

Again we neglected the ${\eta_j}^2$ term and used eq. (102) neglecting the $e^{-2\,i\,\theta_j}$ component.

For the derivative of the bunching factor we get

$$\frac{db}{d\hat{z}} = \frac{d}{d\hat{z}} \left\langle e^{-i\theta_j} \right\rangle = -i \left\langle \left(\frac{d\theta_j}{d\hat{z}} \right) e^{-i\theta_j} \right\rangle = -i \left\langle \eta_j\, e^{-i\theta_j} \right\rangle - i\,\hat{\eta}_0 \left\langle\, e^{-i\theta_j} \right\rangle = -i\, P(\hat{z}) + i\,\hat{\eta}_0 - i\,\hat{\eta}_0\, b = \tag{1.134}$$

$$= -i\, P(\hat{z}) + i\,\hat{\eta}_0 + i\,\hat{\eta}_0\, a' = -a'' \tag{1.135}$$

For immediate use we compile from that

$$2\, i\,\hat{\eta}_0\, a'' = -2\,\hat{\eta}_0\, P + 2\,{\hat{\eta}_0}^2 + 2\,{\hat{\eta}_0}^2\, a' \tag{1.136}$$

For the second derivative we get

$$\frac{d^2 b}{d\hat{z}^2} = -i\,\frac{dP}{d\hat{z}} - i\,\hat{\eta}_0\,\frac{db}{d\hat{z}} = -i\,a - i\left(-i\,\hat{\eta}_0\, P + i\,{\hat{\eta}_0}^2\right) - \hat{\eta}_0\, P(\hat{z}) + {\hat{\eta}_0}^2 + {\hat{\eta}_0}^2\, a' = \tag{1.137}$$

$$-i\,a - \hat{\eta}_0\, P + {\hat{\eta}_0}^2 + {\hat{\eta}_0}^2 - \hat{\eta}_0\, P + {\hat{\eta}_0}^2\, a' = -i\,a - 2\,\hat{\eta}_0\, P + 2\,{\hat{\eta}_0}^2 + {\hat{\eta}_0}^2\, a' =$$

$$= -i\,a + 2\,i\,\hat{\eta}_0\, a'' - {\hat{\eta}_0}^2\, a' \equiv -a''' \tag{1.138}$$

If we now use what we compiled in (136), we finally have the differential equation for the normalized field amplitude for the case of non vanishing initial energy deviation $\hat{\eta}_0$:

$$\boxed{i\,a + {\hat{\eta}_0}^2\, a' - 2\,i\,\hat{\eta}_0\, a'' == a'''} \tag{1.139}$$

Making, as above, the Ansatz $a[\hat{z}] = a_0\, e^{-i\mu\,\hat{z}}$ we get for μ the algebraic equation

$$\mu^3 - 2\,\hat{\eta}_0\,\mu^2 + {\hat{\eta}_0}^2\,\mu - 1 = 0 \tag{1.140}$$

The three general solutions of this equation can be found using *Mathematica* without problem but they are a bit bulky :

$$\left\{\mu_1 \to \frac{1}{3}\left(2\,\hat{\eta}_0 + \frac{2^{1/3}\,\hat{\eta}_0^2}{\left(-2\,\hat{\eta}_0^3 + 3\left(9 + \sqrt{81 - 12\,\hat{\eta}_0^3}\right)\right)^{1/3}} + \frac{\left(-2\,\hat{\eta}_0^3 + 3\left(9 + \sqrt{81 - 12\,\hat{\eta}_0^3}\right)\right)^{1/3}}{2^{1/3}}\right),\right.$$

$$\mu_2 \to \frac{1}{6}\left(4\,\hat{\eta}_0 - \frac{2\,(-2)^{1/3}\,\hat{\eta}_0^2}{\left(-2\,\hat{\eta}_0^3 + 3\left(9 + \sqrt{81 - 12\,\hat{\eta}_0^3}\right)\right)^{1/3}} + (-2)^{2/3}\left(-2\,\hat{\eta}_0^3 + 3\left(9 + \sqrt{81 - 12\,\hat{\eta}_0^3}\right)\right)^{1/3}\right), \tag{1.141}$$

$$\left.\mu_3 \to -\frac{1}{3}\left(-\frac{1}{2}\right)^{1/3}\left(-2\,\hat{\eta}_0^3 + 3\left(9 + \sqrt{81 - 12\,\hat{\eta}_0^3}\right)\right)^{1/3} +\right.$$

$$\frac{1}{3}\hat{\eta}_0\left(2+\left(\text{Root}\left[-2+\#1^3\,\&,\,3\right]\hat{\eta}_0\right)\Big/\left(-2\,\hat{\eta}_0^3+3\left(9+\sqrt{81-12\,\hat{\eta}_0^3}\right)\right)^{1/3}\right)\Bigg\}$$

Just from "inspection" of the terms, it seems to be important if

$$(\hat{\eta}_0)^3 < \frac{81}{12} \tag{1.142}$$

is fulfilled or not.

Again, the general solution will be the sum of the three fundamental solutions with coefficients C_n defined by the initial conditions.

As we made the Ansatz, an exponentially growing term needs a *positive imaginary part* of μ, so we check that graphically.

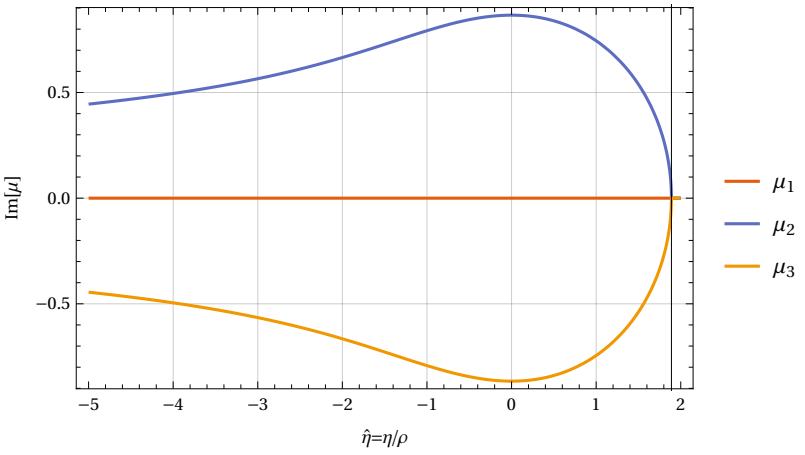


Figure 22: Imaginary part of the three general solutions eq. (141) as function of the initial energy deviation η_0. Only μ_2 has a positive imaginary part and leads to exponential power growth. The vertical line indicates $\eta/\rho = \sqrt[3]{\left(\frac{81}{12}\right)}$.

Notice that again this is a general solution in $\hat{\eta}=\eta/\rho$, the Pierce parameter is *the* ruling constant of the FEL process.

We see that only μ_2 has a positive imaginary part and that there is a hard "cut off" at $\eta_{\max} = \sqrt[3]{\left(\frac{81}{12}\right)}\,\rho \approx 1.88\,\rho$. For larger energy deviations, no exponentially growing term exists, the solution is purely oscillating.

To find the coefficients of the general solution we write it again as matrix equation

$$\begin{pmatrix} a_0 \\ b_0 \\ P_0 \end{pmatrix} = M_\mu \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ i\,\mu_1 & i\,\mu_2 & i\,\mu_3 \\ i\,{\mu_1}^2 & i\,{\mu_2}^2 & i\,{\mu_3}^2 \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} \tag{1.143}$$

In the following, we study the two cases of amplification of an initial field and amplification of the initial density modulation.

Initial field amplification

We start from an initial field a_0, no initial bunching and get for the three coefficients

$$\left\{C_1 \to \frac{a_0\,\mu_2\,\mu_3}{(\mu_1-\mu_2)\,(\mu_1-\mu_3)},\; C_2 \to \frac{a_0\,\mu_1\,\mu_3}{(\mu_1-\mu_2)\,(\mu_3-\mu_2)},\; C_3 \to \frac{a_0\,\mu_1\,\mu_2}{(\mu_1-\mu_3)\,(\mu_2-\mu_3)}\right\} \tag{1.144}$$

Using the μ_n found in eq. (141) gives rather bulky expressions which are not useful to took at in full detail. But we can now plot the gain evolution for various initial energy offsets as it is done in Figure (23).

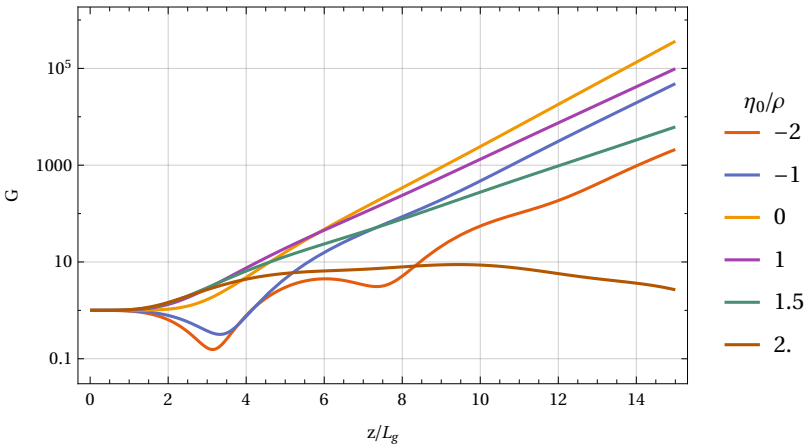


Figure 23: Gain curves for different initial energy deviations η_0 normalized to the Pierce parameter.

The $\eta_0 = 0$ curve is of course identical to the "universal gain curve" derived earlier. We see that for the positive initial energy deviation, the lethargy regime is shorter, actually this is pointing in the direction of the "short undulator FEL" or "low gain operation". For negative η_0, the initial gain is always negative, the electrons suck energy from the field before the exponential growth takes over. For η_0 above the magic threshold $\eta_{\max}$, no exponential regime exists.

For sufficiently large η_0, the positive exponential will dominate and we can define a **"modified power gain length"** in normalized coordinates $\hat{z}$ according to

$$\hat{L}_g = \frac{1}{2\,\mathrm{Im}(\mu_2)} \tag{1.145}$$

or in real coordinates $z = \hat{z}/(2\,k_u\,\rho)$ as

$$L_g(\eta_0) = \frac{1}{4\,k_u\,\rho\,\mathrm{Im}\,(\mu_2)} \tag{1.146}$$

(The factor of 2 comes from the definition as "power gain" length and the power scaling with $|\,a\,|^2$).

It is common practice to define a **"power growth rate function"** as ratio of the modified gain length to the gain length at $\eta_0 = 0$:

$$f_{\mathrm{gr}}(\eta_0) = \frac{L_g(\eta_0)}{L_g(0)} \tag{1.147}$$

For $\eta_0 < \eta_{\max}$ we can write this function in closed form as

$$f_{\mathrm{gr}}(\eta_0) = \frac{L_g(\eta_0)}{L_g(0)} = \frac{-2 \times 2^{1/3}\,\eta_0^2 + \left(2\,\alpha - 4\,\eta_0^3\right)^{2/3}}{6\left(\alpha - 2\,\eta_0^3\right)^{1/3}} \quad \text{with} \quad \alpha = 3\left(9 + \sqrt{81 - 12\,\eta_0^3}\right) \tag{1.148}$$

and it looks as shown in figure (24).

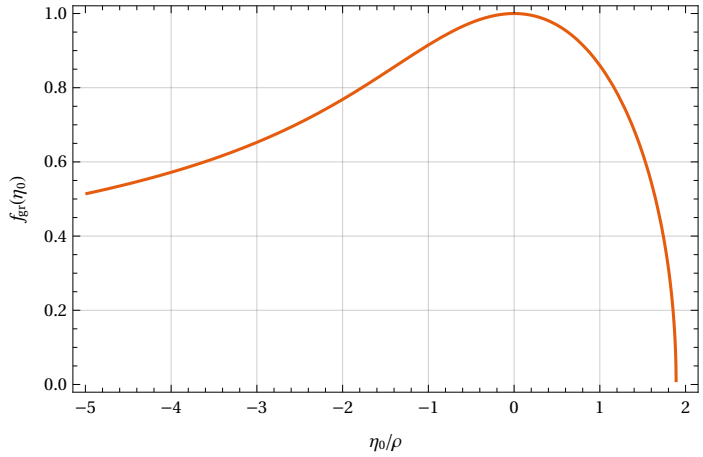


Figure 24: Gain curves for different initial energy deviations η_0 normalized to the Pierce parameter.

The fastes growth rate is always achieved for $\eta_0/\rho =$ Depending on η_0/ρ is as well the 0 but this does not mean that for *a* given undulator length $\eta_0/\rho = 0$ produces the largest gain. undulator length needed to reach the exponential gain regime.

The following figure (25) summarizes the "gain curves" at various undulator depths:

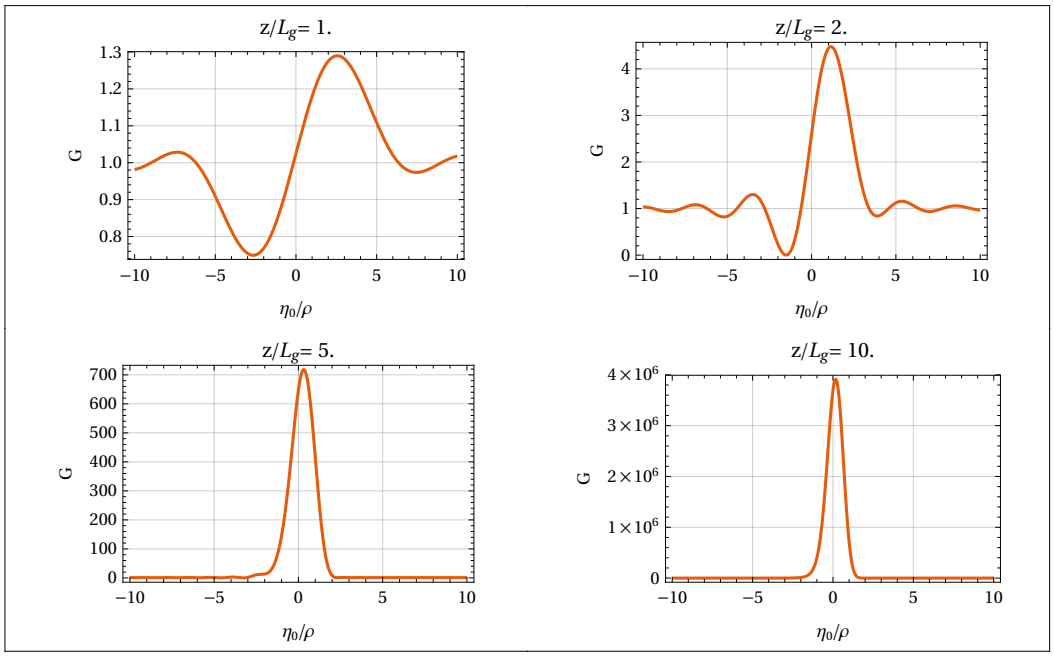


Figure 25: Gain as function of initial energy deviation η_0 at different undulator depth.

While for a short undulator we see the typical "low gain FEL curve", for longer undulators much higher gains are achieved in a narrow regime around $\eta_0/\rho = 0$.

For a sufficiently long undulator, the gain is dominated by $\mathrm{Im}(\mu_2)$ and the relevant regime of $\hat{\eta}_0$ is small (notice: while η_0 is always a small number, $\hat{\eta}_0 = \eta_0/\rho$ can be of order unity or larger. So we now restrict us to the narrow regime of high gains for $z/L_g >> 1$). In this case, we expand μ_2 around $\hat{\eta}_0$=0 to second order:

$$\mu_2 \approx -\frac{1}{2} + \frac{2\,\hat{\eta}_0}{3} - \frac{\hat{\eta}_0^2}{18} + i\left(\frac{\sqrt{3}}{2} - \frac{\hat{\eta}_0^2}{6\sqrt{3}}\right) \tag{1.149}$$

The power gain rises as

$$\sim e^{2\,\mathrm{Im}(\mu_2)\,\hat{z}} = e^{\hat{z}\left(\sqrt{3}-\frac{\hat{\eta}_0^2}{3\sqrt{3}}\right)} = e^{\hat{z}\sqrt{3}}\; e^{-\left(\frac{\hat{z}\hat{\eta}_0^2}{3\sqrt{3}}\right)} \tag{1.150}$$

or rewritten in non-normalized coordinates :

$$G(\eta_0,\, z) \sim e^{\frac{z}{L_g}}\; e^{-\left(\frac{z}{9\,L_g}\left(\frac{\eta_0}{\rho}\right)^2\right)} \tag{1.151}$$

For a fixed $z = L_u$, the gain-curve approaches a Gaussian in $\hat{\eta}_0 = \eta_0/\rho$ with width $\sigma_{\hat{\eta}} = \frac{3}{\sqrt{2\chi}}$ or $\sigma_\eta = \frac{3\rho}{\sqrt{2}}\sqrt{\frac{L_g}{z}}$. Notice again : this is valid for $z >> L_g$, typically $z \approx 10\, L_g$.

Instead of de-tuning the electron beam energy, we can as well de-tune the resonance frequency ω_r, that is the frequency of the initial EM field. Since $\omega \sim 1/\gamma^2$, we have

$$\eta = \left(\frac{\gamma - \gamma_r}{\gamma_r}\right) = \frac{1}{2}\left(\frac{\omega_r - \omega}{\omega_r}\right) \tag{1.152}$$

and thus

$$\frac{\sigma_\omega}{\omega_r} = 2\;\sigma_\eta = 3\sqrt{2}\;\rho\sqrt{\frac{L_g}{z}} \tag{1.153}$$

The relative **gain bandwidth** of the FEL decreases with increasing undulator depth and is, what would you expect, scaling with the ρ-parameter :-). For saturation we need about 10 gain length, that is the bandwidth will be about $\frac{\sigma_\omega}{\omega_r} \approx 1.3\,\rho$.

Start from initial density modulation

The FEL process can start not only from an initial field but as well from an initial modulation of the electron beam. As we discuss in the following section, the intrinsic "noise" of the beam current resulting from the stochastic nature of the electron distribution is sufficient to start-up the "SASE" process.

$$a_0 = 0 \tag{1.154}$$

$$b_0 = \left\langle e^{-i\,\theta_j}\right\rangle\Big|_0 \tag{1.155}$$

Since we have no initial EM field, the definition of the ponderomotive phase is not longer obvious. For the moment we assume that the beam has well defined initial density modulation with wavelength λ_m. This wavelength defines the length scale of the ponderomotive phase, that is the length of the interval $\pm\pi$ along the bunch. The corresponding $\omega_m = \frac{2\,\pi\,c}{\lambda_m}$.

The beam energy together with the undulator parameters define the "resonant λ_r" of the beam according to eq. (7). If λ_r and λ_m are not identical, this can be translated into an effective relative beam energy deviation η_0

$$\eta_0(\omega_m) = \left(\frac{\gamma_m - \gamma_r}{\gamma_r}\right) = \frac{1}{2}\left(\frac{\omega_m - \omega_r}{\omega_r}\right) \tag{1.156}$$

To illustrate the role of b_0 as "level of modulation", two examples : if the beam has a harmonic density modulation with wavelength λ_m and amplitude $a_m < 1$, the bunching strength $|b_0|$ would be $\frac{a_m}{2}$ if the phases are determined from the positions of the electrons according to

$$\theta_j = \frac{z_j}{\lambda_m}\,2\,\pi + \theta_0 \tag{1.157}$$

For any other wavelength λ, $|b_0|$ averages to zero over many periods.

If all electrons are squeezed into very thin slices with distance λ_m (extreme micro-bunching), $|b_0|$ approaches one.

Back to the road : the normalized parameter is defined as before

$$\hat{\eta} = \eta/\rho \tag{1.158}$$

We assume no "genuine" energy modulation, that is all $\eta_j = 0$. But with finite η_0, there is a "collective" energy modulation resulting from the finite density modulation:

$$\left(\frac{d\theta}{d\hat{z}}\right)\Big|_0 = \hat{\eta}_0 \tag{1.159}$$

$$\left(\frac{db}{d\hat{z}}\right)\Big|_0 = -i\left\langle\left(\frac{d\theta}{d\hat{z}}\right)e^{-i\theta_j}\right\rangle\Big|_0 = -i\,\hat{\eta}_0\,b_0 \tag{1.160}$$

$$P_0 = i\left(\frac{db}{d\hat{z}}\right)\Big|_0 = \hat{\eta}_0\,b_0 \tag{1.161}$$

So we have to start with

$$\begin{pmatrix} a_0 \\ b_0 \\ P_0 \end{pmatrix} = \begin{pmatrix} 0 \\ b_0 \\ \hat{\eta}_0\,b_0 \end{pmatrix} \tag{1.162}$$

The coefficients C_l are then

$$C_1 \to i\,b_0\,\frac{(\mu_2+\mu_3-\hat{\eta}_0)}{(\mu_1-\mu_2)\,(\mu_1-\mu_3)},\; C_2 \to -i\,b_0\,\frac{(\mu_1+\mu_3-\hat{\eta}_0)}{(\mu_1-\mu_2)\,(\mu_2-\mu_3)},\; C_3 \to -i\,b_0\,\frac{(\mu_1+\mu_2-\hat{\eta}_0)}{(\mu_1-\mu_3)\,(-\mu_2+\mu_3)} \tag{1.163}$$

on resonance, $\eta_0 = 0$

Putting in the general μ_n results in bulky expression, for $\eta_0 = 0$ (beam energy and modulation in resonance) it comes out handy :

$$C_1 \to -i\,\frac{b_0}{3},\; C_2 \to (-1)^{5/6}\,\frac{b_0}{3}\;,\; C_2 \to (-1)^{1/6}\,\frac{b_0}{3} \tag{1.164}$$

In this case, the scaled radiation power increases as (with $\chi = z/L_g$)

$$|a|^2 = \frac{2}{9}\,{b_0}^2\,G[\chi] \tag{1.165}$$

with

$$G[\chi] = \left(\text{Cosh}[\chi] + \sqrt{3}\;\text{Sin}\Big[\frac{\sqrt{3}\;\chi}{2}\Big]\,\text{Sinh}\Big[\frac{\chi}{2}\Big] - \text{Cos}\Big[\frac{\sqrt{3}\;\chi}{2}\Big]\,\text{Cosh}\Big[\frac{\chi}{2}\Big]\right) = \tag{1.166}$$

$$G[\chi] = \frac{1}{2}\,(e^{-\chi} + e^{\chi}) - \text{Cos}\Big[\frac{\sqrt{3}\;\chi}{2}\Big]\,\text{Cosh}\Big[\frac{\chi}{2}\Big] + \sqrt{3}\;\text{Sin}\Big[\frac{\sqrt{3}\;\chi}{2}\Big]\,\text{Sinh}\Big[\frac{\chi}{2}\Big] \tag{1.167}$$

For large χ the scaled energy density goes as

$$\Big|a\Big|^2 \approx \frac{1}{9}\,{b_0}^2\,e^{\chi} = \frac{1}{9}\,{b_0}^2\,e^{z/L_G} \tag{1.168}$$

Notice that the power scales with the initial bunching factor *squared*. The power increase from an initial bunching is shown in figure (26).

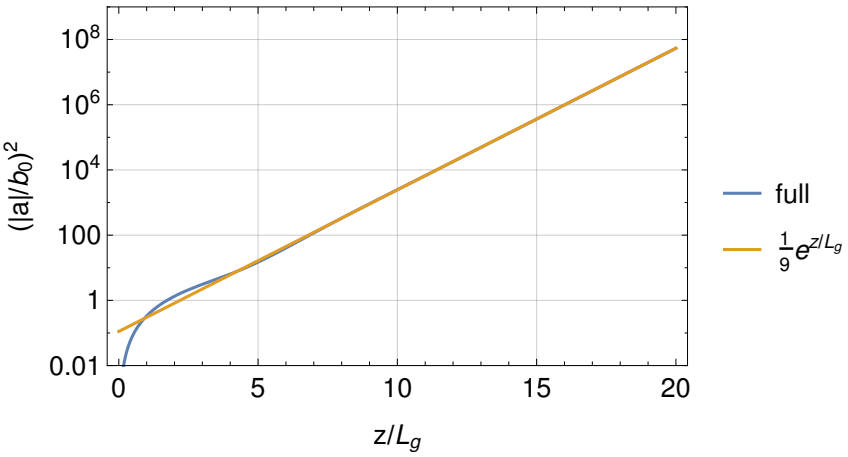


Figure 26: Gain as function of initial energy deviation η_0 at different undulator depth.

Since $|a|$ is restricted to $|a| < 1$, a modulation amplitude of say 10% ($b_0 = 0.1/2$) would lead to saturation after about 8 gain length. With $b_0 = 10^{-4}$, about 20 gain length are needed to reach saturation.

off resonance, $\eta_0 \neq 0$

As we did in the case of field amplification we can now study the influence of an energy deviation, that is a mismatch of beam energy and wavelength of the initial modulation. The procedure is the same as we did in the previous section: insert the general solutions for μ_{1-3} (Eq. 141) into the expressions for C_{1-3} (Eq. 163) and sum up the three contributions according to (Eq. 112). The power evolution is now a function of both, $\chi = z/L_G$ and the (scaled) energy deviation $\hat{\eta_0} = \eta_0/\rho$.

$$|a|^2 = \frac{2}{9} b_0{}^2 \, G[\chi, \hat{\eta}_0] \tag{1.169}$$

The general expressions are lengthy, but for sufficiently large χ (in the exponential regime) it can be approximated by

$$\left|a\right|^2 \approx \frac{1}{9} b_0{}^2 \, e^{\chi} \, e^{-\frac{\hat{\eta}_0{}^2}{9\chi}} \tag{1.170}$$

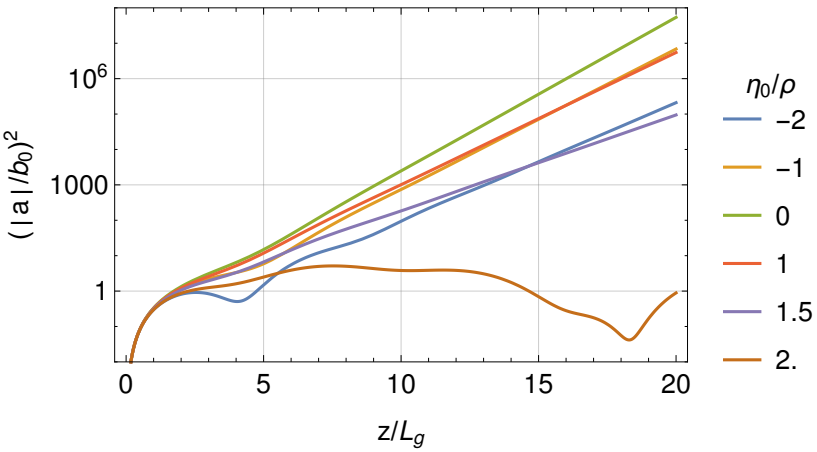


Figure 27: Gain as function of initial energy deviation η_0 at different undulator depth.

As expected the "on-resonance" modulation has the fastest growth. As discussed above, for $(\hat{\eta}_0)^3 > \frac{81}{12}$ no exponential regimes exists, μ_2 has no positive imaginary part, the growth rate function is zero.

If we plot the "normalized power" for fixed undulator depth χ, we get the curves shown in figure (28):

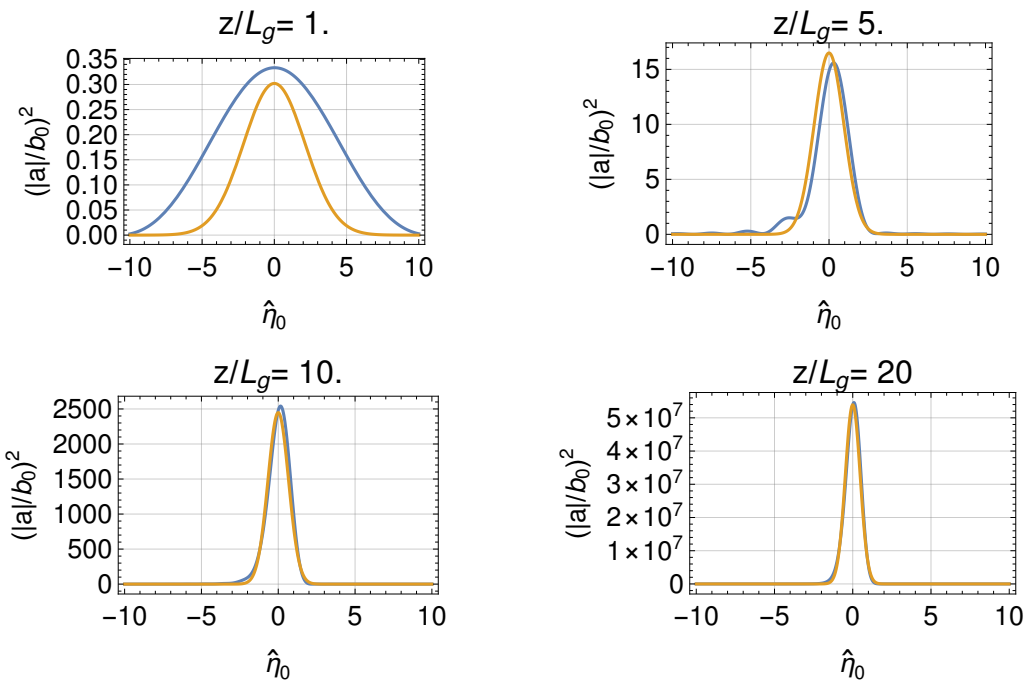


Figure 28: Gain as function of initial energy deviation η_0 at different undulator depth.

Comparing with the gain curves for an initial EM field we observe a striking difference: independent of the undulator depth, the highest power is always resulting from an "on resonance" modulation, that is $\hat{\eta}_0 = 0$. There is no "Madey like" bipolar curve as for the amplification of an initial EM field.

The reason for that is easy to understand. If the beam has an initial density modulation, this modulation generates an electric field with a phase with respect to the modulation such that energy is transferred from the beam to the field. This field has the *correct* phase for amplification, the average phase difference between field and density modulation is zero.

In the case of an initial EM field, this field first generates an energy modulation (in phase) leading to a density modulation shifted by $\pi/2$ in phase. This density modulation is *even* with respect to the zero crossing of the field (θ=0), energy loss and gain are balanced. In the case of $\eta_0 = 0$, the phases are to first order stable and there is no net energy transfer between beam and field. A finite η_0 is needed to make the beam phase "run" with respect to the field. The modulation is shifted with increasing z, becomes unbalanced with respect to θ=0 resulting in a net energy transfer.

As given in Eq. (170), for large z/L_G, the "power curve" is in very good approximation given by a Gaussian of width $\sigma_{\hat{\eta}} = \frac{3}{\sqrt{2}\chi}$ or $\sigma_\eta = \frac{3\rho}{\sqrt{2}}\sqrt{\frac{L_G}{z}}$, identical to what we found for the gain curve from an initial field.

Starting from an initial density modulation is important to understand the physics of the "SASE" process, "Self-Amplified Spontaneous Emission", and a variety of modern seeding schemes. Both will be discussed in the next sections.

Energy spread and phase space distributions

In a real accelerator, the electrons have a certain energy distribution. Especially after compressing the bunches to high peak currents and a variety of nasty collective effects (like space charge and coherent synchrotron radiation in the compressor dipols), the phase-space distribution of the bunch normally looks quite complex. Such "real condition" cannot be easily simulated in our simple 1-D harmonic model.

Nevertheless, the influence of a finite energy distribution can be estimated using the coupled equations and appropriate initial conditions. Another complication arises from the fact the we cannot simulate the real number of electrons in the slices we look at, the number of macroparticles we are using typically is much less. If we would distribute the initial η of the macroparticles randomly, this introduces a very strong "seeding component" in the initial distribution which completely dominates the simulation leading to irregular and very high gain. So we have to further simplify the model by using a regular energy distribution, as we did for the phases before.

As usual, we normalize the width of the flat-top energy distribution to ρ, our "one an only" parameter. The resulting gain evolutions are shown in figure (29).

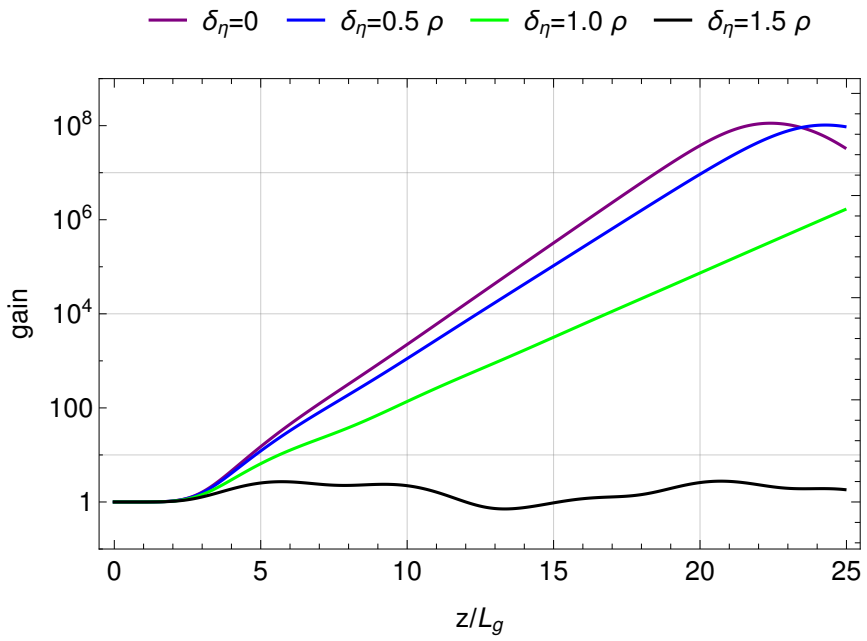


Figure 29: Gain evolution with initial (flat-top) energy spread δ_η.

To keep the gain length close to the optimal value, the **initial energy spread has to be less than ρ**. This is a rather stringent requirement especially for FEL in the nanometer and sub-nanometer regime where ρ typically is a small number below 10^{-3}.

In our simple 1-D model, the bunches are uniform in z, the longitudinal coordinate. In reality, the bunches are not only of finite length but as well are non-uniform. For the FEL process, the decisive length is the "cooperation length", the scale over which electrons act coherently in producing the EM wave. This coherence length is related to the spectral bandwith by the time-bandwidth product, that is $\tau_c * \sigma_\omega \approx 1$. From eq.(153) we learned that for a long undulator $\sigma_\omega \simeq \rho\,\omega$, that is $c\,\tau_c \simeq \lambda/(2\,\pi\,\rho)$. For typical ρ, the coherence length is a few hundred wavelengths, that is for a nm-FEL in the μm regime while the total bunch length is typically some 100 μm. In consequence, it is not the *overall* energy spread influencing the gain evolution but the local or *slice energy spread*.

While an as small as possible (slice) energy spread is advantageous for keeping L_g short, it has the unwanted side effect of creating longitudinal instabilities during the acceleration and bunch compression process. (Notice : the FEL process is exactely such an instability). To avoid these detremental effects, the slice energy spread can be artifically be worsened such that the optimal balance, leading to the shortes L_g, is found. This is done by a "laser heater", a device in which energy from an external laser is transferred *to* the electron beam.

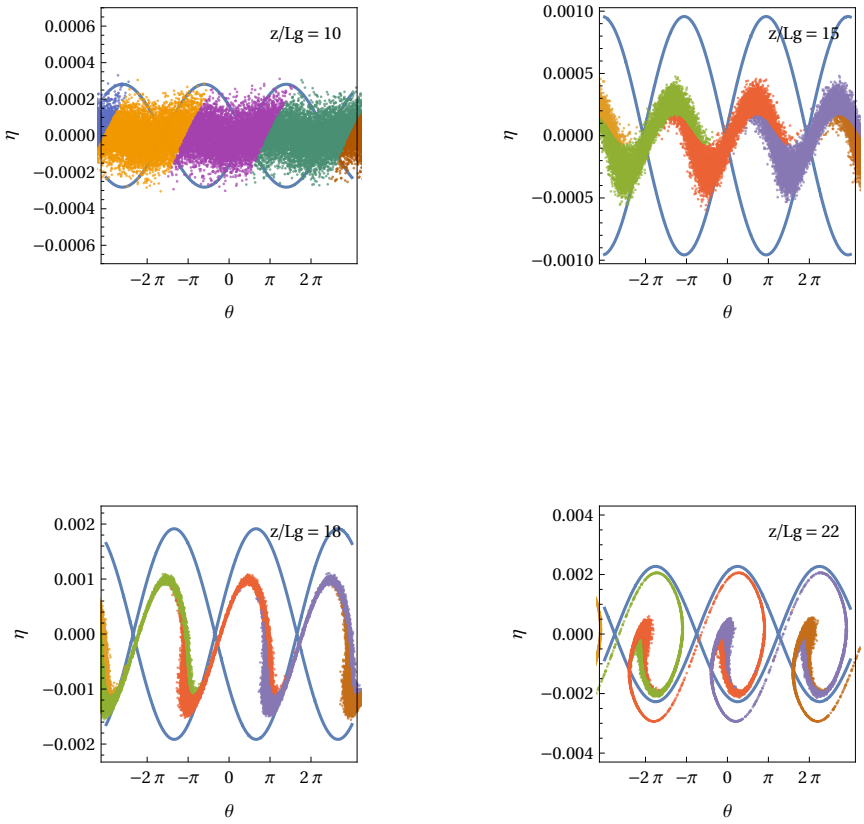


Figure 30: Phase space distribution of at different depths in the undulator. The initial energy spread is $\sigma_\eta = 0.1\,\rho$.

A typical phase space evolution is shown in figure (30). Up to 15 L_g, the energy modulation looks pretty harmonic, the overall energy loss is barely visible. At 18 L_g and deeper, the phase space distribution is dominated by the "trapping" inside the FEL buckets. Notice that the buckets are not at fixed positions along θ but move due to the phase shift of the EM field (figure 21).

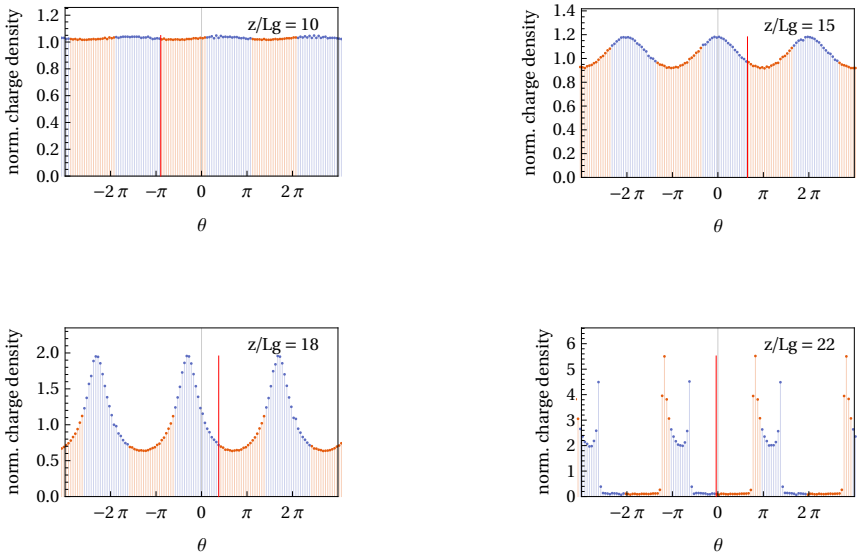


Figure 31: Density modulation in various depths of the undulator. The red line indicates a zero crossing of the EM field.

The phase space plots are not really well suited to see the densitiy modulation and especially how it phases with the EM field. In Fig. (31) finally we look at the densitiy modulation at different depth in undulator. The vertical red line indicates a zero crossing of the EM field, the blue and red colours of the density points indicate parts of the bunch loosing energy (blue) and gaining energie (red) from the field. At 10 L_g, the density modulation is a few percent effect. The phases have shifted such that slightly more electrons lose energy ("stimulated emission") than gain energy ("absorption") from the field. At 15 L_g, we see a rather perfect harmonic density modulation of about 20% peak-peak. At 18 L_g, the modulation becomes unharmonic and at 22 $L_g$the electrons are bunched in slices. The "unphysical" sharp edges result from the "unphysical" flat-top energy distribution at the beginning.

SASE

"Self amplified spontaneous emission" (SASE) was proposed in the 1980th as a method to produce high power FEL radiation for short wavelengths where no powerful initial "seed field" exists. The basic idea is to use the stochastic nature of the charge density distribution in the electron beam to seed the FEL process. This can be considered in two ways:

a) the electron beam generates spontaneous undulator radiation in the beginning of the undulator, this spontaneous radiation is picked up by the FEL process and amplified.

b) the random longitudinal distribution of electrons in the beam leads to a non-vanishing bunching factor b_0 at the resonant frequency of the (beam - undulator) system which starts the FEL process as described in the previous section.

Wise men have proven that both pictures are fully equivalent.

A comprehensive discussion of the SASE process is beyond the scope of this lecture, but some basic understanding can be gained from what we discussed so far.

If we follow the line b) of how the process starts, we have (according to the previous section) the find out the initial bunching amplitude $|b_0|$ of a statistically distributed beam. First of all we have to consider that the distribution of the electrons, that is the θ_j, is purely random. For different "samples" of N beam electrons, $|b_0|$ will not have a unique value but follow a statistical distribution with mean value $(|b_0|)_{\text{avg}}$. So we have carefully to distinguish the two types of averages we are taking about : the "ensemble average" denoted by $\langle\ \rangle$ and the average over many ensembles denoted $(\ )_{\text{avg}}$.

Since SASE starts from a random distribution, each power evolution will look different and lead to a different radiation power at the end of the undulator. This is a fundamental difference to the "field amplification" and "start a from well defined initial modulation" which we discussed so far.

Initial bunching : mean value and fluctuations

The bunching factor is defined as the ensemble average

$$|b_0| = \left|\left\langle e^{i\theta_j}\right\rangle\right|_0 = \frac{1}{N}\left|\sum_{j=1}^{N} e^{i\theta_j}\right| \tag{1.171}$$

with N, the number of contributing electrons, to be defined in a minute. Since the power resulting from the initial modulation is proportional to $|b_0|^2$, the average SASE power should be proportional to

$$\left(|b_0|^2\right)_{\text{avg}} = \frac{1}{N^2}\left(\left|\sum_{j=1}^{N} e^{i\theta_j}\right|^2\right)_{\text{avg}} \tag{1.172}$$

If we consider $e^{i\theta_j}$ as a unit vector in the complex plane with arbitrary angle, we ask : what is the length squared of the sum of N such unit vectors of random orientation ? Without lengthy mathematics we look at an analogy : a person does N successive steps starting from the origin, each step has unit length but goes into random orientation. (This is the classical random walk problem). The mean value of his position after Nsteps is still the origin, since no direction is preferred, but the mean distance of his position to the origin increases with $\sqrt{N}$. For large N, the "distance distribution " is a Gaussian with variance $\sigma_d{}^2 = N$.

Thus we conclude (or "guess"), that the averaged sum in Eq. (172) is N and the average bunching factor is

$$\left(|b_0|^2\right)_{\text{avg}} = \frac{1}{N} \tag{1.173}$$

This result is not so surprising : the more electrons contribute to the ensemble average, the "smoother" the resulting distribution is and thus the less the average bunching factor resulting from the stochastic nature.

To verify this ideas and to get hand on the distribution of $|b_0|^2$ we do a bit of "experimental mathematics", using *Mathematica*. For a fixed number of ensemble particles N, we create random phase angles and determine the ensemble average b_0 and its modulus squared. This we repeat many times and ask for the resulting mean value of the modulus squared :

```
ensembleavg[n_Integer] := Abs[Mean[Exp[-I RandomReal[{-π, π}, n]]]]^2

manysammples[npart_Integer, nchecks_Integer] := Table[avgbun[npart], {nchecks}]
```

So we check for various N between 100 and 10^6 with 1000 samples each :

```
setlist = {100, 1000, 10 000, 100 000, 10^6}

fulltab = Table[manysammples[npart, 1000], {npart, setlist}];
```

Here is the result :

```
TableForm[Transpose[{1 / Mean /@ fulltab, (setlist - 1 / Mean /@ fulltab) / setlist}],
 TableHeadings → {ToString /@ setlist, {"exp.", "rel.dev."}}]
```

	exp.	rel.dev.
100	99.0125	0.00987471
1000	928.407	0.0715928
10000	9760.24	0.0239763
100000	99 773.2	0.00226833
1000000	979 430.	0.0205697

Our considerations seem to be OK, Eq. (173) is correct.

Since we have the data at hand, we can as well look at the distribution of the ensemble averages.

Figure (32) shows the "experimental" probability distribution for the two extreme cases, $N = 100$ and $N = 10^6$. In both cases find an exponential distribution, the solid lines are the PDF

$$p(|b_0|^2) = N\, e^{-|b_0|^2\, N} \tag{1.174}$$

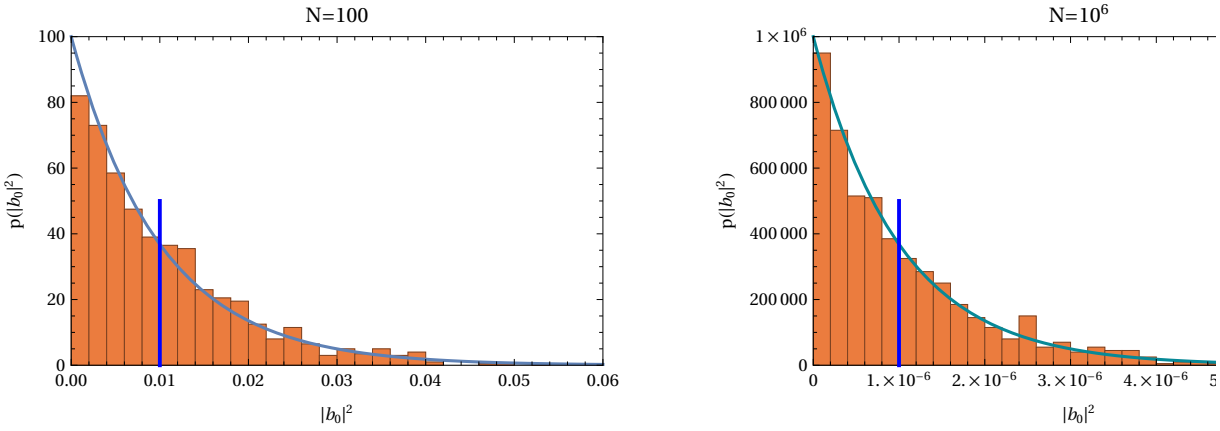


Figure 32: Probability distribution of the bunching factors for 100 and 10^6 electrons in the sample.

the blue lines the average $1/N$.

Before saturation, the FEL process is a "linear amplifier" and these initial fluctuations will be translated into proportional fluctuations of the output power. More about this in the next section.

Back to the average $|b_0|^2$ which we found out to be $1/N$. How many electrons are contributing to the SASE start-up ? The full bunch ? One FEL bucket ?

The correct answer to this question is tricky, so handle it with a simplified but useful consideration based on the well known relation between frequency bandwidth $\Delta\omega$ and duration $\Delta\tau$ of an EM wave packet.

The EM radiation field has, as we learned, a finite frequency bandwidth and the Fourier transform tells us, that this is coupled with a finite duration of the pulse which is defined by the "time-bandwidth product". So we have for the coherence time of the radiadion field

$$\tau_c = \sqrt{\pi} \,\Big/\, \sigma_\omega \approx \sqrt{\pi} \,\Big/\, (\rho\, \omega) = \frac{\lambda}{2\sqrt{\pi}\ \rho\, c} \tag{1.175}$$

where we used from equ. (153) that for a long undulator the bandwidth is about $\sigma_\omega \approx \rho\, \omega$.

The number of electrons starting up the SASE process is then simply the number of electrons within one coherence length (c τ_c) $= \frac{\lambda}{2\sqrt{\pi}\ \rho}$.

So we conclude: for a typical $\rho = 10^{-3}$, the coherence length is about 300 wavelengths or for a nanometer FEL below 1 μm and thus much shorter than the typical bunch lengths (10 - 100 μm).

If we have n_e the electron density in the bunch and σ_x its lateral width, the number of "cooperating electrons" becomes

$$N_c = n_e\left(2\,\pi\,{\sigma_x}^2\right)(c\,\tau_c) \;=\; \frac{\sqrt{\pi}\;\; n_e\,\lambda\,{\sigma_x}^2}{\rho} \tag{1.176}$$

Thus we have for the average bunching factor squared:

$$\left\langle {b_0}^2\right\rangle = \frac{1}{N_c} = \frac{\rho}{\sqrt{\pi}\;\; n_e\,\lambda\,{\sigma_x}^2} = \frac{2\,c\,e\,\sqrt{\pi}\;\,\rho}{I_b\,\lambda} \tag{1.177}$$

where we replaced the electron density by the (local !) beam current $I_b = \; n_e\left(2\,\pi\,{\sigma_x}^2\right)c$.

Typical numbers for eg. FLASH : $I_b = \;3$ kA, $\lambda = 4$ nm , $\rho = 5.\,10^{-4}$ results in $\left\langle {b_0}^2\right\rangle \approx 10^{-8}$! From the considerations in the section "starting from initial modulation" we know that this means that about 20 gain lengths are needed to reach |a| ~ 1, that is saturation.

wavelength and time structure fluctutions

Another importat conseqence of the fact that the "cooperation length" being considerably shorter than the total bunch length is the fact, that within one bunch, several "areas" can start a SASE process individually. These independent regimes are called "modes". For each mode, the frequency with the largest bunching factor will grow most rapidly and thus dominate, as long as the frequency is within the gain bandwidth. This results in a rather complex time and frequency pattern of the radiation produced by a SASE-FEL. Neither the time profile nor the frequency spectrum are "smooth", they both consist of a number (the number of surviving "modes") of spikes.

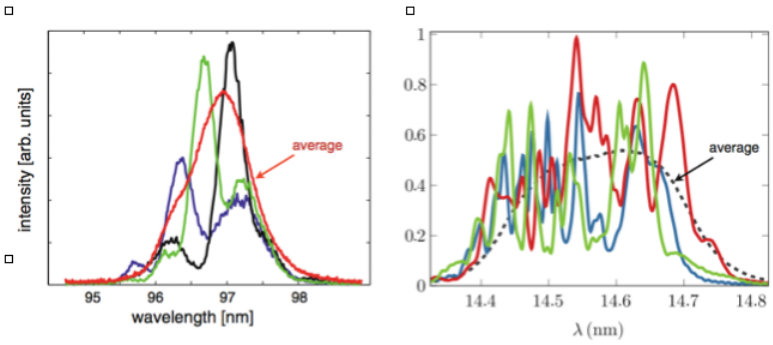


Figure 33: Single shot SASE spectra measured at FLASH at two different wavelength (from the blue book).

Figure (33) shows two examples of single shot spectra measured at FLASH. Since the modes fluctuate on a statistical basis, the wavelength spectrum fluctuates accordingly from shot to shot.

intensity fluctuations

As we have analyzed above (eq. 174), a fixed number of cooperating electrons leads to a certain *average* $\left\langle {b_0}^2\right\rangle$, but the probability distribution of ${b_0}^2$ is very wide with an exponential distribution function. The most probable ${b_0}^2$ is zero, that is nothing happens, no FEL process is started, the fluctuations width σ_{b^2} is 100%. From eq. (168) we see that for a given depth in the undulator z in the exponential regime, the resulting FEL power (people call it "SASE power") is directly proportional to ${b_0}^2$ and fluctuates as well from shot to shot with an exponential distribution. A highly unstable lightsource, indeed.

If only one "mode" contributes, the relative power distribution is exponential:

$$p(u)\,\mathrm{du} \;=\; e^{-u}\,\mathrm{du} \tag{1.178}$$

$\mathrm{u} = U_{\mathrm{rad}}/\langle U_{\mathrm{rad}}\rangle$.

But as we learned, not only one "mode" but several individual modes produce radiation from a single electron bunch and we have to sum up their individual contributions. The resulting statistical distribution from summing M distributions with identical exponential distribution is given by

$$p(u)\,\mathrm{du} = \frac{M^M\, u^{M-1}}{\Gamma(M)}\, e^{-u}\,\mathrm{du} \tag{1.179}$$

with Γ the Gamma function and thus p(u) called a "Gamma distribution".

This can also be veryfied experimentally, figure (34) shows measured intensity distributions for two different average number of modes compared to the expected Gamma distributions (eq. 179):

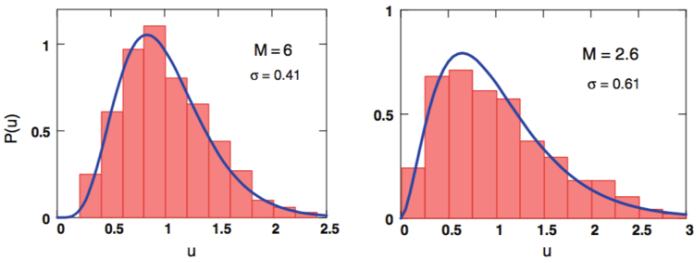


Figure 34: Experimental power distributions for two different numbers of contributing modes M. To vary M, the effective bunch length was changed.

The different M were realized by using two different overall bunch lengths. To measure the pure exponential M=1 distribution is possible as well, even without compressing the bunch to (unrealistic) short lengths, by selecting a very narrow frequency by using a monochromator. The result is shown in figure (35).

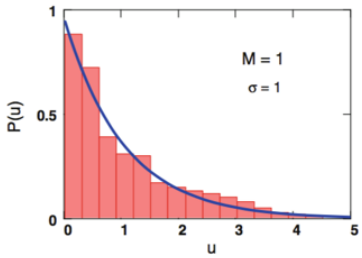


Figure 35: Experimental power distribution for a single FEL mode. The mode was selected by using a narrow band spectral filter for the FEL radiation (monochromator).

As mentioned above, the intensity fluctuations as described here are valid in the exponential regime of the gain curve. If it is possible to drive the power up to the level where it levels off even for the modes with the smalles bunching factor, the intensity fluctuations are supressed since the saturation level is independent of the initial bunching factor. In practise, the fluctuation level in saturation is typically of the order 20%, compared to the 100% in the M=1 case.

Seeding

To overcome the problems of the "SASE" process, a variety of "seeding schemes" have been proposed and used since quite some time.

If a radiation source of sufficientyl large power exists, seeding with an initial EM field is the most direct way. The problem is, that the intensity of this seeding field has to be large enough to overcome the start-up from the unavoidable SASE process. For short wavelengths, this quite demanding, in the sub-nm regime no such sources exist.

Nevertheless, a variety of tricky methods have been found to to seeding in the nm regime and even for wavelengths as short as 0.1 nm.

Unfortunately, I don't have the time now to describe them here even shortly, I have to refer to the many publications and few textbooks about the subject.