Numerical Studies for a Single Pass High Gain Free-Electron Laser

Dissertation
zur Erlangung des Doktorgrades
des Fachbereichs Physik
der Universität Hamburg

vorgelegt von
Sven Reiche
aus Jork

Hamburg
1999

Gutachter der Dissertation:	Prof. Dr. B. Sonntag Dr. J. Roßbach
Gutachter der Disputation:	Prof. Dr. P. Schmüser Prof. Dr. G. Materlik
Datum der Disputation:	6.1.2000
Dekan des Fachbereichs Physik und Vorsitzender des Promotionsausschusses:	Prof. Dr. F.-W. Büßer

Deusque Dixit Fiat Lux
Et Facta Est Lux

Genesis 1.3

Abstract

The importance of single pass high gain Free-Electron Lasers (FEL) is due to its the unique characteristic of the short, high intense radiation pulse at a well defined, tunable frequency. Extending the radiation wavelength to the ultraviolet or X-ray region is an on-going research project and will open new branches of experiments based on this radiation.
Several projects are proposed or currently under construction such as the integrated X-ray FEL at the TESLA linear collider and the VUV FEL at the TESLA Test Facility. For all of these projects a detailed study is of importance, including all effects which might influence the FEL performance. In particular, for an FEL operating at X-ray wavelengths, the amplification process is more sensitive to any kind of disturbance.
Some of these perturbations influence the FEL performance in a complex way. Because existing codes are not capable to cover these problems the simulation code GENESIS 1.3 has been developed. The new features of GENESIS 1.3 includes the discretization of the radiation field on a Cartesian grid and the input of arbitrary profiles of the undulator field and the electron beam in longitudinal direction. The code is capable of covering aspects such as beam halos, wake fields and non-periodic focusing structures to name a few.
Using GENESIS 1.3 special aspects of the TTF-FEL and TESLA FEL have been studied. The simulation covers the transverse motion of the electron beam, the impact of wake fields as well as the increase of the energy spread due to the quantum fluctuation of the incoherently emitted radiation.

Zusammenfassung

Die Bedeutung von Freie-Elektronen Lasern (FEL), die eine große Verstärkung innerhalb eines einzelnen Durchlaufs des Elektronenpakets durch den Undulator erzielen, liegt in der einmaligen Charakteristik der kurzen, hochintensiven Strahlungspulse bei einer durchstimmbaren Frequenz. Die momentante Forschung beschäftigt sich mit der Verkürzung der erreichbaren Wellenlängen bis zum Röntgenbereich. Dieses würde neue Experimente, die sich diese Strahlung zu Nutze machen, ermöglichen.

Mehrere Projekte sind vorgeschlagen oder befinden sich bereits im Aufbau, die in diesen Wellenlängenbereich vorstoßen, so auch der im Linearbeschleuniger TESLA integrierte Röntgenlaser und der Freie-Elektronen Laser der TESLA Test Facility (TTF). Solche Projekte können nicht ohne detaillierte Untersuchungen vor und während des Betriebs realisiert werden, wobei alle störenden Effekte mit eingeschlossen seien sollten. Gerade bei kurzen Wellenlängen ist die FEL Verstärkung anfällig für jeglicher Art von Störung.

Einige der Störungen beeinflussen die FEL Verstärkung auf komplexer Weise. Da existierende Simulationsprogramme diese Probleme nicht korrekt behandeln können, wurde das Programm GENESIS 1.3 entwickelt. Die neuen Eigenschaften von GENESIS 1.3 beinhalten die Diskretisierung des Strahlungsfeldes auf einem kartesischen Gitter sowie die Eingabe von beliebigen Profilen des Undulatorfeldes und Elektronstrahls in longitudinaler Richtung. Damit sind Simulationen möglich, die Effekte wie z.B. Elektronenhalos, externe Störfelder oder eine nicht-periodische Fokusierung beinhalten.

Unter Ausnutzung dieser Möglichkeiten wurden bestimmte Aspekte des TESLA FEL und TTF-FEL untersucht. Die Spannweite reicht von der transversalen Bewegung der Elektronen, über induzierte elektrische Störfelder des Elektronenpakets bis zur Aufweitung der Energieverteilung durch die Quantenfluktuation von inkohärenter Röntgenstrahlung.

Contents

Chapter 1

Introduction

The achievements of the research on radiation sources have already found a firm place in today's life. The most common offspring of this research is the laser. The acronym stands for **L**ight **A**mplification by **S**timulated **E**mission of **R**adiation. The span of laser applications is wide and covers many aspects such as network communication or CD players.

The main principle of these lasers is based on the transition of electrons between quantum states in an optical medium always connected with the emission or absorption of photons (the quantum of radiation). Beside the spontaneous emission as the natural decay of excited electrons states the transition can be stimulated if an external field is applied with the same wavelength of the spontaneous emission. In the case that the higher quantum states are populated by more electrons than the lower ones, the emission dominates and a small radiation field, initially generated by the spontaneous emission, gets amplified. Beside the well defined properties of this radiation source in comparison to conventional light sources (light bulbs), the fixed frequency f of these lasers is defined by the quantum states of the optical medium with $\Delta E = hf$, where ΔE is the difference in the electron energy of the two quantum states and h is the Planck constant.

Although the radiation of accelerated electrons differs from radiation of conventional lasers, the bremsstrahlung of electrons in a periodic magnetic field can be stimulated by an externally applied radiation field as pointed out by Madey in 1971 [1]. Using a quantum mechanical description, where the stimulated emission is based on population inversion between quantum states of the electron beam, the analogy to conventional lasers became obvious. Because the electrons are not bound to any optical medium, such as atoms, molecules of crystals, Madey named this device **F**ree-**E**lectron **L**aser (FEL). Five years later Colson [2] published an equivalent description using classical mechanics. The experimental verification was in 1976 [3].

In the early stage of FEL research the amplification of the seeding radiation field was rather low, which creates the necessity of an optical feedback system. Enclosing the undulator or wiggler, which provides the periodic magnetic field, by an optical cavity reflects the amplified radiation back to the interaction region with the electrons. The gain is accumulated over several oscillation of the radiation field in this cavity, defining the FEL oscillator.

With a new generation of injectors, based on photo-electron guns [4], the electron beam quality became sufficient to reach saturation of the FEL amplification within a single pass of the electron

bunch through the undulator. This opens also the possibility to extend the wavelength region to the VUV and X-ray regime, where mirrors are not opaque and an FEL oscillator would not work.
The increase of the interaction efficiency between radiation field and electron beam yields a collective instability of the electron beam, which was first theoretically derived by Kontradenko and Saldin in 1980 [5] and 4 years later by Bonifacio, Pellegrini and Narducci [6]. This collective instability exhibits an exponential growth in the radiation power with an amplification of several orders of magnitude per single pass. Several experiments [7] – [16] have confirmed this high gain amplification down to wavelengths below 1 μm. To extend the wavelength range to the VUV or X-ray region, new Free-Electron Lasers have been proposed [17] – [20].

1.1 The Working Principle of Free-Electron Lasers

The working principle of a Free-Electron Laser is best explained in three steps: the spontaneous undulator radiation, the low gain FEL and the high gain FEL. This order coincides with the historical evolution of the Free-Electron Laser as briefly discussed above.
An undulator or wiggler provides a periodic magnetic field along its main axis. A relativistic electron, propagating along the undulator axis, is deflected by the magnetic field and oscillates transversely with the period λ_U of the magnetic field. Any accelerated charge will radiate, which can be approximated in this case by a Hertz dipole propagating through the undulator with almost the speed of light. The relativistic Doppler effect bundles most of the radiation in the forward direction and shortens the resonant wavelength to

$$\lambda_0 = \frac{\lambda_U}{2\gamma^2}(1 + K^2) \quad . \tag{1.1}$$

Because the magnetic field excites a transverse oscillation of the electron the longitudinal velocity is slowed down. This fact is covered by the dimensionless undulator parameter K in Eq. 1.1, which is a measure for the strength and period of the magnetic field and thus for the amplitude of the transverse motion.
Due to the finite length of the undulator the spontaneously emitted radiation pulse has a total length of $N_U \lambda_0$, where N_U is the total number of undulator periods. Even without knowing the exact rate of emission, the shape of the power spectrum (Fig. 1.2) can easily be determined by a Fourier transformation. The general dependence on the radiation frequency is

$$I(\omega - \omega_0) \propto \frac{\sin^2\left(\frac{(\omega-\omega_0)}{\Delta\omega}\right)}{\left(\frac{(\omega-\omega_0)}{\Delta\omega}\right)^2} \tag{1.2}$$

with $\omega_0 = 2\pi c/\lambda_0$ and $\Delta\omega = \omega_0/\pi N_U$. The narrow bandwidth $\Delta\omega$, which is inversely proportional to the undulator length, makes even the spontaneous undulator radiation attractive to many scientific users.
The stimulated emission of the low gain FEL is rather complicated if the analogy to a conventional laser is followed. In this frame the shape of the spectrum in Eq. 1.2 is regarded

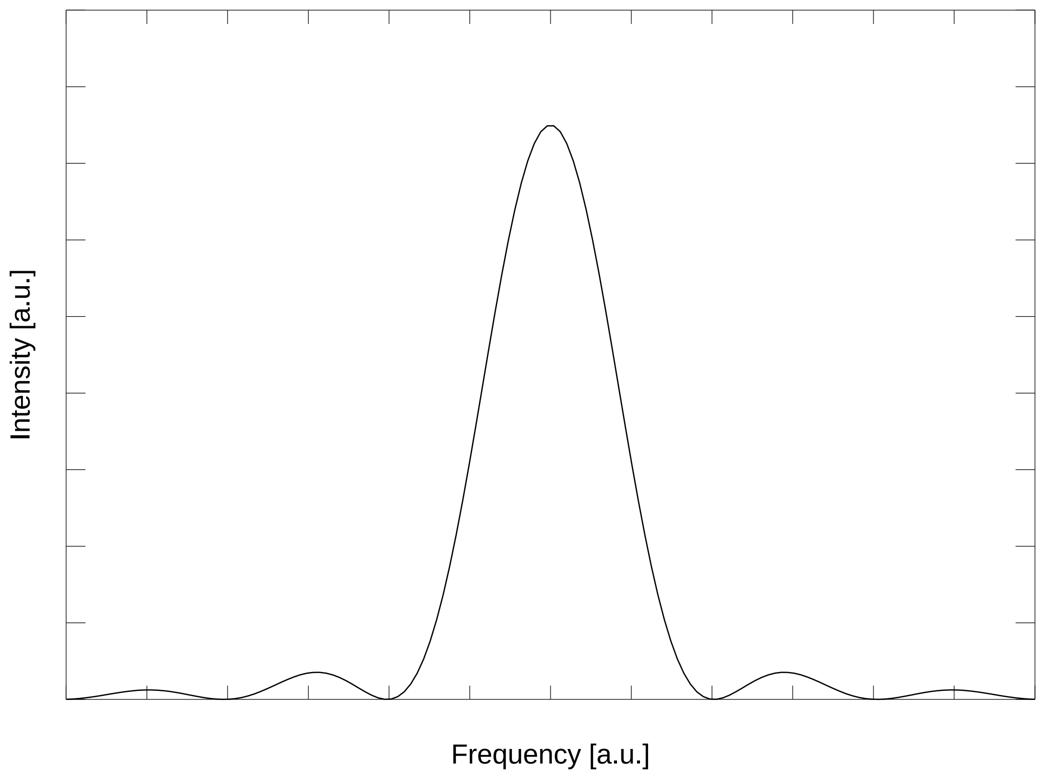


Figure 1.1: *Spectrum of a homogeneously emitted radiation pulse.*

as the probability distribution of an electron to either emit or absorb a photon of a certain wavelength. A more careful analysis has to include the recoil of the electron. The absorption and emission spectrum of the electrons is therefore slightly shifted in opposite direction. This separation yields a difference in the absorption and emission rate, which is necessary to provide a stimulated emission at certain frequencies. Because the recoil shift is small the probability distribution can be expand into a Taylor series around the emitted frequency. The explicit calculation gives a dependence of the gain for stimulated emission (Fig. 1.2) on the frequency ω which is the negative differential of the spectrum of the spontaneous emission. This is knows as Madey's theorem [21].

In an equivalent and more classical approach the low gain FEL amplifier can be understood if the electron motion in the longitudinal phase space is considered. Due to the transverse motion the electrons are able to couple to the electric field component of the radiation field. Depending on the phase the electrons are either accelerated or decelerated. This modulation in the velocity results in a different longitudinal drift compared to the synchronized electron at resonance energy, where the radiation wave advances exactly one period of the resonant wavelength λ_0 during one transverse oscillation of the electrons. If the electron beam is injected at resonance energy and equally distributed over one radiation wavelength the energy gain of a

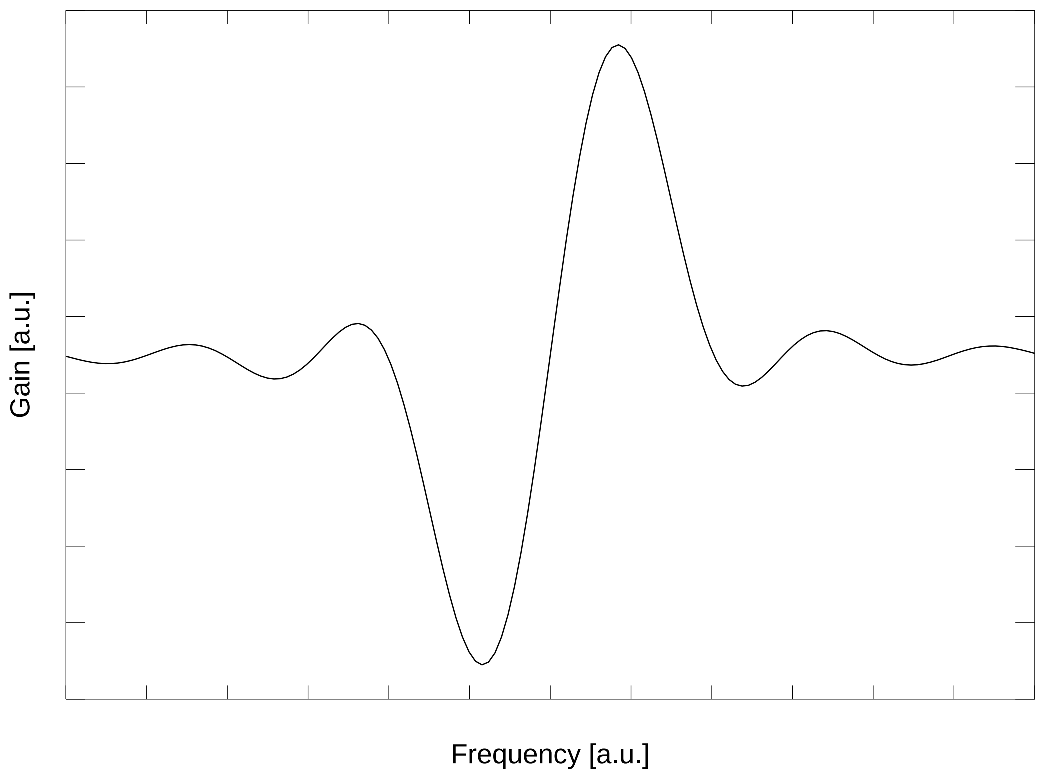


Figure 1.2: *Spectrum of the low gain Free-Electron Laser.*

certain electron is compensated by the energy loss of an electron at a different radiation phase. The net electron beam energy does not change and the radiation field remains unamplified. The situation differs if the electron beam is injected above resonant energy. This causes a slippage of the electrons with respect to the radiation field. Electrons, which decalerate by losing energy, become more synchronized with the radiation field. Gaining energy has a different effect because the slippage is enhanced. Therefore electrons are accumulated in a certain phase range, where the energy loss dominates. The net transfer of energy from the electron beam to the radiation field is positive, providing the amplification of the radiation field. The explicit calculation yields the same results as Madey's model, is easier to understand, but lacks the intrinsic definition of a laser.

The classical model is also used to derive the theory of the high gain Free-Electron Laser. The model is extended by Maxwell's equations to include the growth of the radiation field in a self-consistent manner. A high gain FEL works slightly different than the low gain one because the driving force is not the net energy transfer of the electron beam to the radiation field but the bunching of the beam at a certain radiation phase. The modulation of the electron beam energy causes different longitudinal velocities where the faster electrons catch up with the slower ones. The beam tends to bunch periodically with the radiation wavelength and the coherent emission is enhanced. The stronger radiation field forces the electrons to bunch even faster and a collective instability occurs, where the radiation field amplitude increases exponentially. The amplification saturates when the maximum bunching of the electron beam has been obtained. If the characteristic length of the exponential growth is much shorter than the undulator length the FEL saturates within a single pass of the electron beam. This radiation source is tunable by varying the electron energy and the radiation wavelength can be extended to the X-ray region. One problem for these future devices is that the FEL cannot be seeded by conventional radiation sources anymore. The start-up of the laser has to rely on the initial spontaneous emission at the resonance frequency. These **S**elf-**A**mplified **S**pontaneous **E**mission (SASE) Free-Electron Lasers are the topic of the current research to develop a high intense, tunable radiation source to any kind of experiments at wavelengths in the VUV and X-ray regime.

1.2 Outline of this Thesis

Although the basic working principle of a Free-Electron Laser is rather fundamental any calculation of the power evolution involves the coupled Maxwell's equations for the electric and magnetic field and Hamilton equations of motion. Under some assumptions and approximation the FEL model is simplified for the analytical discussion. The core process of the FEL amplification is driven by the correlation in the longitudinal phase space of the electron beam and the growth of the radiation field amplitude along the undulator axis. Any additional aspect modifies only the physics slightly while the main working principle remains unchanged. This includes a finite bunch length and spot size, transverse motion of the electrons, longitudinal fluctuation in the electron positions, diffraction of the radiation field and the slippage of the radiation field in the forward direction relative to the electron beam.
At a certain state of designing or operating an FEL or analyzing the properties of the radiation field, a more detail study can only be done by numerical simulation. All existing codes such as TDA3D [22] and GINGER [23] are somehow limited by their underlying model of the FEL. While TDA3D excludes any longitudinal variance in the electron beam or radiation field, GINGER assumes an axi-symmetric electron beam and radiation field. Undulator field errors or wake fields are treated by these codes – if at all – only in approximation.
Most of the time for the work of this thesis was devoted to develop a tool to study these kind of problems. The name of the simulation code is GENESIS 1.3, which is a reference to the bible. the code solves the differential equation of FEL process in a complete three dimensional Cartesian coordinate system. The model used focuses on the high gain FEL, where the only major assumption, namely an ultrarelativistic electron beam, is well justified.
This thesis is structured in the following way:
Chapter 2 discusses the analytical frame for the description of the Free-Electron Laser. A self-consistent set of differential equations is derived as the theoretical base of the simulation code. In the remaining sections of this chapter these equations are discussed by first reducing the problem to the fundamental FEL process in the longitudinal phase space and then including a transverse and longitudinal dependence such as beam profile and slippage of the radiation field. To achieve a highly efficient simulation code the numerical methods of solving coupled ordinary

or partial differential equations must be sophisticated. Therefore Chapter 3 is devoted to the analysis of numerical methods to solve all problems which are encountered at FEL simulations. The fast and efficient algorithms have been incorporated into GENESIS 1.3
In Chapter 4 the performance of GENESIS 1.3 is analyzed. Beside the standard test of stability and consistency the results are compared to the analytical model of Chapter 2 and the results of the UCLA/LANL/RRCKI/SSRL experiment [16] on a high gain SASE FEL. For standard problems such as the dependence of the radiation power on the energy spread or emittance of the electron beam, GENESIS 1.3 has been compared to other codes. The simulations are based on the APS-FEL parameters at Argonne [18] and agree well with each other [24].
The last two chapters present simulation results for two projects of SASE FELs at the Deutsches Elektronen Synchrotron (DESY) in Hamburg. For the future linear collider TESLA [20] a test facility is currently under construction at DESY. This TESLA Test Facility (TTF) [17] provides the electron beam suitable for an FEL operating in the VUV region. Two problems, which occur at these wavelengths and which are most suitable to be studied with GENESIS 1.3, are the impact of a coherent transverse motion of the electron beam and the energy modulation by wake fields. The results are discussed in Chapter 5.
Similar to the TTF linear accelerator the TESLA linear collider is proposed to include the possibility to operate an X-ray FEL. At these short wavelengths additional problems limit the performance of the Free-Electron Laser. The design and the operation has to deal with a slow diverging radiation beam, stronger effect on the amplification by large transverse emittance and the increase of the energy spread due to quantum fluctuation of the incoherently emitted radiation. The simulation of these aspects is the topic of Chapter 6.

Chapter 2

Theory of Free-Electron Lasers

This chapter gives an overview of the theoretical description of a Free-Electron Laser (FEL). A Free-Electron Laser differs from common types of lasers because the electrons are not bound to quantum states of an optical medium but to a magnetic system instead. They pass the alternating magnetic field of an undulator or wiggler thus being forced to oscillate transversely. This motion causes the emission of radiation, called spontaneous undulator radiation, sharply peaked at resonant frequencies.
The emission can be stimulated by the presence of an external radiation field. Due to this stimulated emission the FEL can be classified as a laser with the additional feature that the radiation frequency is tunable by varying the electron beam energy. This is the major advantage of the Free-Electron Laser.
Free-Electron Lasers can be grouped into three types of operation

- FEL amplifier,
- FEL oscillator,
- Self-Amplified Spontaneous Emission (SASE) FEL.

The FEL amplifier is the simplest device. An external radiation fields seeds the FEL and gets amplified by the interaction with the electron beam. The basic working principle of an FEL can be explained best by this device.
If the undulator is enclosed by mirrors and the radiation is reflected back to the entrance of the undulator the FEL operates as an FEL oscillator. The amplification is accumulated over several passes until an equilibrium state is reached at saturation. The characteristics of the FEL, such as the transverse size of the radiation field and the synchronization between radiation pulse and electron beam, are strongly influenced by the optical cavity. The FEL oscillator is started by the spontaneous emission of the electron beam.
The SASE FEL operates almost as an FEL amplifier, where the seeding field is supplied by the spontaneous emission. Because the bandwidth of the spontaneous emission spectrum is larger than the FEL amplification bandwidth the SASE FEL is always tuned to the resonant frequency with the largest growth rate. The disadvantage is that the radiation pulse consists of spikes. They are caused by the random fluctuation in the longitudinal electron positions.

In most of this chapter the case of an FEL amplifier is treated, because it exhibits the FEL physics in its simplest form. The FEL oscillator is mentioned where it seems to be useful although this thesis does not emphasize this kind of device.
To begin the motion of the electrons within the undulator field, excluding interaction with the radiation field, is derived (Section 2.1). Section 2.2 includes an external radiation field. Then in the approximation of a nearly constant amplitude of the radiation, the equations of a low gain FEL are discussed.
On the basis of the self-consistent FEL equations (Section 2.3), including Maxwell's equation for the radiation field, the 1D FEL model is discussed first (Section 2.4 and 2.5). This model is capable to analyze the fundamental characteristics of a high gain FEL.
The remaining two sections extend the 1D model to radiation field diffraction (Section 2.6) and longitudinal variation of the radiation field (Section 2.7). In particular the SASE FEL, which is the main topic of the last section, can only be treated in a time-dependent formalism.

2.1 Electron Motion in an Undulator or Wiggler

The hardware part of a Free-Electron Laser (FEL) is an undulator or wiggler. Its main purpose is to force the electrons to oscillate ('wiggle') while moving through the undulator or wiggler. This transverse motion causes the electron beam to emit synchrotron radiation. For relativistic electrons the synchrotron radiation is confined to a forward cone. The opening angle is the inverse of the Lorentz factor $\gamma = E/mc^2$, where E is the electron energy, m is the electron mass and c is the speed of light.
The main feature of an undulator and wiggler is a series of paired magnets along the main axis. They are placed opposite to each other, separated by a gap of width g. The magnetic flux has only a transverse component on the undulator axis. If the plane of the gap is fixed the undulator or wiggler is planar. Another type involves the rotation of the magnets along the main axis in the form of a double helix. This type is called helical.
A Cartesian coordinate system, where the z-axis coincides with the undulator axis, will be used throughout this thesis. The transverse coordinates x and y are chosen so that the magnetic field for a planar undulator or wiggler is parallel with the y-axis. Due to the rotational symmetry the choice of the coordinate system orientation for the helical undulator is arbitrary. So here it is defined so that the magnetic field at the undulator entrance ($z = 0$) has only field components in the y-direction.
A higher magnetic field strength can be achieved by hybrid magnets, where iron poles with high permeability are placed between permanent magnets [25]. Fig. 2.1 shows a schematic cross section of a planar undulator or wiggler based on hybrid magnets. The magnetic field of the permanent magnets points either in positive or negative z-direction. The flux of two adjoining magnets is bent into the transverse direction by the iron pole. The advantage of this method is that the cross section of the iron pole faces is smaller than the permanent magnets themselves. Therefore the maximum achievable magnetic field can be increased by compressing the magnetic flux. A magnetic field strength larger than 2 T can then be obtained.
Wigglers and undulators differ in the deflection strength of the magnetic field. If the max-

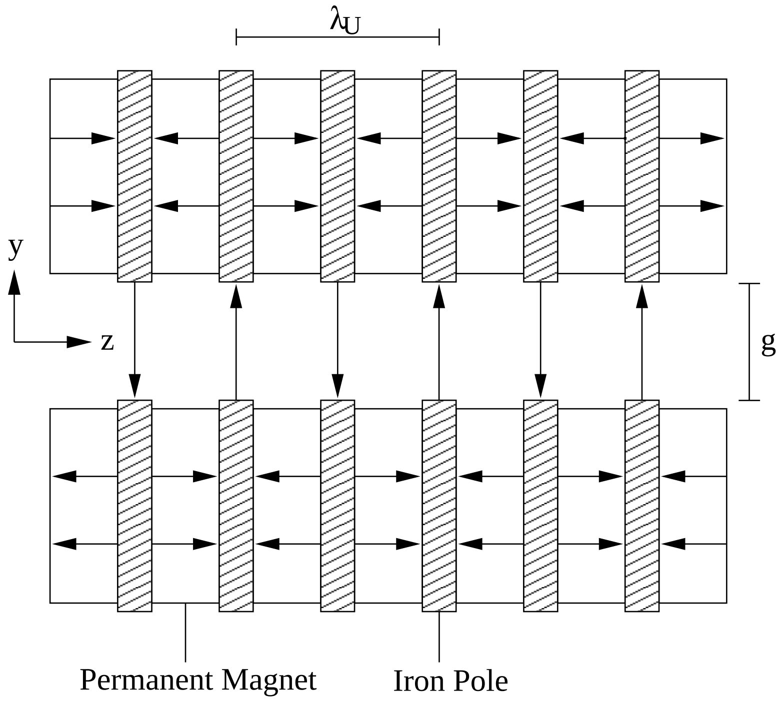


Figure 2.1: *Schematic cross section of a planar undulator with a gap width g and a periodicity λ_U. The direction of the magnetic field is indicated by arrows.*

imum deflection angle is larger than the opening angle of the spontaneous emission there is no continuous emission in the forward direction, resulting in a wiggler. The spectrum observed is enriched by higher harmonics of the periodic signal of the detected radiation. Undulator radiation is modulated but not pulsed in the forward direction and the number of higher harmonics in the spectrum is reduced. A typical spectrum for the TESLA Test Facility is shown in Fig. 2.2. A more quantitative criterion to distinguish undulators and wigglers is given later in this section. Although both undulators and wigglers are used for Free-Electron Lasers, for the sake of simplicity the remaining part of this thesis refers only to undulators unless necessity requires that the two types must be distinguished.

2.1.1 The Planar Undulator

The discussion begins with the derivation of the electron trajectories within the planar undulator. The calculation for the helical case is similar and is given in the next subsection in a more compressed form.
The magnetic field on the undulator axis is a harmonic function of the longitudinal position z:

$$B_y(z, x=0, y=0) = B_0 \cos(k_U z) \quad .$$

The field points in the y-direction and has an amplitude B_0 and wavenumber $k_U = 2\pi/\lambda_U$, respectively. Although it might be desirable, the field cannot be constant over the whole

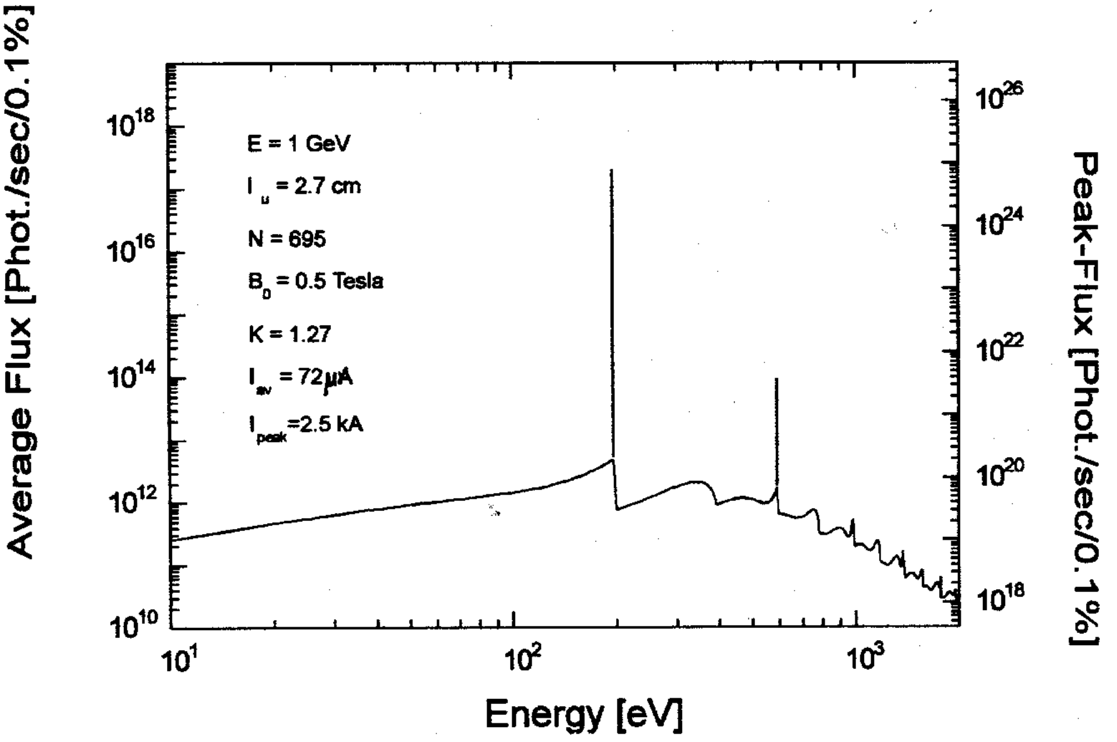


Figure 2.2: *Radiation spectrum of the Free-Electron Laser at the TESLA Test Facility.*

transverse plane. Within the free space of the undulator gap Maxwell's equations for a static magnetic field require that the divergence and curl vanish ($\vec{\nabla} \cdot \vec{B} = 0$ and $\vec{\nabla} \times \vec{B} = 0$). The second condition determines the dependence of the magnetic field on the transverse coordinates. It also allows to derive the magnetic field from a scalar potential ϕ with $\vec{B} = -\vec{\nabla}\phi$. In order to fulfill Maxwell's equations the scalar potential ϕ must be a solution of the Laplace equation $\Delta\phi = 0$.
A good starting assumption is

$$\phi = -\frac{B_0}{k_y} \cosh(k_x x) \sinh(k_y y) \cos(k_U z) \quad , \tag{2.1}$$

which gives the desired magnetic field on axis. Inserting Eq. 2.1 into the Laplace equation the scalar potential is a physically reasonable solution if the relation

$$k_x^2 + k_y^2 = k_U^2 \tag{2.2}$$

is valid [26]. In general, to a good approximation a magnetic field is perpendicular to the pole faces. This implies that the pole faces can be identified with equipotential surfaces, where the scalar potential ϕ is constant. For any arbitrarily chosen position z the curvature of the equipotential surface is defined by the relation $\cosh(k_x x) \sinh(k_y y) = const.$ It can be seen that y must be constant for $k_x = 0$ and that the pole faces are plane. The case of an outward

bent pole face is covered by an imaginary value of k_x or, which is equivalent, be replacing the cosh-function in Eq. 2.1 by the cosine function. In this case k_y^2 becomes larger than k_U^2. For real values of k_x with $k_x > 0$ the two opposite poles are bent towards each other and k_y is either reduced ($k_x < k_U$), zero ($k_x = k_U$) or imaginary ($k_x > k_U$).
For x and y small compared to the undulator period length so that $k_x x, k_y y \ll 1$, the hyperbolic function can be expanded into Taylor series up to second order. In this approximation, which is reasonable for most undulators up to a beam radius of typically 1 mm, the magnetic field becomes

$$\vec{B} = B_0 \begin{pmatrix} k_x^2 xy \cos(k_U z) \\ \left(1 + \frac{k_x^2 x^2}{2} + \frac{k_y^2 y^2}{2}\right) \cos(k_U z) \\ -k_U y \sin(k_U z) \end{pmatrix} \quad . \tag{2.3}$$

The extra field caused by curved pole faces is equivalent to a sextupole field with the amplitude $B_0 k_x^2$. As shown later in this section it provides focusing of the electron beam in the x-direction.
For further discussion it is useful to know the vector potential $\vec{A}$ of the undulator field. This is given by

$$\vec{A} = \frac{B_0}{k_U} \begin{pmatrix} \left(1 + \frac{k_x^2}{2} x^2 + \frac{k_y^2}{2} y^2\right) \sin(k_U z) \\ -k_x^2 xy \sin(k_U z) \\ 0 \end{pmatrix} \tag{2.4}$$

with $\vec{B} = \vec{\nabla} \times \vec{A}$.
The equations of motion for the position $\vec{r}$ and canonical momentum $\vec{P}$ of a single electron [27] are obtained from the Hamilton formalism, using the Hamilton function of a relativistic electron

$$H = \sqrt{(\vec{P} - e\vec{A})^2 c^2 + m^2 c^4} + e\Phi \quad , \tag{2.5}$$

where Φ is the scalar potential of the electric field $\vec{E}$ with $\vec{E} = -\vec{\nabla}\Phi - \partial\vec{A}/\partial t$.
If the electron is relativistic with $\gamma \gg 1$ the motion of the electron is mainly defined by the magnetic field of the undulator. Interaction with a radiation or electrostatic field can be regarded as a perturbation. These effects, which are important for the FEL process, are discussed in later sections.
With this assumption the Hamilton function is a constant of motion because it does not depend explicitly on the time t. Due to the absence of an electric field ($\Phi = 0$) the electron energy γmc^2 is constant as well and identical in value to the Hamilton function.
It is difficult to solve the equations of motion directly. Therefore the electron motion is split into two parts,

$$\vec{r}(t) = \vec{r}_0(t) + \vec{R}(t) \quad ,$$

separating the main oscillation $\vec{r}_0(t)$ due to the periodic undulator field from a drift $\vec{R}(t)$ in the transverse position. The drift is slow compared to the quickly varying term $\vec{r}_0(t)$ and has

a characteristic length on the scale of many undulator periods. As a first step the solution for $\vec{r}_0$ is obtained by assuming that $\vec{R}(t)$ is constant.
The equations of the motion for the transverse canonical momentum $\vec{P}$ are

$$\dot{P}_x = -\frac{\partial}{\partial x}H = \frac{e}{\gamma m}\left(\frac{\partial}{\partial x}\vec{A}\right)\cdot\left(\vec{P}-e\vec{A}\right) \quad , \tag{2.6}$$

$$\dot{P}_y = -\frac{\partial}{\partial y}H = \frac{e}{\gamma m}\left(\frac{\partial}{\partial y}\vec{A}\right)\cdot\left(\vec{P}-e\vec{A}\right) \quad . \tag{2.7}$$

For the vector potential Eq. 2.4 the lowest order term of the time derivative $\dot{\vec{P}}$ is linear in $k_x x$ or $k_y y$, respectively. As mentioned at the expansion of the hyperbolic function in Eq. 2.3 these linear terms are small compared to unity. Thus the change of the canonical momentum contributes either to the 'slow' motion $\vec{R}(t)$ or to the higher orders solutions of $\vec{r}_0$, which are not regarded in this discussion.
The remaining equations of the transverse motion

$$\dot{x} = \frac{\partial}{\partial P_x}H = \frac{P_x - eA_x}{\gamma m} \quad , \tag{2.8}$$

$$\dot{y} = \frac{\partial}{\partial P_y}H = \frac{P_y - eA_y}{\gamma m} \tag{2.9}$$

have only one dominant and quickly oscillating source term, given by the x-component of the vector potential in Eq. 2.8. The resulting motion takes place in the xz-plane with the 'fast' velocity

$$\dot{x}_0 = -\frac{\sqrt{2}cK}{\gamma}\sin(k_U z). \tag{2.10}$$

Eq. 2.10 suggests the definition of the dimensionless undulator field

$$K = \frac{e\hat{B}}{mck_U}\left(1+\frac{k_x^2}{2}X^2+\frac{k_y^2}{2}Y^2\right) \tag{2.11}$$

depending to second order on the transverse position $X = X(t)$ and $Y = Y(t)$ of the 'slow' trajectory $\vec{R}(t)$. This definition differs from that in other publications where the on-axis peak field B_0 is used instead of the root-mean-square value $\hat{B}$. In the case of a planar undulator $\hat{B}$ is $B_0/\sqrt{2}$. The advantage of this definition is that many equations remain the same for the case of the helical undulator. The value of K at the undulator axis $(X, Y = 0)$ defines the undulator parameter. Because the second order corrections to the undulator field are of the order of 10^{-3} the transverse dependence of the undulator field has a negligible impact on most of the calculations. Therefore it is sufficient to use the constant value of the undulator parameter instead. One exception can be found in Chapter 6.

Eq. 2.10 exhibits the distinction between wiggler and undulator. If the electron is relativistic ($z \approx ct$) the maximum divergence $x' = \dot{x}_0/c$ of the electron is $\sqrt{2}K/\gamma$. The opening angle of the synchrotron radiation is γ^{-1} and thus the device is an undulator for $K \leq 1/\sqrt{2}$ and a wiggler otherwise.
There is no dominant component of the vector potential in y and the motion in this direction consists only of the 'slow' motion ($y_0(t) = 0$).
Due to energy conservation the longitudinal velocity can directly be obtained from the definition of the Lorentz factor γ and the normalized velocity $\vec{\beta} = d\vec{r}/cdt$. Then the longitudinal velocity is

$$\begin{aligned} \beta_z &= \sqrt{1 - \frac{1}{\gamma^2} - \beta_x^2 - \beta_y^2} \\ &\approx 1 - \frac{1+K^2}{2\gamma^2} - \frac{\beta_R^2}{2} + \frac{K^2}{2\gamma^2}\cos(2k_U z) \quad , \end{aligned} \tag{2.12}$$

where β_R is the transverse velocity of the slow drift, normalized to c. The cross term proportional to $\beta_R K/\gamma \cdot \sin(k_U z)$ has been neglected because it is either small compared to the leading oscillating term ($\propto K^2 \cos(2k_U z)$) or not resonant with variation of β_z as is the case for $\beta_R^2/2$. The transverse motion within the undulator slows down the electron by roughly $\Delta\beta_z = K^2/2\gamma^2$ with a superimposed longitudinal oscillation with a period half as long as the transverse oscillation.
To obtain the trajectory $x_0(t)$, the longitudinal position is approximated by $z = c\beta_z t \approx c\beta_0 t$ and then Eq. 2.10 is integrated in first order, using the averaged velocity

$$\beta_0 = 1 - \frac{1+K^2}{2\gamma^2} \quad . \tag{2.13}$$

The integration yields

$$x_0(t) = \frac{\sqrt{2}K}{\gamma k_U \beta_0}\cos(ck_U\beta_z t) \quad . \tag{2.14}$$

The longitudinal oscillating term in Eq. 2.12 is the source of a phase modulation in the cosine function in Eq. 2.14. As a consequence the transverse oscillation exhibits higher harmonics of the fundamental wavenumber k_U. In addition the synchronization of the electron position with a phase front of an electromagnetic wave, propagating along the undulator axis, is reduced. The impact of both facts will be discussed in the next section.
Only slowly varying terms in the equations of motion can contribute to $\vec{R}$. By averaging over one undulator period Eqs. 2.8 and 2.9 are reduced to $\dot{X} = P_x/\gamma m$ and $\dot{Y} = P_y/\gamma m$. The vector potential $\vec{A}$ has only terms proportional to $\sin(k_U z)$ or $\sin(2k_U z)$ and they vanish after averaging.
In the remaining equations Eqs. 2.6 and 2.7 all terms are zero except for $(\partial A_x/\partial x)A_x$ and $(\partial A_x/\partial y)A_x$, respectively.
The resulting differential equations

$$\dot{P}_x = -\gamma m c^2 \frac{K^2 k_x^2}{\gamma^2} X \quad , \tag{2.15}$$

$$\dot{P}_y = -\gamma m c^2 \frac{K^2 k_y^2}{\gamma^2} Y \tag{2.16}$$

describe reaction forces proportional to the displacement.

The magnetic field of the undulator provide a natural focusing of the electrons if the pole faces are flat or bent towards each other ($k_x^2 \leq 0$). Although the focusing strength in both planes depends on the curvature of the magnetic poles the combined strength $K^2 k_U^2/\gamma^2$ does not due to Eq. 2.2. For flat horizontal pole faces there is no focusing in the x-plane. Increasing the focusing strength in this plane involves a reduction in the y-plane. A more precise calculation shows that the finite width of the undulator magnets introduces a small change in the magnetic field so that a slight defocusing term is noticeable in the x-direction for $k_x = 0$ [29].

The trajectories of the transverse slow motion are harmonic functions with frequencies $\Omega_x = K k_x/\gamma$ and $\Omega_y = K k_y/\gamma$, for the x- and y-plane, respectively. The period length λ_β is typically of the order $\lambda_\beta \approx (\gamma/K)\lambda_U$ and thus much larger than the undulator period for a highly relativistic electron ($\gamma \gg 1$). The index β refers to the definitions, used in the accelerator physics, where this oscillation is called a betatron oscillation [28].

The transverse beam size is strongly related to the focusing strength. The calculations for the y-direction are identical to those in the x-direction, which are presented here. The general betatron oscillation of a single electron is given by $X(t) = X_0 \cos(\Omega_x z) + (X_0'/\Omega_x)\sin(\Omega_x z)$, where X_0 is the initial offset of the electron and X_0' is the initial angle relative to the undulator axis.

The emittance

$$\epsilon_x = \sqrt{\overline{(x-\overline{x})^2}\;\overline{(x'-\overline{x'})^2} - \overline{(x-\overline{x})(x'-\overline{x'})}^2} \quad , \tag{2.17}$$

where the bar of a parameter denotes an average over all electrons, is a constant of motion in linear optics [30]. Regarding this definition of the emittance, $\pi\epsilon_x$ can be identified as an equivalent volume of the electron distribution in the transverse (x, x') phase space.

In contrast to the emittance, the root-mean-square envelope of the electron beam is usually not a constant of motion [28]. The general expression of the envelope $\sigma_x(z)$ for $k_x^2 > 0$ within an undulator is

$$\sigma_x(z) = \sqrt{\sigma_x(0)^2 \cos^2(\Omega_x z) + \frac{\sigma_x(0)\sigma_x'(0)}{2\Omega_x}\sin(2\Omega_x z) + \frac{\epsilon_x^2 - \sigma_x^2(0)\sigma_x'^2(0)}{\sigma_x^2(0)\Omega_x^2}\sin^2(\Omega_x z)} \quad , \tag{2.18}$$

where $\sigma_x(0)$ and $\sigma_x'(0)$ are the initial beam size and its derivative in z, respectively. For a matched beam, when the beam size remains constant over the full undulator length, the electron beam must go through a waist directly at the entrance of the undulator ($\sigma_x'(0) = 0$) with an rms size of $\sigma_x(0) = \sqrt{\epsilon_x/\Omega_x}$. If the undulator focuses equally in both planes with $k_x = k_U/\sqrt{2}$

the constant size is $\sigma_x(0) = \sqrt{(\sqrt{2}mc/e)\epsilon_x\gamma/\hat{B}}$. All other initial settings cause a modulation of the envelope. If a smaller beam size is desired it can be achieved by superimposing a lattice of quadrupoles. Normally this is referred to as strong focusing in contrast of the natural or weak focusing given by the undulator field itself.

2.1.2 The Helical Undulator

The treatment of the helical undulator is very similar to that of the planar one. Indeed most of the results are the same. The magnetic field $\vec{B}$ as well as the vector potential $\vec{A}$ consists of a linear combination of the first order modified Bessel functions I_0 and I_1 [31] depending just on $k_U r$, where r is the transverse distance between the electron position and the undulator axis. Using the assumption that $k_U r$ is much smaller than unity the Bessel functions are expanded into Taylor series. Up to second order in $k_U r$ the vector potential in the Cartesian coordinate system is given by

$$\vec{A} = \frac{B_0}{k_U} \begin{pmatrix} \left[1 + \frac{k_U^2}{8}(3y^2 + x^2)\right] \sin(k_U z) - \frac{k_U^2}{4} xy \cos(k_U z) \\ \left[1 + \frac{k_U^2}{8}(3x^2 + y^2)\right] \cos(k_U z) - \frac{k_U^2}{4} xy \sin(k_U z) \\ 0 \end{pmatrix} . \tag{2.19}$$

The magnetic field is derived in the usual way by evaluating $\vec{B} = \vec{\nabla} \times \vec{A}$.
The trajectory is split into a quickly oscillating term $\vec{r}_0(t)$ and the slow betatron oscillation $\vec{R}(t)$. The velocity of the fast motion $\dot{\vec{r}}_0$ is proportional to the vector potential. Close to the undulator axis $\dot{x}_0$ and $\dot{y}_0$ have the same amplitude of oscillation but they have a phase difference of $\pi/2$. This is an obvious result which can be expected due to the symmetry of a helical undulator. Off-axis the velocity differs in both direction due to the higher order terms in x and y of the vector potential. The transverse motion of the electron becomes more elliptical with the short axis pointing in the radial direction. To eliminate this azimuthal dependence and for a better comparison with the results for the planar undulator, the vector potential is averaged in the azimuthal direction.
The normalized longitudinal velocity is

$$\beta_z \approx 1 - \frac{1 + K^2}{2\gamma^2} - \frac{\beta_R^2}{2} \tag{2.20}$$

with the undulator field

$$K = \frac{e\hat{B}}{mck_U}\left(1 + \frac{k_U^2}{4}(X^2 + Y^2)\right) \tag{2.21}$$

in the Taylor series expansion up to second order in X and Y.
The major difference between helical and planar undulators becomes apparent here. Because the electron oscillates in both transverse directions but with $\pi/2$ phase difference, the longitudinal velocity is almost constant. The terms proportional to $\beta_R \cos(k_U z)$ or $\beta_R \sin(k_U z)$ are negligible and not included in Eq. 2.20. The absence of a longitudinal oscillation excludes the generation

of higher harmonics in the transverse motion of the electron. The helical undulator field Eq. 2.21 agrees with that of a planar undulator (Eq. 2.11) if the planar undulator provides equal focusing in both planes with $k_x^2 = k_y^2 = k_u^2/2$. This similarity is an advantage of the undulator field definition based on the root-mean-square value $\hat{B}$.
With the definition of β_0 in analogy with Eq. 2.13 the transverse velocity can be integrated to obtain the trajectory $\vec{r}_0$. The electrons moves along a helix with a pitch of λ_U. Due to the asymmetry in the azimuthal and radial motion for larger transverse offsets the helix is slightly distorted [32]. The average radius of the motion is independent of the azimuthal angle with

$$r_0 = \frac{K}{\gamma k_U \beta_0} \quad . \tag{2.22}$$

By averaging the transverse equations of motion over the length of one period the fast oscillation drops out. Some basic algebra yields the differential equations

$$\dot{X} = \frac{P_x}{\gamma m} - c\frac{\Omega_U^2}{k_U} Y \quad , \tag{2.23}$$

$$\dot{P}_x = -\gamma m c^2 \Omega_U^2 X - c\frac{\Omega_U^2}{k_U} P_y \quad , \tag{2.24}$$

$$\dot{Y} = \frac{P_y}{\gamma m} - c\frac{\Omega_U^2}{k_U} X \quad , \tag{2.25}$$

$$\dot{P}_y = -\gamma m c^2 \Omega_U^2 Y - c\frac{\Omega_U^2}{k_U} P_x \tag{2.26}$$

with $\Omega_U = K k_u / \sqrt{2}\gamma$.
These differential equations describe two coupled oscillations but they can be decoupled into ordinary differential equations for harmonic oscillations with the frequencies

$$\hat{\Omega} = \Omega_U \left(\sqrt{1 + \frac{\Omega_U^2}{k_U^2}} \pm \frac{\Omega_U}{k_U} \right) \tag{2.27}$$

by transforming to the variables $X \pm iY$. The ratio Ω_U / k_U is of the order $1/\gamma$. The characteristic length of the orbit beat by the coupling is roughly $\gamma^2 \lambda_U$ and even for moderately relativistic electrons much longer than the undulator length itself. This term is only important for storage ring based undulators because it is a major source of coupling of the betatron motion [32]. By neglecting the coupling term, Eqs. 2.23 – 2.26 become identical with the corresponding equations for the planar undulator with $k_x^2 = k_y^2 = k_U^2/2$. The conditions for optimum matching of the electron beam are valid for the helical undulator as well.

2.2 The Interaction of Electrons with a Radiation Field in an Undulator

In this section the interaction of electrons with a radiation field while they move through the undulator is analyzed. The approach to this problem is similar to that in the previous section except that an additional term in the Hamilton function describes the vector potential of the radiation field. If the emission of radiation is stronger than the absorption the electrons are losing energy in average and the radiation field is amplified. As long as this amplification is small the radiation field amplitude can be assumed to be constant in the Hamilton function for deriving the equations of motion. The limitations of this model of a 'low gain' Free-Electron Laser are given at the end of this section. A more self-consistent model of an FEL can be found in the next section, including Maxwell's equation for the radiation field description. Nevertheless a discussion of the low gain FEL is fruitful, because it shows the basic principle of how an FEL works with rather simple equations.
The interaction of charged particles with a radiation field shows two major aspects. The first is the change of the particle momentum and energy. The Hamilton equations of motion are the mathematical representation of this process. The method for solving these equations is very similar to the treatment in the previous section, but differs in the point that the electron energy is not constant anymore due to the electric field components of the radiation field.
The second aspect is the change of the radiation field itself. The fast transverse oscillation of the electrons is a source of radiation. For relativistic particles this radiation points mainly in the forward direction of the electron beam motion. If the radiation wavelength is shorter than the electron bunch length the electrons emit at almost all phases and the radiation adds up incoherently. The emission is strongly enhanced if the longitudinal beam profile is modulated on the scale of the radiation wavelength.
Under special conditions both processes, the change of the particle energy and the emission of radiation, are the source of a collective bunching of the electrons on a resonant frequency and the radiation field is strongly amplified. The next section analyzes this instability — the working principle of the 'high gain' FEL. In contrast to the high gain FEL the low gain FEL provides an amplification without the necessity of a strong modulation in the electron density.
The discussion begins with the assumption on the radiation field. If a radiation field propagates along the undulator together with the electron bunch the interaction time is maximized. The electric field components are lying in the transverse xy-plane, thus only a transverse motion, along or against the field orientation, changes the electron energy. Due to the symmetry of the magnetic field the radiation emitted in a planar undulator is linear polarized while it is circular polarized for the case of a helical undulator. In this section the case of a planar undulator is regarded. Most of the results are similar or identical for a helical undulator and only important differences are mentioned in the text.
The electric field component of the radiation field

$$\vec{E} = \vec{E}_0 \cos(k(z - ct) + \Psi) \quad , \tag{2.28}$$

is defined by its amplitude $\vec{E}_0$, its wavenumber $k = 2\pi/\lambda$ or wavelength λ, and its initial phase

Ψ at the undulator entrance.
The magnetic field component is perpendicular to $\vec{E}$ as well as to the unit vector in the direction of propagation, which mainly coincides with $\vec{e}_z$. Compared to the strong undulator field the magnetic field of the radiation field is negligible and can be ignored in the further discussion. The amplitude $\vec{E}_0$ and the phase Ψ depend on z due to diffraction. The dependence becomes negligible small if the transverse extension of the radiation wavefront is much larger than the radiation wavelength. Appendix A gives more details on propagation of electromagnetic waves in free space.
The change of the electron energy is caused only by the electric field components, which, depending on the radiation phase, accelerate or decelerate the electron with

$$\dot{\gamma} = \frac{e\vec{E}\cdot\vec{\beta}}{mc} \quad . \tag{2.29}$$

Only the parallel components of $\vec{E}$ and $\vec{\beta}$ contribute to Eq. 2.29. In the case of the planar undulator they are pointing in the x-direction resulting in a linear polarization of the radiation field.
To obtain the transverse velocities $\vec{\beta}$ the vector potential $\vec{A}_r$ of the electro-magnetic wave has to be added to the Hamiltonian Eq. 2.5. From the potential

$$\vec{A}_r = \frac{1}{ck}\sin(k(z-ct)+\Psi)\begin{pmatrix} E_0 \\ 0 \\ 0 \end{pmatrix} \quad . \tag{2.30}$$

the electric field is derived by the time derivative $\vec{E} = -\partial\vec{A}_r/\partial t$. In this section the Lorentz gauge is chosen, which allows to omit the scalar potential in the derivation of the electric field. For an assumed pulse length $L \gg \lambda$ the dependence of the amplitude $\vec{E}_0$ as well as the phase Ψ on the time is negligible and A_r is a valid vector potential for the radiation field Eq. 2.28. Inserting the vector potential of the radiation field and the undulator field into the Hamilton function, the transverse velocities are

$$\dot{x} = -\frac{\sqrt{2}cK}{\gamma}\sin(k_U z) - \frac{\sqrt{2}cK_r}{\gamma}\sin(k(z-ct)+\Psi) + \dot{X}, \tag{2.31}$$

$$\dot{y} = \dot{Y}. \tag{2.32}$$

The dimensionless radiation amplitude

$$K_r = \frac{e\hat{E}}{mc^2k} \tag{2.33}$$

is defined in an analogous way as the undulator parameter K. The motivation to use the root-mean-square value $\hat{E}$ of the electric field is the same. Most results will be identical for the helical undulator. The velocity terms $\dot{X}$ and $\dot{Y}$ of the betatron oscillation are the same as in the last sections.

For sake of simplicity any transverse variation of the radiation field is excluded. A radiation field with a finite transverse extension is more difficult to analyzed (3D FEL model in Section 2.6 and free space propagation in Appendix A).
For small transverse momenta the longitudinal velocity is approximately

$$\begin{aligned} \beta_z \approx\ & 1 - \frac{1+K^2+K_r^2}{2\gamma^2} - \frac{\beta_R^2}{2} \\ & + \frac{K^2}{2\gamma^2}\cos(2k_U z) + \frac{K_r^2}{2\gamma^2}\cos(2k(z-ct)+2\Psi) \quad (2.34) \\ & - \frac{2KK_r}{\gamma^2}\sin(k_U z)\sin(k(z-ct)+\Psi) \quad . \quad (2.35) \end{aligned}$$

This expression is very similar to Eq. 2.12 except for three additional terms. The electric field forces an additional transverse oscillation with the frequency of the electromagnetic wave. As for the undulator field the longitudinal velocity is slowed down and modulated with an oscillation of twice the frequency of the radiation field. It will be shown later in this chapter that the longitudinal modulation by the radiation field is much smaller than that by the undulator field and can be neglected.
The cross term $\propto KK_r$ can be split into two independent oscillations. If one of them has a small frequency it can significantly change the longitudinal velocity β_z on a time scale different to the dominant oscillating term $\propto K^2$. The explicit calculation of this term is postponed till β_z is further discussed (Eq. 2.42).
Combining all constant or slow varying terms to β_0, the integration of Eq. 2.35 up to first order yields

$$z = \beta_0 ct + \frac{K^2}{4\gamma^2 k_U \beta_0}\sin(2k_U\beta_0 ct) \quad . \quad (2.36)$$

With the given expression of the transverse velocities $\dot{x}$ and $\dot{y}$, Eq. 2.29 can be evaluated. Most of the cross terms between E_x and β_x are fast oscillating. Over many undulator periods the net change of the electron energy is negligible. The only possible term that might be constant is the product of $\cos(k(z-ct)+\Psi)$ and $\sin(k_U z)$, similar to the term in Eq. 2.35. This term is split into two independent oscillations with the phases $(k \pm k_U)z - kct + \Psi$. If one of the phases remains almost constant the energy change is accumulated over many periods.
With an average longitudinal velocity of $c\beta_0$ the phase relation between electron and radiation field remains unchanged if the condition

$$\beta_0 = \frac{k}{k \pm k_U} \quad (2.37)$$

is fulfilled. As shown later in this chapter the interaction between the electron beam and the radiation field needs to add up resonantly over many undulator periods to result in a significant change of the electron energy or radiation amplitude and phase. This implies that for a given beam energy and undulator wavelength the radiation wavelength of the radiation field is well defined according to Eq. 2.37. The case of the '–' sign is excluded because it would demand an

electron velocity faster than the speed of light to keep the electrons in phase with the radiation field for any time. The restriction to a well defined resonant radiation wavelength is called resonance approximation. A quantitative expression for the limits of this approximation will be given in Section 2.3, when the characteristic length of this resonant interaction has been calculated.
In the limit of a weak electric field ($K_r \rightarrow 0$) and a small beam emittance β_0 is identical with Eq. 2.13. The resonant radiation wavelength is

$$\lambda_0 = \frac{\lambda_U}{2\gamma^2}(1 + K^2) \quad . \tag{2.38}$$

This important equation is valid for a planar and a helical undulator as well. A transverse betatron motion and a stronger radiation field shift slightly the resonance condition towards longer wavelength. If Eq. 2.38 is exactly fulfilled the energy change is constant over many undulator periods pushing the electron off-resonance.
So far the longitudinal oscillation of the electron has not been taken into account. As mentioned in the previous section it induces higher harmonics in the motion of the electrons.
Inserting Eqs. 2.28 and 2.31 into Eq. 2.29 yields the resonant term

$$\dot{\gamma} = -\frac{2ckKK_r}{\gamma}\cos(k(z - ct) + \Psi)\sin(k_U z) \quad . \tag{2.39}$$

Note that the choice of the radiation wave number k is free and does not need to agree with the resonant wavenumber $k_0 = 2\pi/\lambda_0$, defined by the undulator properties and the particle energy. To evaluate Eq. 2.39 the sine and cosine function are replaced by complex exponential functions. The oscillating part of the longitudinal motion (Eq. 2.36) can be expanded into a series of Bessel functions [33] by the identity

$$e^{ia\sin b} = \sum_{m=-\infty}^{\infty} e^{imb} J_m(a) \quad .$$

The result is a sum of exponential functions with the frequencies $[(k + (2m+1)k_U)\beta_0 - k]c$. Beside the ground mode with $m = 0$ some terms are resonant at different wavelengths. The frequencies of these are the odd harmonics of the resonant frequency $\omega_0 = ck_0$.
Collecting all terms belonging to one mode Eq. 2.39 becomes

$$\begin{aligned} \dot{\gamma} \;\; &= \;\; -\frac{2ckKK_r}{\gamma}\frac{1}{4i}\left[e^{i\theta+i\Psi}\sum_{m=-\infty}^{\infty} e^{i2mk_U\beta_0 ct}(J_m(\chi) - J_{m+1}(\chi))\right. \\ &\quad\;\; \left. - e^{-i\theta-i\Psi}\sum_{m=-\infty}^{\infty} e^{-i2mk_U\beta_0 ct}(J_m(\chi) - J_{m+1}(\chi))\right] \end{aligned} \tag{2.40}$$

with $\chi = kK^2/4\gamma^2 k_U$ and the so-called ponderomotive phase

$$\theta = (k + k_U)z - ckt \quad . \tag{2.41}$$

For completeness it is noted that a transverse non-uniform radiation field couples the particle motion also to the even harmonics of ω_0 [34, 35]. If the radiation field is expanded into a Taylor series around the electron position of the betatron oscillation ($x = X + x_0$)

$$\vec{E}(x) = \vec{E}(X) + \left.\frac{d\vec{E}}{dx}\right|_X x_0 \quad ,$$

the factor $x_0\dot{x_0}$ is proportional to $\sin(2k_U z)$ in Eq. 2.29. Carrying out the same calculation as for Eq. 2.40 the complex exponential functions have the arguments $[(k+(2m+2)k_U)\beta_0 - k]ct$, being resonant at all even harmonics. The additional pre-exponential factor is $(K/2K_r\gamma k_U\beta_0)dK_r/dx$. The postponed calculation of the cross term $\sin(k(z-ct)+\Psi)\sin(k_U z)$ in Eq. 2.35 is performed in a very similar way. If the phase Ψ is temporarily replaced by $\tilde{\Psi} = \Psi - \pi/2$ to convert the sine function into a cosine function, the expansion into Bessel functions yields

$$\begin{aligned} \beta_z &= 1 - \frac{1+K^2+K_r^2}{2\gamma^2} - \frac{\beta_R^2}{2} \\ &\quad + \frac{KK_r}{2\gamma^2}\left[e^{i\theta+i\Psi}\sum_{m=-\infty}^{\infty} e^{i2mk_U\beta_0 ct}(J_m(\chi) - J_{m+1}(\chi))\right. \\ &\quad \left. + e^{-i\theta-i\Psi}\sum_{m=-\infty}^{\infty} e^{-i2mk_U\beta_0 ct}(J_m(\chi) - J_{m+1}(\chi))\right], \quad . \end{aligned} \tag{2.42}$$

The resonant frequencies are well separated such that only one resonance frequency is of importance for a given radiation field. The coupling factor is smaller for higher modes. Thus the interaction is the strongest for the fundamental mode [36], which is the only mode considered in the following discussion.

In the case that the FEL operates at the fundamental frequency, the non-linear terms in the FEL equations will induce an enhanced bunching in the longitudinal position at higher harmonics. This bunching grows faster than operating on the higher frequency itself.

For the case of a helical undulator the amplification of higher modes are much smaller because the dominant longitudinal oscillation, which is the reason for the coupling to higher harmonics, is strongly suppressed. At the fundamental frequency the synchronization of the phase front of the ponderomotive wave and the electrons is almost perfect, while it is reduced by the factor $(J_0(\chi) - J_1(\chi))$ for the planar undulator.

Compared to the fast changing position of the electron $z \approx \beta_0 ct$ the ponderomotive phase $\theta = (k+k_U)z - ckt$ of the electron is almost constant. It is convenient to change to a moving coordinate system, which is synchronized with the ponderomotive wave. With a simple canonical transformation [37], which keeps the energy unchanged, the equation of motion for the new variable θ becomes $\dot{\theta} = (k+k_U)c\beta_z - kc$. Replacing β_z by Eq. 2.42, the differential equations for the 'low gain' Free-Electron Laser are obtained:

$$\dot{\theta} = ck_U - \omega\frac{1+K^2+K_r^2 - 2f_c KK_r\cos(\theta+\Psi)}{2\gamma^2} - \omega\frac{\beta_R^2}{2} \tag{2.43}$$

and

$$\dot{\gamma} = -\omega f_c \frac{KK_r}{\gamma} \sin(\theta + \Psi) \quad . \tag{2.44}$$

With the definition of the coupling factor

$$f_c = \begin{cases} J_0(\chi) - J_1(\chi) & \text{planar undulator} \\ 1 & \text{helical undulator} \end{cases} \tag{2.45}$$

and $\chi = kK^2/4\gamma^2 k_U = K^2/2(1+K^2)$ for the fundamental resonant wavelength, the FEL equations are valid for both types of undulators.
Another way to derived the differential equations is the rigorous canonical and Legendre transformation of the Hamilton function Eq. 2.5 [38]. The new Hamilton function, depending on the canonical variable and momentum θ and γ, respectively, is

$$H = ck_U\gamma + \omega \frac{1 + \gamma^2\beta_R^2 + K^2 + K_r^2 - 2f_c KK_r \cos(\theta + \Psi)}{2\gamma} \quad . \tag{2.46}$$

The independent variable is the time t. As long as the electric field and the transverse momenta do not change significantly, they can be kept constant in the Hamiltonian. This is the basic assumption of the 'low gain' FEL. The limitation of this model will be given at the end of this section.
In the limit of a 'low gain' FEL the Hamilton function is regarded as independent of t and therefore a constant of motion. Setting the Hamiltonian to $H = 2ck_U(1+\alpha)\gamma_R$ with $\gamma_R^2 = k(1 + \gamma^2\beta_R^2 + K^2 + K_r^2)/2k_U$, the particle energy γ depends on θ as

$$\gamma = \gamma_R(1+\alpha) \pm \sqrt{\gamma_R^2\alpha(2+\alpha) + \frac{kf_c KK_r}{k_U}\cos(\theta + \Psi)} \quad . \tag{2.47}$$

The lowest boundary of α is $\alpha > -1$ to avoid unphysical negative values of the energy. Other limitations are given by the square root in Eq. 2.47. Two values of α are of particular interest for the lowest possible value of the Hamilton function and for an existing solution of γ for all phases θ, respectively.
The smallest value of α is found if the cosine function in the argument of the square root is unity. At $\theta = -\Psi$ the root becomes real for

$$\alpha_0 = -1 + \sqrt{1 - \frac{kf_c KK_r}{k_U\gamma_R^2}} \quad . \tag{2.48}$$

Inserting α_0 into Eq. 2.47 yields the corresponding energy

$$\gamma_0 = \sqrt{\gamma_R^2 - \frac{kf_c KK_r}{k_U}} \quad .$$

The position $(-\Psi, \gamma_0)$ in the longitudinal phase space is a stable fix point, where the electron remains in its position. For any small deviation the differential equation 2.43 and 2.44 can be linearized and combined to a second order differential equation of $\Delta\theta = \theta + \Psi$ with

$$\Delta\theta'' + \Omega^2 \Delta\theta = 0 \tag{2.49}$$

and $\Omega = \sqrt{2 f_c k k_U K K_r}/\gamma_0$.

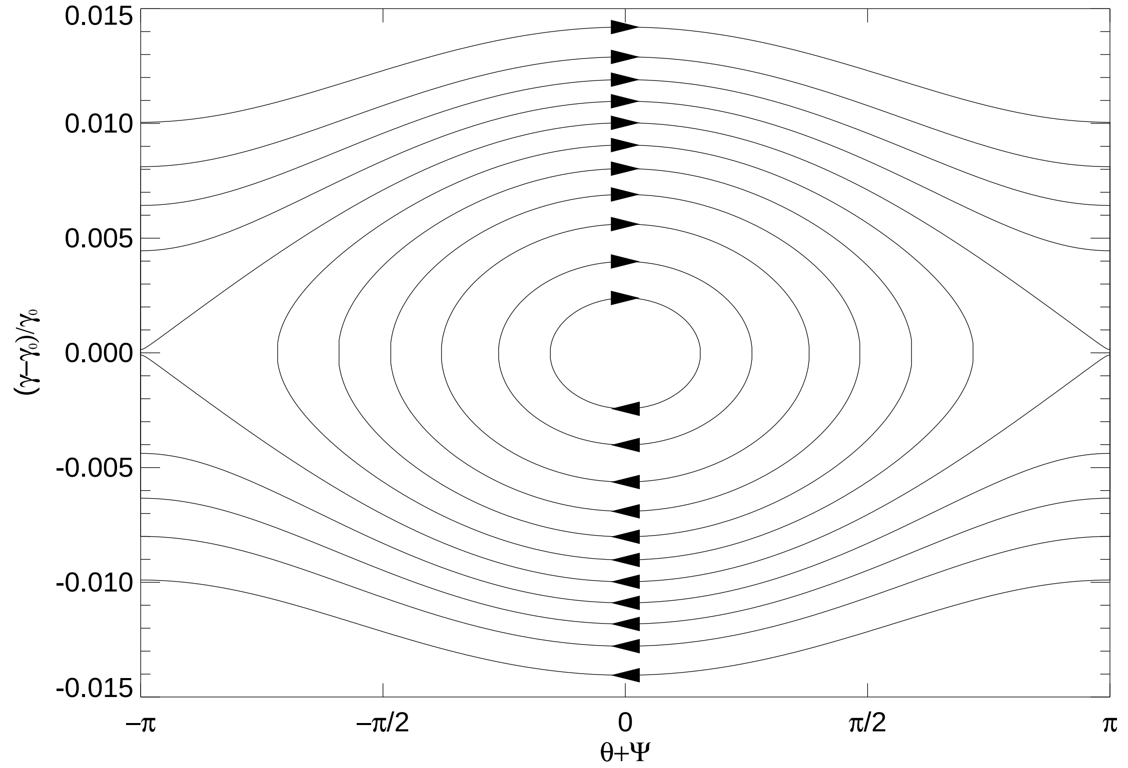


Figure 2.3: *Electron trajectories in the longitudinal phase space for different initial settings.*

This equation is solved by any sine or cosine function with the frequency Ω. The motion in the longitudinal phase space is bound. This is typical for a stable fix point. For a larger amplitude of $\Delta\theta$ non-linear terms are not negligible any longer and the frequency depends on the initial condition of the electron.

Solutions of γ for all phases θ are found for α larger than

$$\alpha_1 = -1 + \sqrt{1 + \frac{k f_c K K_r}{k_U \gamma_R^2}} \quad . \tag{2.50}$$

The trajectory in phase space is not closed and the electrons have either energy above or below γ_R. A transition is not possible.

The phase space surface for $H = 2ck_U(1+\alpha_1)\gamma_R$ is called separatrix. It separates the bound and unbound motion. Any electron within the separatrix is trapped in the ponderomotive wave and oscillates around $-\Psi$. Referring to acceleration of charged particle in rf-cavities this enclosed area of the separatrix is often called 'bucket' [28]. The width of the bucket is given by the properties of the undulator and the radiation field and it is $\Delta\gamma = \sqrt{8kf_cKK_r/k_U}$. Electrons outside the separatrix are moving unlimited in θ either faster than the ponderomotive wave or slower.

Fig. 2.3 shows several phase space trajectories for different initial conditions calculated by Eq. 2.47. Within the bucket the electrons are moving clockwise, above toward larger phases ($\dot{\theta} > 0$) or below towards smaller phases. This implies that an electron, injected at the ponderomotive phase $0 < \theta + \Psi < \pi$ loses energy. If the undulator length is shorter than the period length of the phase space oscillation $2\pi/\Omega$ the electron will mainly remain in this phase region. Due to energy conservation the radiation field has been amplified. This can be generalized for the whole electron bunch. As long as the initial distribution in the longitudinal phase space changes to a final distribution of a mean energy smaller than the initial energy, the gain of the FEL is positive.

Unfortunately the most obvious way by injecting all electrons at $0 < \theta + \Psi < \pi$ is not realizable. The radiation wavelength depends on the energy as γ^{-2} (Eq. 2.38) and is much smaller than a typical bunch length of about 1 mm. The initial ponderomotive phases of the electrons are almost uniformly distributed over 2π. Due to the finite number of electrons over one radiation wavelength a small modulation of the electron beam remains. This spontaneous emission provides the initial radiation field for Self-Amplified Spontaneous Emission Free-Electron Laser (SASE FEL), discussed in the last section of this chapter.

With an rf-photo gun driving the injector for an FEL, relative energy spreads smaller than 1% can be achieved. This width is typically smaller than the width of the bucket and fills it unevenly. For a large energy spread the bucket is filled almost homogeneously. Any motion of the electrons within the homogeneously filled bucket would not change the mean energy, because the phase space density remains constant according to Liouville's theorem [27].

Operating as an FEL amplifier the injection at resonance energy γ_R would not provide any gain at all. For the unmodulated beam the energy change of one electron is always compensated by a complementary electron, which moves on the same trajectory but which has a phase difference of $2(\theta + \Psi)$. The only visible effect is the increase of the energy spread, because electrons at $-\pi < \theta + \Psi < 0$ gain energy while the complementary electrons at $0 < \theta + \Psi < \pi$ lose energy. If the injection is off-resonance ($\gamma \neq \gamma_R$) the change of the phase space distribution is not symmetric anymore. For $\gamma > \gamma_R$ electrons at $-\pi < \theta + \Psi < 0$ tend to change rather the phase than the energy while it is opposite for the remaining electrons. Averaging over all electrons the electron beam loses energy and the radiation field is amplified. For injection below the resonant energy the electron beam will gain energy and the radiation field is weakened.

The gain dependence on the injection energy can be calculated by perturbation theory [39]. The rather long but straight forward calculation is not presented here. The next section derives the explicit results in an alternative and more general way. Till then only the dependence on the injection energy is stated with

$$G \propto -\frac{d}{d(\eta/2)} \frac{\sin^2(\eta/2)}{(\eta/2)^2} \quad , \tag{2.51}$$

$\eta = 4\pi N_U(\gamma - \gamma_R)/\gamma_R$ and N_U the total number of undulator periods. The gain of the low gain

amplifier is related to the spectrum of the spontaneous undulator radiation [40, 41] by taking the frequency derivative of the intensity spectrum of the spontaneous radiation. This relation is known as the Madey-theorem [21].
For the FEL oscillator as well as for the SASE FEL the situation is slightly different, because both types of FEL start from the spontaneous emission with a broad bandwidth in the frequency domain. As a consequence the electron beam is always in resonance with the frequency of the largest gain. A energy dependence as the argument of Eq. 2.51 is not meaningful anymore and must be replace by the frequency dependence. The results are similar by redefining η as $\eta = 2\pi N_U(\omega - \omega_0)/\omega_0$ with ω_0 as the resonant frequency.
In this low gain approximation the interaction between the electrons is almost negligible and the gain is proportional to the total number of electrons. In this one dimensional model of a 'low gain' FEL a higher beam current means a larger amplification of the radiation field. Unless the gain does not exceed several percents the usage of the FEL equations 2.43 and 2.44 is justified. Otherwise the assumption of a constant field K_r is not valid anymore. The radiation power can grow which might change the strength of the electron interaction. To cover this aspect a self-consistent set of FEL equations must be derived as discussed in the following section.

2.3 Self-Consistent FEL Equations

The constraint of the previous section is that the gain of the radiation field must be small over the whole undulator length in order to keep the radiation amplitude K_r constant in the Hamilton function. If the electrons get bunched in the ponderomotive bucket the coherent emission on the resonant wavelength is enhanced. The total emitted power of coherent radiation is proportional to the square of the number of electrons. Thus it is a question of the electron current whether the model of the low gain FEL is valid or not.
Another point is that for a high current beam the electrostatic interaction of the electrons becomes significant. The FEL process is inhibited by these space charge forces because work against the electrostatic field must be done to bunch the charged electrons at a certain phase. To include both effects in a self-consistent manner Maxwell's equations

$$\left[\vec{\nabla}^2 - \frac{d^2}{c^2 dt^2}\right] \vec{A} = -\mu_0 \vec{J} \tag{2.52}$$

$$\left[\vec{\nabla}^2 - \frac{d^2}{c^2 dt^2}\right] \phi = -\frac{\rho}{\epsilon_0} \tag{2.53}$$

have to be solved in addition to the Hamilton equations of motion, providing the current density $\vec{J}$ as the source term for the vector potential $\vec{A}$ and the charge distribution ρ for the scalar potential, with ϵ_0 as the dielectric constant and μ_0 as the magnetic permeability.
The current density and charge distribution are given by

$$\vec{J} = ec \sum_j \vec{\beta}_j(t) \delta(\vec{r} - \vec{r}_j(t)) \quad \text{and} \quad \rho = e \sum_j \delta(\vec{r} - \vec{r}_j(t)) \quad ,$$

where δ is the Dirac-function and $\vec{r}_j(t)$ the trajectory of the jth electron.
The longitudinal electrostatic field is Lorentz contracted as γ^{-2} for higher beam energies. Because the characteristic length of the density modulation of the electron beam is the resonant radiation wavelength (Eq. 2.38), which has the same dependence on the energy as the Lorentz contraction, the longitudinal dependency of the electrostatic field remains the same in the frame of the ponderomotive bucket. Therefore it cannot be neglected for ultrarelativistic electrons.
Although the transverse electric field is even enhanced by a factor γ for relativistic electrons, the induced magnetic field compensates the repulsive forces of the electric field. The residual transverse force scales as γ^{-1}. As discussed later in Section 2.6 as well in Chapter 6 the transverse beam size has to be chosen sufficiently large in order to reduce the FEL degradation by the electron betatron oscillation. For most FELs the transverse electric and magnetic fields have a negligible influence on the FEL performance and therefore they are not regarded in the discussion.
The main component of the longitudinal electric field $E_z = -\partial A_z/\partial t - \partial\phi/\partial z$ is periodic with the ponderomotive wavelength and can be expanded into a Fourier series $E_z = \sum_l \tilde{E}_l \exp[il\theta]$ with $\theta = (k + k_U)z - kct$. The constant term $\tilde{E}_0$ must vanish if the electron beam is regarded as an isolated system of particles. Maxwell's equation for the longitudinal electric field depends on the charge distribution and the current density. One of these source terms can be eliminated in the Fourier series representation by the continuity equations $\partial\rho/\partial t + \vec{\nabla}\vec{J} = 0$. Some basic algebra yields [38]

$$\left[\nabla_\perp^2 - \frac{l^2k^2(1+K^2)}{\gamma_R^2}\right]\tilde{E}_l = i\frac{elk(1+K^2)}{\epsilon_0\gamma_R^2}\sum_j \delta(\vec{r}-\vec{r}_j)e^{-il\theta_j} \quad , \tag{2.54}$$

where θ_j is the ponderomotive phase of the jth electron.
The resonant energy γ_R is defined by $\gamma_R^2 = k(1+K^2)/2k_U$. For a wide electron beam, in particular for the 1D model of a Free-Electron Laser, the contribution of the transverse Laplace operator $\nabla_\perp^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is small and the electric field E_z is proportional to the bunching factor $< \exp(i\theta) >$ or its higher harmonics. The longitudinal oscillation of the electrons for a planar undulator does not influence the electrostatic field because this oscillation is a collective motion and the distance of the electrons remains constant.
The longitudinal motion is almost constant and does not contributes to a radiation field and Maxwell's equation Eq. 2.52 can be restricted to the transverse components alone. The complex linear combination $J = J_x + iJ_y$ and $A = A_x + iA_y$ turns it further into a scalar equation.
The complex vector potential of the radiation field is approximated by

$$A_r = \frac{mc}{e}ue^{ik(z-ct)} \tag{2.55}$$

introducing the complex amplitude $u = -iK_re^{i\Psi}$. Although it might seem that the case of a planar undulator should be treated differently it is not necessary. The previous section has shown that there is no major difference between a helical and a planar undulator for the FEL amplification of the radiation field. The longitudinal oscillation is covered by the coupling factor f_c (Eq. 2.45). The use of the root-mean-square value in K and K_r expresses the weaker field strength in the case of a planar undulator.

The dependence of the complex amplitude u for the free space propagation on $\vec{r}$ and t is weak resulting in the paraxial approximation of the radiation field ([42] or Appendix A). This approximation is valid for the FEL equations as long as the change in the radiation field amplitude u is smaller than the dominant oscillation $\exp[ikz]$ of the radiation field fulfilling the constraint $ku \gg \partial u/\partial z$. A more quantitative expression for this constraint will be given later. In analogy to the radiation field the vector potential of the undulator field is given by

$$A_U = i\frac{mc}{e}Ke^{-ik_Uz} \quad . \tag{2.56}$$

For convenience the transverse dependence of the undulator field is included in the undulator amplitude K. Again this complex vector potential is sufficient to obtain universal FEL equations for both undulator types.

Inserting the fast oscillating velocities β_x and β_y into the current density, neglecting the second order t- and z-derivatives of the radiation field amplitude u the scalar Maxwell's equation becomes in its paraxial approximation

$$\left[\vec{\nabla}_\perp^2 + 2ik\left(\frac{\partial}{\partial z} + \frac{\partial}{c\partial t}\right)\right] u = \frac{ie^2\mu_0}{m}\sum_j\left[\frac{f_cK}{\gamma_j}e^{-i\theta_j} - i\frac{u}{\gamma_j}\right]\delta(\vec{r}-\vec{r}_j) \quad . \tag{2.57}$$

The source term of the radiation field equation can be split into two terms. The second term is proportional to the radiation field u itself. This term changes only the phase but not the amplitude of the radiation field. An interpretation is that the electron beam can be regarded on a macroscopic scale as a dielectric medium. The resulting frequency shift of the radiation field is $\Delta\omega/ck = \Omega_p^2/2c^2k^2$ revealing the Taylor series approximation of the dispersion relation $\omega^2 = c^2k^2 + \Omega_p^2$ for an electron plasma with the plasma frequency [43]

$$\Omega_p = \left[\frac{\mu_0 n_e e^2 c^2}{m\gamma_0}\right]^{\frac{1}{2}} \quad . \tag{2.58}$$

Together with the differential equations of the electron motion a self-consistent set of FEL equations is derived. The change of the electron phase (Eq. 2.43) is in its complex representation

$$\dot{\theta}_j = ck_U - \omega\frac{1+K^2}{2\gamma_j^2} - \omega\frac{\beta_R^2}{2} - \omega\frac{|u|^2}{2\gamma^2} + \omega\frac{f_cK}{2\gamma^2}(ue^{i\theta_j} - \text{c.c.}) \quad , \tag{2.59}$$

where 'c.c.' indicates the complex conjugate.

Similarly the electron energy (Eq. 2.44) changes as

$$\dot{\gamma}_j = -\omega\frac{f_cK}{2\gamma_j}(ue^{i\theta_j} + \text{c.c.}) + \frac{e}{mc}\sum_{l=1}^{\infty}[\tilde{E}_l e^{i\theta_j} + c.c] \quad . \tag{2.60}$$

To complete the set of differential equations the slow betatron motion is given by

$$\dot{X} = \frac{P_x}{\gamma m} \quad , \tag{2.61}$$

$$\dot{P}_x = -\gamma mc^2 \frac{K^2 k_x^2}{\gamma^2} X \quad , \tag{2.62}$$

$$\dot{Y} = \frac{P_y}{\gamma m} \quad , \tag{2.63}$$

$$\dot{P}_y = -\gamma mc^2 \frac{K^2 k_y^2}{\gamma^2} Y \quad , \tag{2.64}$$

where the case of the helical undulator demands to set k_x^2 and k_y^2 equal to $k_U^2/2$.
Eqs. 2.54 and 2.57 – 2.64 are the basic differential equations to describe the physics of a high gain Free-Electron Lasers. Only a few but reasonable assumptions have been made, which are

- relativistic energy of the electrons,
- small transverse extension compared to the undulator period length,
- resonant interaction between radiation field and electron beam.

However these equations are rather complex and it is difficult to analyze them analytically without any further assumptions. Methods to solve them numerically are the topic of Chapter 3.
The discussion is started with the one dimensional model of the Free-Electron Laser. The transverse betatron oscillation has only a weak influence on the FEL dynamics and is neglected in the 1D treatment. Stronger is the dependence on the transverse extension of the radiation field and the electron beam. The main aspect here is the diffraction, where the radiation field tends to spread out transversely and thus to weaken the field amplitude and to change the phase at the center of the electron beam. Although it is important, diffraction does not change the basic working principle of a high gain Free-Electron Laser. Section 2.6 analyze this effect in the extended model of an FEL including the transverse dimension. For the 1D model the transverse Laplace operator in Eqs. 2.54 and 2.57 is dropped.
Another assumption is an infinitely long electron bunch and radiation pulse. At a certain position in the undulator an observer would not see any change in the amplitude of the radiation field or in the modulation of the electron beam in time. In this steady state model the partial time derivative in Eq. 2.57 is neglected. Time-dependent effects such as the slippage of the radiation field with respect to the electron beam are discussed in Section 2.7. The remaining variables of interest in the 1D FEL model are γ, θ and u.
To further simplify the model, it is assumed that the initial electron beam is almost unmodulated. Higher harmonics of the electrostatic field are therefore much reduced compared to the fundamental harmonic ($l = 1$) in Eq. 2.54. Only if the beam is strongly modulated higher modes need to be taken into account. For the applicable range of the following analytic discussion this is never the case. When the equations are solved numerically (see Section 2.5 or Chapter 3) the higher harmonics are included in the calculations.
Another approximation is that only electron beams are regarded which energy spread σ_γ is much smaller than the mean energy γ_0. The energy in the denominator Eqs. 2.57 – 2.60 can be expanded into a Taylor series.

Before trying to find a solution of the 1D FEL equations it is convenient to transform the phase, energy and radiation field amplitude as well as the independent variable t to normalized and scaled variables. The new dimensionless variables are

$$\Phi_j = \theta_j \quad , \tag{2.65}$$

$$\eta_j = \frac{\gamma_j - \gamma_0}{\rho\gamma_0} \quad , \tag{2.66}$$

$$A = \frac{Kkf_c}{4\gamma_R^2 k_U \rho^2} u \quad , \tag{2.67}$$

$$\hat{z} = 2ck_U\rho \left(\frac{\gamma_R}{\gamma_0}\right)^2 t \quad . \tag{2.68}$$

Using these normalized variables, the FEL equations depend only on three parameters. The most important one is the FEL or Pierce parameter [6]

$$\rho = \left[\frac{Kf_c\gamma_0\Omega_p}{4c\gamma_R^2 k_U}\right]^{\frac{2}{3}} \quad . \tag{2.69}$$

The FEL parameter ρ is an expression of the electron density n_e relevant for the FEL amplification.
The space charge parameter

$$\sigma^2 = \frac{\Omega_p^2}{\omega^2}\frac{\gamma_0^2}{\rho^2}\frac{1}{1+K^2} \tag{2.70}$$

describes mainly the repulsing forces of the electrostatic field inhibiting the bunching on the resonant wavelength [36]. Due to the dependence on the plasma frequency, σ^2 drops in the same manner as ρ for a vanishing electron density.
The last parameter is the detuning

$$\delta = \frac{\gamma_0^2 - \gamma_R^2}{2\rho\gamma_R^2} \quad , \tag{2.71}$$

a measure of how much off-resonance in energy the electron beam is injected relative to the resonant wavelength of the seeding radiation field.
It is of importance to estimate the limits of the resulting FEL model due to the approximations made so far. The resonant approximation (see Eq. 2.37) has the strongest impact on the accuracy of the analytic model in this chapter. This approximation states that the electrons and the radiation field have a constant phase relation and that the interaction adds up resonantly. In order to neglect the resonant mode which corresponds to an unphysical velocity of the electrons faster than the speed of light, the phase relation between the electrons and the radiation field must be fast oscillating for this mode. The phase difference with respect to the other mode is $2k_Uz$. Using the normalized variables, where $\hat{z} = 1$ is the characteristic length, the resonance approximation yields the constraint

$$\rho \ll 1 \quad . \tag{2.72}$$

The FEL parameter ρ can be regarded as the error of the accuracy in the normalized variables due to the resonance approximation. Deriving the normalized FEL equations from Eqs. 2.57, 2.59 and 2.60 all terms proportional to ρ are neglected because the resonance approximation omits other terms $\propto \rho$, which are needed to improve the accuracy one order in ρ.
The resonance approximation where the two resonant modes must clearly be separated, has also some limitation for the valid range of the detuning δ, the normalized energy spread $\Delta = \sigma_\gamma/\rho\gamma_0$ and the space charge parameter σ^2. The similar argument that the phase shift of the resonant mode per characteristic length $(2k_U\rho)^{-1}$ should be much smaller than ρ^{-1} yields the constraints

$$\delta \cdot \rho \ll 1 \quad , \tag{2.73}$$

$$\Delta \cdot \rho \ll 1 \quad , \tag{2.74}$$

and

$$\sigma \cdot \rho \ll 1 \quad . \tag{2.75}$$

The last constraint implies that a plasma oscillation of the electron beam should be much slower than the transverse oscillation in the undulator.
Using the dimensionless variables (Eqs. 2.65 – 2.68) the differential equations Eqs. 2.54, 2.57 and 2.59 – 2.60 become the normalized 1D FEL equations

$$\begin{aligned} \Phi'_j &= \delta + \eta_j \quad , && (2.76)\\ \eta'_j &= -\left[\left(A + i\sigma^2\left\langle e^{-i\Phi_j}\right\rangle\right)e^{i\Phi_j} + c.c\right] \quad , && (2.77)\\ A' &= \left\langle e^{-i\Phi_j}\right\rangle \quad , && (2.78) \end{aligned}$$

where the derivative is taken with respect to the normalized position $\hat{z}$. The square brackets $< \ldots >$ indicate to take the average over all electrons of the enclosed argument.
The maximum growth rate of A is about unity, when all electrons emit coherently at a certain phase. In the frame of the electron beam this case occurs if all electrons are bunched with a periodicity of the ponderomotive wavelength $\lambda\lambda_U/(\lambda+\lambda_U) \approx \lambda$ which is much smaller than the electron bunch length. This kind of modulation is called microbunching.
A strong modulation of the electron bunch on the scale of the resonant wavelength is difficult to generate without an FEL. The performance of the FEL is at its optimum when this maximum bunching is achieved. The radiation field cannot be amplified beyond the maximum bunching.
A trivial solution of the FEL equations is the injection of an unmodulated beam with no correlation between energy and phase and no initial radiation field ($A = 0$). If an energy-phase correlation is present the bunching factor $| < \exp(-i\Phi) > |$ may grow. This effect is similar to the reduction of the bunch length and thus an enhancement of coherent emission due to an energy gradient over the whole length, sometimes referred to as ballistic bunching [44]. For any

realistic application of an FEL the initial conditions are at least close to the fix point mentioned above.
If the radiation field A is regarded as a canonical variable it is possible to derive all FEL equations 2.76 – 2.78 from the Hamiltonian [45]

$$H = \sum_j \left[\delta\eta_j + \frac{\eta_j^2}{2} - [iAe^{i\Phi_j} + c.c.] - \sigma^2(\left\langle e^{-i\Phi_j} \right\rangle e^{i\Phi_j} + c.c) \right] . \tag{2.79}$$

The pairs of canonical conjugated variables are (Φ_j, η_j) and $(\sqrt{N_e}A, i\sqrt{N_e}A^*)$ with the equations of motion:

$$\Phi_j' = \frac{\partial H}{\partial \eta_j}, \quad \eta_j' = -\frac{\partial H}{\partial \phi_j}, \quad A' = \frac{1}{N_e}\frac{\partial H}{\partial iA^*}, \quad iA^{*\prime} = -\frac{1}{N_e}\frac{\partial H}{\partial A}.$$

Using this Hamiltonian an efficient method exists to find the solutions of the FEL equations.
If the number of electrons is large the phase space distribution can be approximated by a continuous function $f(\Phi, \eta, \hat{z})$. According to Liouville's theorem [27] the local phase space density does not change while the electrons are moving in phase space. Taking the total 'time' derivative $df/d\hat{z} = \{f, H\} + \partial f/\partial \hat{z}$, where $\{f, H\}$ is the Poisson bracket [37], yields the Vaslov equation

$$\left(\frac{\partial}{\partial \hat{z}} + \Phi' \frac{\partial}{\partial \Phi} + \eta' \frac{\partial}{\partial \eta} \right) f(\Phi, \eta, \hat{z}) = 0 \quad . \tag{2.80}$$

The phase space distribution is expanded into a Fourier series in Φ.
With the trial solution

$$f(\Phi, \eta, \hat{z}) = f_0(\eta) + f_1(\eta) e^{i\Phi + i\Lambda\hat{z}} \quad , \tag{2.81}$$

the problem of solving the FEL equation is transformed into the problem to find a physical reasonable value of Λ for the phase space distribution f. The motivation of this trial solution is given by the scalar Maxwell's equation Eq. 2.78. The source term of the radiation field is proportional to $f_1(\eta)$. Restricting the Fourier series to the fundamental f_0, which is the trivial solution of the Vaslov equation, and the resonant harmonic f_1, this approach refers to the linear regime of the FEL amplification.
The solutions in this limit are discussed in the next section. Unfortunately the equations cannot be solved analytically in the non-linear regime. Section 2.5 shows the results for the non-linear regime, where the 1D FEL equations Eqs. 2.76 – 2.78 have been solved numerically.
With Eq. 2.81 the differential equation of the radiation field is solved by

$$A = -\frac{ie^{i\Lambda\hat{z}}}{\Lambda} \int f_1(\eta) d\eta \quad . \tag{2.82}$$

As mentioned before only the first harmonic of the phase space distribution contributes to the radiation field. Although an unknown distribution f_1 appears in Eq. 2.82 the integration of f_1 over the entire parameter space of η is independent of η and therefore constant.

The linear terms of the differential equations for Φ' and η' are inserted into the Vaslov equation and ordered according to their harmonics in Φ. The evolution of the radiation field and the electron bunching in the linear regime are given by the terms in the Vaslov equation proportional to $\exp(i\Phi)$. The dependence on f_1 can be elimnated by taking a functional derivative using the identity

$$\frac{\partial}{\partial f_1(\eta_1)} f_1(\eta_0) = \delta(\eta_0 - \eta_1) \quad .$$

The resulting expression is the dispersion relation

$$1 + \left(\frac{1}{\Lambda} - \sigma^2\right) \int \frac{\partial f_0}{\partial \eta} \frac{1}{\Lambda + \eta + \delta} d\eta = 0 \quad . \tag{2.83}$$

Any solutions Λ of the dispersion relation solves the Vaslov equation and depends on the detuning δ and the space charge parameter σ^2 as well as on the explicit energy distribution f_0 of the electrons.

An alternative way to derive a dispersion equation is by linearizing the FEL equations. With a similar trial solution for the radiation field ($A \propto \exp(i\Lambda)$) the dispersion equation is cubic [46]. The disadvantage is that this results is valid only for a mono-energetic beam. For any general solution the solution for a certain energy γ_0 has to be convoluted with the energy distribution. Other possible methods to solve the FEL equations are by finding the Green's function [47] or by a different trial solution for the phase space distribution [48]. Quantum mechanical treatments have shown that except for extreme parameter sets an FEL operates in the quasi classical limit, where any quantum effects are almost negligible [49, 50].

2.4 The 1D FEL in the Linear Regime

Starting the discussion for a mono-energetic beam ($f_0(\eta) = \delta(\eta)$) the dispersion relation is transferred to a cubic equation

$$((\Lambda + \delta)^2 - \sigma^2)\Lambda + 1 = 0 \quad . \tag{2.84}$$

Because the parameters δ and σ^2 are real the cubic equation has always 3 solutions being either 3 real numbers or one real number and a pair of complex conjugate numbers.

Each solution Λ defines a mode for the evolution of the radiation field in time as $A \propto \exp(i\Lambda\hat{z})$. The general solution is the sum of all three independent solutions.

In the case of the solution with three real values for Λ each independent mode changes only its phase but not the amplitude. However the observed power $P \propto |A(\hat{z})|^2$ is fluctuating due to the interference of the three modes. The low gain FEL exploits this feature as discussed later in this section. Despite the interference the absolute power of the radiation field and therefore the maximum gain of the FEL is limited to the order of the initial seeding field.

Of particular interest is a solution of one real and two complex values for Λ. The mode belonging to the real solution is oscillating while the others are exponentially decaying and growing. This is the phenomenon of a collective instability. At the beginning of the undulator all modes

have almost identical amplitude and the driving mechanism is the interference of these mode, defining the 'start-up' regime.
After a certain distance the growing mode dominates showing an exponential amplification of the initial field. This is the typical characteristic of a high gain FEL, where the 'start-up' regime is succeeded by the 'exponential' regime. To avoid confusion the term 'exponential' refers only to the growth of the radiation field. Start-up and exponential growth are still given by the same differential equations of the FEL-model. Because this model used only the linear terms, both regimes of the radiation field evolution are combined to the 'linear' regime of the FEL.
If the normalized radiation field A has grown to an amplitude of approximately 0.1 the non-linear terms in the differential equations are not negligible anymore. Numerical simulation are showing (Section 2.5) that the radiation field will be limited to amplitudes of the order of unity. The exponential growth ends in this so called 'saturation' regime.
The growth rate $\Re e(i\Lambda)$ in the linear regime depends only on two independent parameters: the detuning δ and the space charge parameter σ^2.
The dependence on δ is shown in Fig. 2.4. For the calculation the space charge parameter has been set to zero. The region of exponential amplification exists up to $\delta < \delta_{\mathrm{Th}} \approx 1.89$.

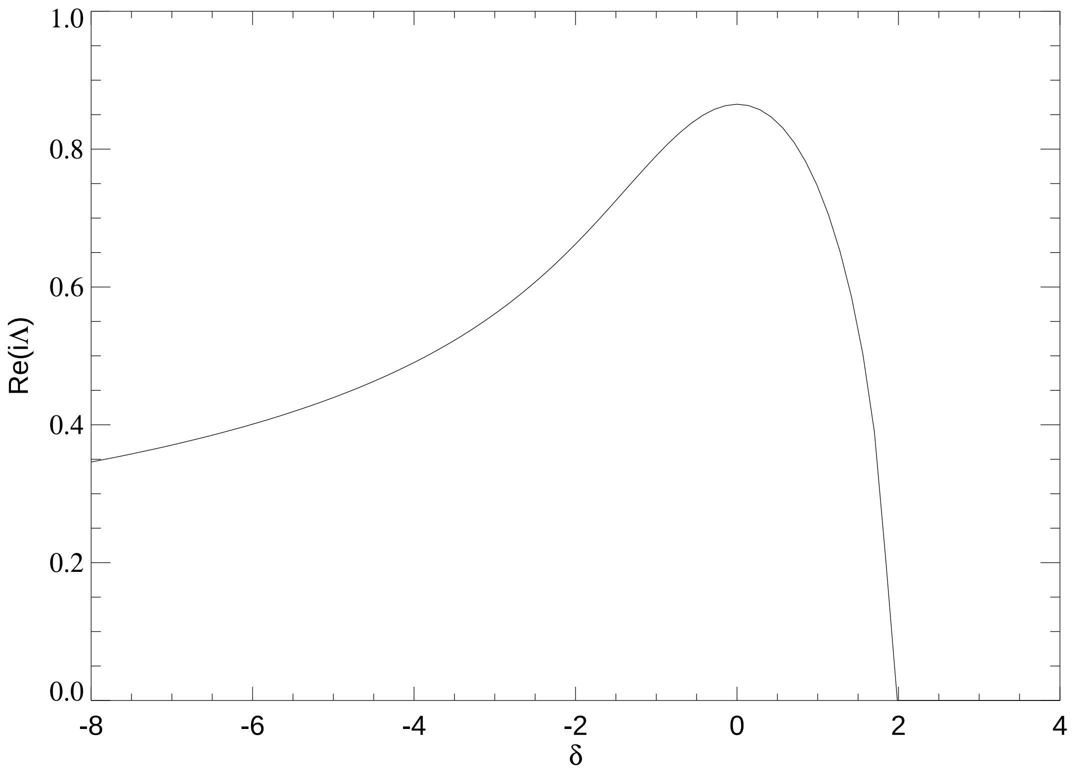


Figure 2.4: *Growth rate versus detuning δ ($\sigma^2 = 0$).*

In the limit of a low current, mono-energetic beam the cubic dispersion relation is reduced to

$(\Lambda+\delta)^2\Lambda+1=0$. The maximum growth rate occurs for $\delta = 0$ with $\Re\mathrm{e}(i\Lambda) = \sqrt{3}/2$. This result seems to be in contradiction with the result of the low gain FEL, where at resonance energy no gain is visible (Section 2.2). The explanation of this difference is that the low gain Free-Electron Laser remains in the start-up regime with $\hat{z} \ll 1$.
To simplify the cubic equation the substitution $\Lambda = \tilde{\Lambda} + \delta$ is performed to eliminate the linear term. The result of this substitution is that the phase slippage of the ponderomotive wave due to the detuning is simply given by an additional phase factor $A(\hat{z}) \rightarrow A(\hat{z})\exp(i\delta\hat{z})$.
The general solution A for an unbunched electron beam and an initial seeding field A_0 is

$$A(\hat{z}) = A_0\left[\frac{\tilde{\Lambda}_1^2 e^{i\tilde{\Lambda}_1\hat{z}}}{(\tilde{\Lambda}_1-\tilde{\Lambda}_2)(\tilde{\Lambda}_1-\tilde{\Lambda}_3)} + \frac{\tilde{\Lambda}_2^2 e^{i\tilde{\Lambda}_2\hat{z}}}{(\tilde{\Lambda}_2-\tilde{\Lambda}_1)(\tilde{\Lambda}_2-\tilde{\Lambda}_3)} + \frac{\tilde{\Lambda}_3^2 e^{i\tilde{\Lambda}_3\hat{z}}}{(\tilde{\Lambda}_3-\tilde{\Lambda}_1)(\tilde{\Lambda}_3-\tilde{\Lambda}_2)}\right] \quad , \tag{2.85}$$

where $\tilde{\Lambda}_i$ is the ith root of the dispersion relation.
Because the detuning parameter δ is proportional to ρ^{-1} the valid parameter range is large for the low gain FEL with $\rho \ll 1$. The solution $\tilde{\Lambda}_i$ can be replaced by its approximation in the limit $|\delta| \rightarrow \infty$. The energy bandwidth, where this approximation is not applicable, shrinks as $\propto \rho$ and is therefore negligible. The roots of the cubic equation $\tilde{\Lambda}^2(\tilde{\Lambda}-\delta) = -1$ up to the order $\mathcal{O}(\delta^{-3})$ are

$$\tilde{\Lambda}_1 = \delta\left(1-\frac{1}{\delta^3}\right) \quad \tilde{\Lambda}_2 = \frac{1}{\sqrt{\delta}}+\frac{1}{2\delta^2} \quad \tilde{\Lambda}_3 = -\frac{1}{\sqrt{\delta}}+\frac{1}{2\delta^2} \quad .$$

For $\delta < 0$ the last two roots have complex values indicating an exponential growth and decay of the radiation field. In the opposite case ($\delta > 0$) all roots are real.
With the definition of the Gain $G = (|A(z)|^2 - |A_0|^2)/|A_0|^2$ in the low gain limit the amplification [51] is

$$G = \frac{4}{\delta^3}\left(1-\cos(\delta\hat{z}) - \frac{\delta\hat{z}}{2}\sin(\delta\hat{z})\right) \tag{2.86}$$

Replacing the normalized variables the total gain G for an undulator with N_U periods is given by

$$G = -\frac{\pi n_e e^2}{4\epsilon_0 mc^2}\frac{f_c^2 K^2\lambda_U^2 N_U^3}{\gamma_0^3}\frac{d}{d(\eta/2)}\frac{\sin^2\eta/2}{(\eta/2)^2} \tag{2.87}$$

with $\eta = \delta\hat{z} = 4\pi N_U(\gamma_0-\gamma_R)/\gamma_R$. This expression is the quantitative verification of the low gain curve, which has be stated in the previous section (Eq. 2.51).
In the same limit the spectrum of spontaneous radiations is covered by the initial conditions of no seeding field ($A(0) = 0$) but a pre-bunched electron beam ($A'(0) \neq 0$). The general solution is obtained if Eq. 2.85 is integrated over $\hat{z}$. Regarding this problem in the opposite direction the low gain FEL radiation field is proportional to the derivative of the radiation field of the spontaneous undulator radiation. This is the main statement of Madey's theorem.

Because the growth rate $\Re\text{e}(i\Lambda)$ is of the order of unity the parameter $2k_U\rho$ (see Eq. 2.68) is a rough estimate for the exponential growth of the radiation field. The inverse value of this parameter defines the gain length

$$L_g = \frac{\lambda_U}{4\pi\rho}. \tag{2.88}$$

So far the presented results are based on a vanishing energy spread of the electron distribution. It can be expected that an energy spread would reduce the field gain because only a fractional part of the distribution covers the resonance of maximum gain. Electrons in the tail of $f_0(\eta)$ will contribute to the amplification process with a much smaller growth rate.
To estimate the impact of the energy spread the dispersion relation Eq. 2.83 has to be solved for the initial distribution f_0, being normalized to unity with $\int_{-\infty}^{\infty} f_0 d\eta = 1$. Analytic solutions exist only for few types of distributions.
The Lorentz distribution

$$f_0(\eta) = \frac{1}{\pi}\frac{\Delta}{\eta^2 + \Delta^2} \tag{2.89}$$

with a normalized energy spread $\Delta = \sigma_\gamma/\rho\gamma_0$ is one of them, which is at least an approximation for the more realistic Gaussian distribution.
A lengthly calculation yields the cubic dispersion equation [46]

$$((\Lambda - i\Delta + \delta)^2 - \sigma^2)\Lambda + 1 = 0 \quad , \tag{2.90}$$

similar to Eq. 2.84 but with complex coefficients. The three roots are not symmetrically located in the complex plane and there might be more than one growing and decaying modes.
In Fig. 2.5 the largest growth rate of all three solution is plotted depending on different settings of the energy spread Δ. In the limit $\Delta \to 0$ the gain curve agrees with the result shown in Fig. 2.4. With increasing energy spread the growth rate for a fixed value of $\delta < 0$ is reduced almost in a linear way. The steep edge of the gain curved at $\delta \approx 2$ is smeared out yielding a slightly increased growth rate for $\delta > 1.9$ with increasing spread.
Beside the predicted reduction of the growth rate the gain curve becomes more and more antisymmetric for $\Delta > 0.5$. For an injection below the resonant energy ($\delta < 0$), where all roots have positive imaginary parts, all modes are exponentially decaying. The complete energy of the radiation field is transfered to the electron beam.
This can be understood regarding the electron motion in the longitudinal phase space. The bucket of the ponderomotive wave is filled by the initial distribution in such a way that the distribution thins out towards higher energy. This happens if the mean energy of the distribution lies at the lower border of the separatrix and the tail towards higher energy covers the bucket. Due to the FEL interaction more electrons in the center of the distribution gain energy than the electrons lose in the tail. The width of the bucket is reduced and electrons may be detrapped if the center of the electron distribution is close to the separatrix of the bucket. This principle is repeated till the bucket completely vanishes.

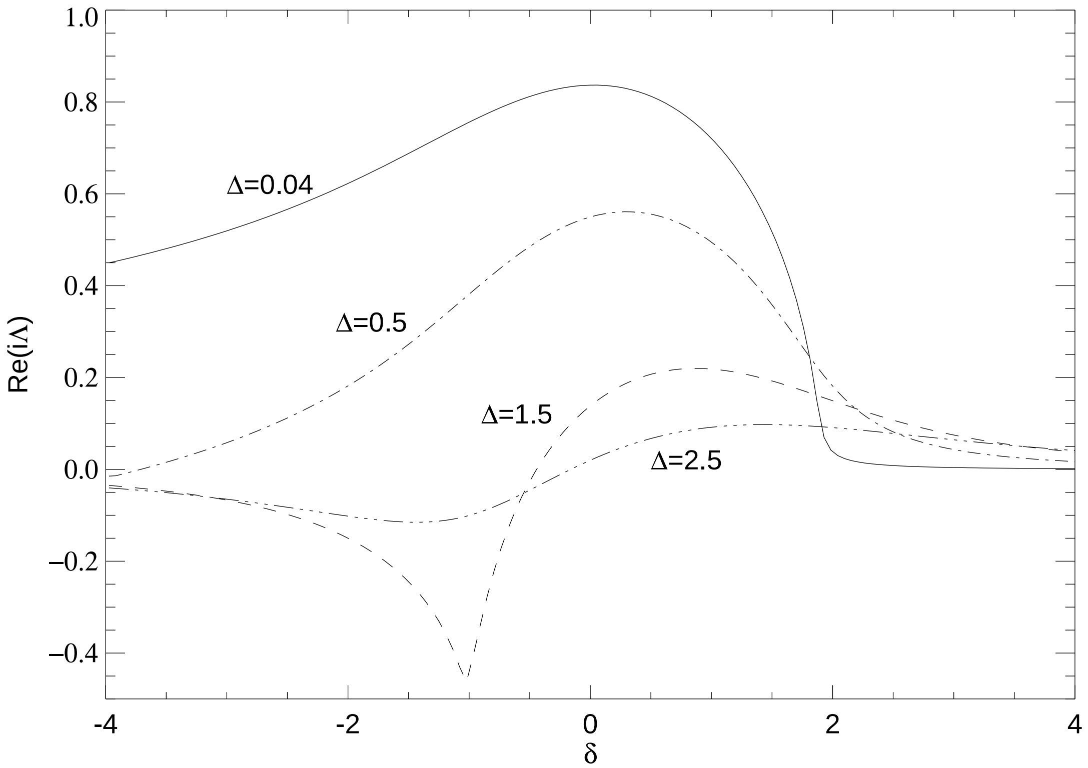


Figure 2.5: *Growth rate versus detuning δ for a Lorentz energy distribution with the energy spread $\Delta = \sigma_\gamma/\rho\gamma_0$, $\rho = 0.001$ and $K = 2$.*

2.5 The 1D FEL in the Non-Linear Regime

For initial conditions, which are typically small initial radiation fields and a nearly unbunched beam, the start-up of the amplification can be described by the FEL equations in the linear regime. The general solution is a sum of several independents modes which are either oscillating, decaying or growing. The initial amplitudes of these modes are of the same order of magnitude and exhibit interference till the exponential growing mode dominates.

An important questions is the limit of the linear model towards large amplitudes. The limiting parameter is the bunching factor $< \exp(-i\Phi) >$, which cannot exceed unity. In the regime of an exponential amplification the radiation field is closely related to the bunching factor by Eq. 2.78. With the rough approximation $|\Lambda||A| = | < \exp(-i\Phi) > |$ and a growth rate close to unity ($|\Lambda| \approx 1$) the radiation field amplitude is of the same order as the bunching factor. The exponential growth is significantly reduced if the radiation field comes close to this limit. At this point the non-linear terms in the FEL equations are not negligible anymore.

To compare with the linear model Eqs. 2.57 – 2.60 are solved numerically. The growth of the radiation field is plotted in Fig. 2.6, both for the linear and non-linear model. For the simulation

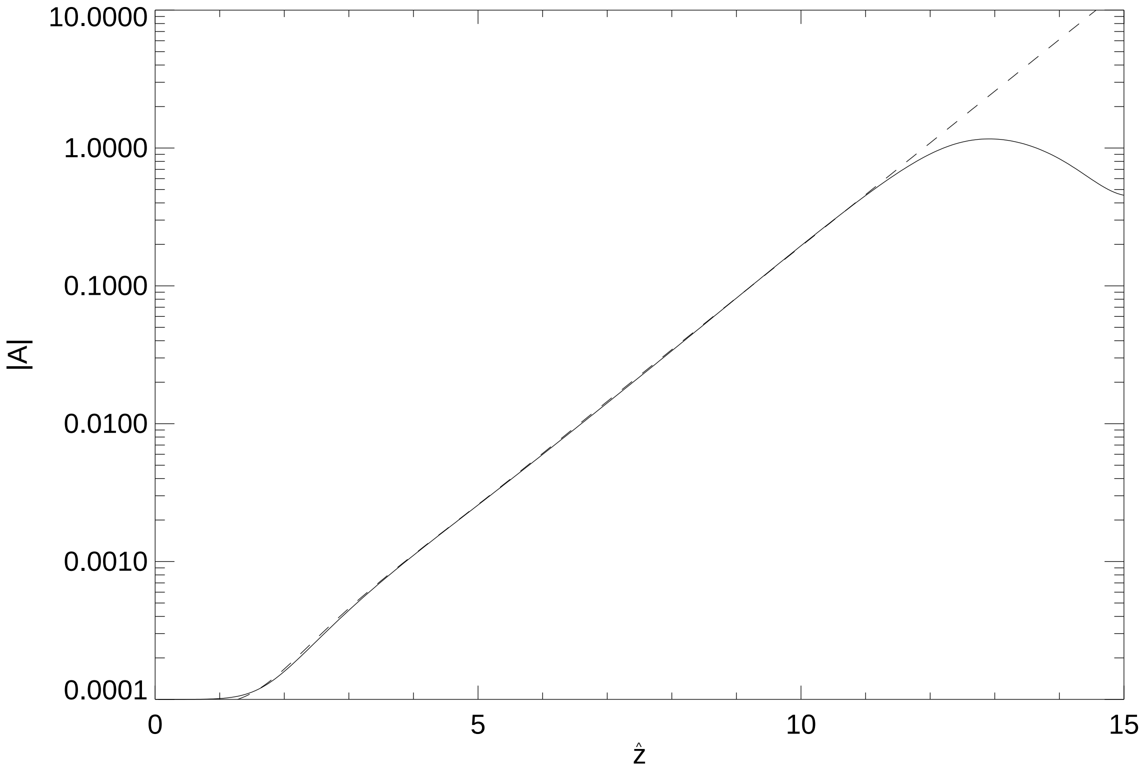


Figure 2.6: *Radiation field amplitude $|A|$ versus normalized position z for the non-linear and linear model, solid and dashed and line, respectively. ($\delta = 0.1$, $\sigma^2 = 0$, $\Delta = 0$).*

macro particles are used to describe the case of a mono-energetic electron beam. The initial homogeneous distribution of the macro particles in Φ provides a vanishing bunching factor. The remaining parameters for this simulation are $\delta = 0.1$ and $\sigma^2 = 0$. The seeding radiation field has the amplitude $A_0 = 10^{-4}$ and the electron bunch is unmodulated. In reality the last statement is not completely true. Due to the finite number of the electrons the beam is slightly modulated. The resulting spontaneous emission might grow faster than the amplification of A_0 in the start-up regime of the FEL. Starting from the spontaneous emission is discussed in Section 2.7.
The linear models agrees well up to a field amplitude of $A \approx 0.3$ with the numerical simulation. The bunching factor, not plotted in Fig. 2.6, is smaller of about roughly 15% compared to the normalized radiation field amplitude A. The growth rate of further amplification is reduced till the radiation field reaches a maximum amplitude of $A \approx 1.2$. Right before the radiation field saturates the maximum value of the bunching factor is obtained with a value of ≈ 0.7.
In addition to start-up and exponential growth of the radiation field the reduction of the FEL amplification process is observed in the so-called 'saturation regime'. The radiation field evolution in this regime can only be calculated by numerical simulation. For this simulation

the maximum gain of the radiation power ($\propto |A|^2$) is $G = 10^8$, about seven orders of magnitude larger than for the low gain FEL.
To be compatible with a single-pass, high gain FEL, a typical low gain FEL operates as an oscillator. Over several passes the trapped radiation field is amplified till it reaches saturation, too. The demands to operate such an FEL oscillator is that the loss factor of the optical cavity is smaller than the single pass gain. The radiation is coupled out either by a hole in one of the mirrors or by the usage of a partial transmitting mirror.

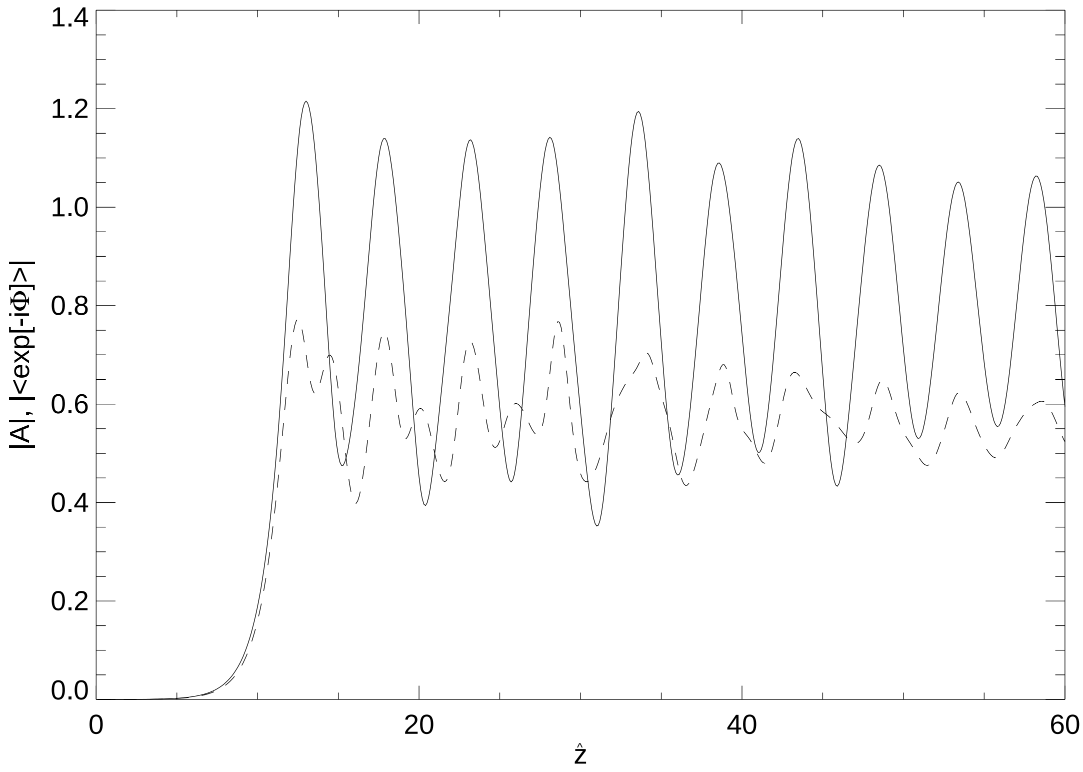


Figure 2.7: *Evolution of radiation field and bunching factor (solid and dashed line, respectively) in the far saturation regime ($\delta = 0.1$, $\sigma^2 = 0$, $\Delta = 0$).*

In Fig. 2.7 the long term evolution of the radiation field and the bunching factor is displayed. After reaching the maximum amplification in the beginning of the saturation regime no further significant amplification is visible. The saturation point is succeeded by an oscillation of the radiation field amplitude. The period length is roughly five gain length and the maximum growth and decay rate is comparable to the exponential regime. The bunching factor tends to follow the oscillation of the radiation field.
For economical reason it is not useful to extend an undulator beyond the point of saturation if the FEL is seeded by an external radiation field unless the undulator parameters (K and λ_U) are matched to compensate the energy loss of the electron beam [52]. With such an undulator

tapering the resonance conditions remains within the bandwidth of the FEL amplifier.
Starting from initial fluctuation of the electron phases the limit of $A \approx 1$ can be exceeded even with an untapered undulator (Section 2.7).

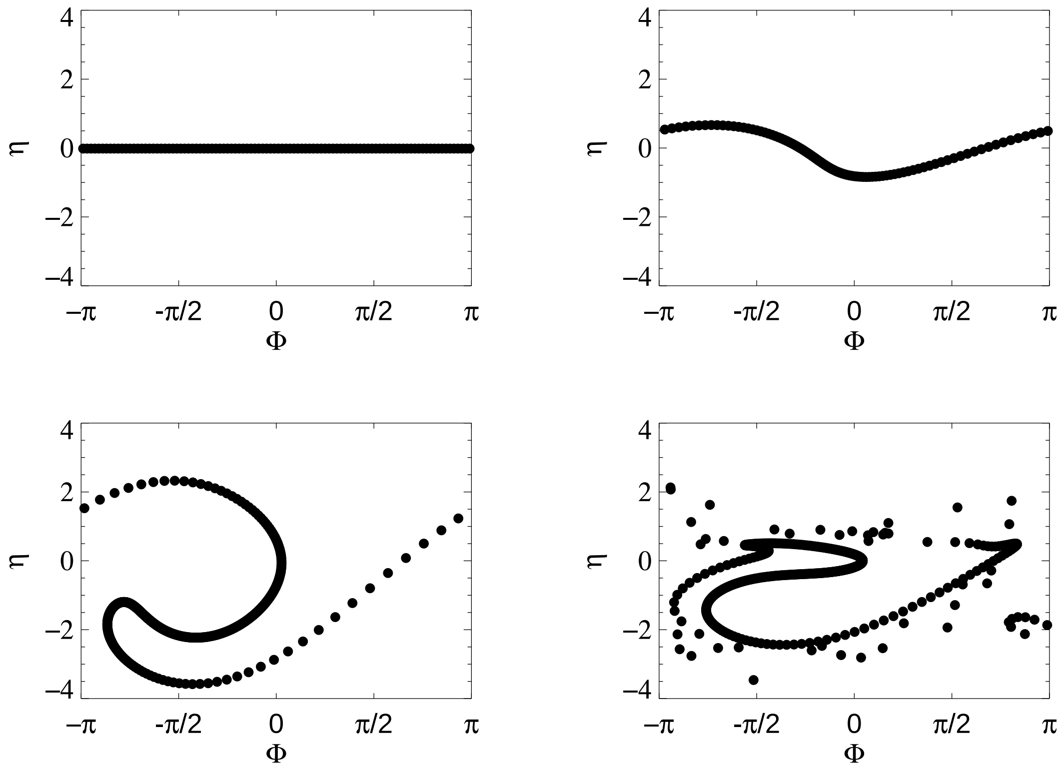


Figure 2.8: *Longitudinal phase space distribution of the electron beam at $\hat{z} = 0$ (upper left), 10.5 (upper right), 13.0 (lower left) and 21.0 (lower right).*

Typical phase space distributions of the electron beam are shown in Fig. 2.8 for different positions in the undulator. The initial distribution (upper left) gets deformed in the linear regime (upper right) until most of the electrons are nearly vertically placed at saturation. Beyond saturation the distribution gets wound up (lower left). Electrons, which are located close to the separatrix of the ponderomotive wave bucket, are detrapped when the radiation field amplitude starts to oscillate (lower right).
Due to the FEL amplification a certain amount of energy is transfered to the radiation field. The energy conservation

$$|A|^2 + \left\langle \frac{\gamma}{\rho\gamma_0} \right\rangle = \text{const} \quad , \tag{2.91}$$

directly derived from the equation of motions, determines the efficiency of the FEL at saturation ($|A| \approx 1$)

$$\text{efficiency} \equiv \frac{\Delta\gamma}{\gamma_0} = \rho \quad . \tag{2.92}$$

Simultaneously the energy spread is increased. To obtain an estimate Eq. 2.77 is multiplied with $\exp[-i\Phi]$. Neglecting space charge it becomes

$$\left\langle \eta' e^{-i\Phi} \right\rangle = -A \quad . \tag{2.93}$$

In the exponential regime the radiation field A is dominated by one mode and Eq. 2.93 can easily be integrated. Multiplying with the complex conjugated of Eq. 2.93 yields the variance of the energy with

$$\left\langle \frac{\Delta\gamma^2}{\gamma_0^2} \right\rangle = \rho^2 \left\langle \eta^2 \right\rangle = \rho^2 \frac{|A|^2}{|\Lambda|^2} \quad . \tag{2.94}$$

If the electron beam is injected close to the resonant energy the absolute value of the growth rate is $|\Lambda| \approx 1$ and the energy spread is related to the efficiency by

$$\frac{1}{\rho} \left\langle \left(\frac{\Delta\gamma}{\gamma_0} \right)^2 \right\rangle = \left\langle \frac{\Delta\gamma}{\gamma_0} \right\rangle \quad . \tag{2.95}$$

The induced energy spread at saturation is ρ, which is identical with the mean energy loss of the electron beam.

2.6 3D Effects

So far only the model of the one dimensional FEL has been discussed. Neither the slow transverse motion of the electrons nor the limited size of the radiation field was taken into account. This section presents an extended model of the FEL including these three dimensional effects. The approach is similar to the previous section where the FEL equations are linearized, but differs in the point that the formal solution of the phase space density f is inserted into Maxwell's equation and not vice versa. Although the formulae are more complicated the basic principle of the FEL is the same as in the previous section.

An easier problem is the slow transverse motion of the electron beam. As seen in Eq. 2.59 the transverse betatron oscillation reduces the longitudinal velocity. In average the resonant wavelength is shifted towards longer wavelength as done by a virtually increased undulator parameter $\tilde{K} = \sqrt{K^2 + \gamma^2 < \beta_R^2 >}$. The variation in β_R^2 for all electrons can be compared to a fluctuation in energy and the energy spread is replaced by an effective energy spread $\tilde{\sigma}_\gamma$ including both, the spread in the energy distribution σ_γ and the impact of the transverse momenta [53, 54].

Because the two terms are independent of each other the effective energy spread is calculated as

$$\tilde{\sigma}_\gamma = \sqrt{\sigma_\gamma^2 + \left(\frac{\epsilon_N \lambda_U}{4\beta\lambda} \right)^2} \quad . \tag{2.96}$$

For the calculation a round beam is assumed where ϵ_N is the normalized transverse emittance and β is the average beta function [28], a measure of the transverse focusing of the undulator. If no extra strong focusing quadrupole lattice is present, the beta function for a matched beam is $\beta = \Omega_U^{-1} = \sqrt{2}\gamma/Kk_U$ (see for comparison Eq. 2.18). The root-mean-square beam radius σ_r is related to the beta function and normalized emittance by $\sigma_r = \sqrt{\epsilon_N \beta/\gamma}$. To significantly improve the FEL performance either the emittance or the energy spread has to be reduced, depending which parameter causes the largest term in the square root of Eq. 2.96.
For undulators, where an external quadrupole lattice is superimposed, the envelope of the electron beam is not constant but oscillates periodically. In the model of an effective energy spread the width of the electron distribution is pulsing in energy. Electrons close to the separatrix of the longitudinal phase space get either trapped or detrapped, depending on the change of the beam size. Another impact is that source term of the wave equation 2.57 changes with any beam size variation. The discussion returns to this point after the field equation has been analyzed.
In the previous sections the space charge field of Eq. 2.54 has been easily implemented into the FEL equations because the transverse Laplace operator can be neglected for a wide electron beam. The resulting field is proportional to the local bunching factor, the average phasor sum of electrons close to the observation point. This remains valid if the transverse beam radius σ_r is larger than

$$\sigma_r^2 \gg \frac{2}{kk_U} \quad . \tag{2.97}$$

As an example, for the FEL of the TESLA Test Facility (TTF FEL) with an undulator period of 2.73 cm and a resonant wavelength of 70 nm the radius of the transverse beam radius must be larger than 9.8 μm. These parameters correspond to a 300 MeV beam and an undulator parameter $K = 0.896$.
Under the assumption that Eq. 2.97 is valid the equations of motion for the electrons remain the same, except that the bunching factor for space charge calculation and the radiation field have to be evaluated locally at the electron position.
In contrast to the previous section the normalization of the variables demands a different definition of the FEL parameter $\hat{\rho}$ with

$$\hat{\rho} = \left[\frac{I}{\gamma_0 I_A} \frac{f_c^2 K^2}{1+K^2} \right]^{\frac{1}{2}} \quad , \tag{2.98}$$

where $I_A = 4\pi mc/e\mu_0 \approx 17$kA is the Alfven current. In order to obtain an identical differential equation for the electron energy the space charge parameter is modified to

$$\hat{\sigma}^2 = 4 \frac{\hat{\rho}}{B} \frac{1+K^2}{K^2} \tag{2.99}$$

where

$$B = 2r_0^2 k k_U \hat{\rho} \tag{2.100}$$

is the diffraction parameter [55]. The definition of the characteristic radius r_0 of the electron beam does not need to agree with the root-mean-square size of the beam. The name 'Diffraction Parameter' might be confusing because a large value of B does not intuitively imply a strong diffraction, but a reduced diffraction instead.

Compared to the 1D model the 3D FEL parameter depends rather on the electron beam current than on the electron density. The reason lies in the diffraction, where the information about the radiation field amplitude and phase is also propagating in the transverse direction. This can be regarded as a kind of integration over the transverse plane and thus replacing the electron density with the current. In the limit of negligible diffraction the 3D model should be consistent with the 1D model as shown later in this section. The 1D and 3D FEL parameter (Eqs. 2.69 and 2.98) are related to each other by the identity

$$\rho = \hat{\rho} B^{-\frac{1}{3}} \quad . \tag{2.101}$$

Similar

$$\sigma = \hat{\sigma} B^{\frac{1}{3}} \tag{2.102}$$

connects the 1D and 3D space charge parameter (Eqs. 2.70 and 2.99).

Energy, phase and radiation field are normalized as in the case of the 1D model (Eqs. 2.65 – 2.67). Also in analogy to the 1D model the longitudinal position within the undulator is scaled with the inverse of the 3D gain length $\hat{L}_g = \lambda_U / 4\pi\hat{\rho}$. The resulting new position $\hat{z} = 2k_U\hat{\rho}z$ is dimensionless. The transverse coordinates are normalized by the beam size r_0 with $\hat{x} = x/r_0$ and $\hat{y} = y/r_0$. It is not compelling that r_0 has to agree with the root-mean-square value of the electron distribution. For uniform or parabolic distributions it is convenient to set r_0 to the maximum radius.

The self-consistent 3D FEL equations become in the linear approximation

$$
\begin{aligned}
\Phi'_j &= \delta + \eta_j \qquad &(2.103)\\
\eta'_j &= -\left[\left(A + i\hat{\sigma}^2 \left\langle e^{-i\Phi_j}\right\rangle\right) e^{i\Phi_j} + c.c.\right] \qquad &(2.104)\\
\left[\hat{\nabla}^2_\perp + 2iB\frac{d}{d\hat{z}}\right] A &= 2i\left\langle e^{-i\Phi_j}\right\rangle \qquad &(2.105)
\end{aligned}
$$

The approach to the equations is similar to the 1D model (Section 2.3) by solving the equation [62]

$$\left[\frac{\partial}{\partial\hat{z}} + \Phi'\frac{\partial}{\partial\Phi} + \eta'\frac{\partial}{\partial\eta}\right] f = 0 \quad . \tag{2.106}$$

The distribution function f is expanded into a Fourier series of the electron phase Φ in the ponderomotive wave. For the linear regime only the fundamental f_0 and the first harmonic f_1 are of interest.

Although Eq. 2.106 is formally identical with Eq. 2.80 it is not a Vaslov equation because the dependence of the phase space distribution f on the transverse variables ($\vec{\beta}'_\perp \partial f/\partial\vec{\beta}_\perp$ and

$\vec{x}'_\perp \vec{\nabla}_\perp f)$ is omitted. They can be neglected because $\vec{x}'_\perp$ and $\vec{\beta}'_\perp$ are fast oscillating ($\propto \exp(ik_U z)$) and neither resonant to the fundamental f_0 nor to the first harmonic f_1 but to the second or higher harmonics, which are not discussed here.
Collecting terms, which are resonant to the first harmonic, the $\hat{z}$ and phase derivative of f_0 as well as the energy derivative of f_1 drops out. Eq. 2.106 has the formal solution [56]

$$f_1 = -\int_0^{\hat{z}} d\hat{z}' \eta' \frac{\partial f_0}{\partial \eta} e^{-i\Phi'(\hat{z}-\hat{z}')} \quad . \tag{2.107}$$

The major difference for the 3D model arises due to the limited transverse extension of the radiation field. The diffraction tends to spread out the radiation field transversely and to thin out the field amplitude at the electron beam position. In addition the radiation field might have an extra phase shift relative to the electron beam disturbing the synchronization of the FEL amplification process.
Appendix A treats the propagation of a radiation field in free space. The Rayleigh length

$$z_R = \frac{k r_G^2}{2} \tag{2.108}$$

is the characteristic parameter for diffraction, where r_G is the root-mean-square radius of a fundamental Gaussian distribution of the radiation field. If r_G is identical with the radius r_0 of the electron beam the diffraction parameter B is proportional to the ratio of the Rayleigh length z_R to the gain length $\hat{L}_g$. Therefore it can be predicted that for large value of B the 3D model must yield the same results as the 1D model, because diffraction becomes less noticeable.
Closely related to the diffraction parameter B is the gain guiding of the radiation field. The diffraction of the radiation field is compensated in a high gain FEL by the field gain. At its equilibrium state the radiation field has a constant radius. Without any explicit calculations this radius is comparable to the electron radius if B is large because the amplification dominates. For a low gain FEL oscillator the properties such as waist position and waist radius is defined by the cavity mirrors, although the modification of the radiation field by the amplification has to be regarded for a stable condition of such an oscillator.
The three differential equations Eqs. 2.103 – 2.105 are combined by inserting the solution f of the Eq. 2.106 into the source term $< \exp(-i\Phi) >= \int f_1(\eta) d\eta$ of Maxwell's equation. The bunching factor in the differential equation of η' is proportional to the source term of the radiation field and is replaced by the left hand side of Eq. 2.105. With the trial solution $A \propto \exp(i\Lambda)$ the resulting differential equation of the radiation field amplitude [55] is

$$[\hat{\nabla}_\perp^2 + \mu^2] A(\hat{x}, \hat{y}, \hat{z}) = 0 \tag{2.109}$$

with

$$\mu^2 = \frac{-2D}{1-\hat{\sigma}^2 D} - 2B\Lambda \quad , \tag{2.110}$$

$$D = \int d\eta \frac{\partial f_0 / \partial \eta}{\eta + \delta + \Lambda} \quad . \tag{2.111}$$

To derived Eq. 2.109 Λ is restricted to have at least a negative imaginary part, exhibiting an exponentially growing instability. A more detailed way to derived Eq. 2.109 can be found in Appendix B, where the initial value problem is solved by Laplace transformation.
It is difficult to find a general solution of Eq. 2.109 because D depends on the transverse electron distribution $f_0 \equiv f_0(\eta, \hat{x}, p_x, \hat{y}, p_y)$ [57]. Any transverse motion of the betatron oscillation can be expressed in terms of the effective energy spread (Eq. 2.96). Further an axi-symmetric electron beam with the radius r_0 is assumed. The arguments of f_0 is reduced to the energy η and radial position $\hat{r}$.
Under these assumption the general solution A can be expanded into series of azimuthal eigenfunction $\exp[im\phi]$, where m is an integer number. Although different radial modes exist they do not form a set of complete and orthonormal eigenfunctions. Therefore the initial value problem is not trivial to solve (Appendix B).
The radiation field

$$A(\phi, \hat{r}, \hat{z}) = \sum_{m=-\infty}^{\infty} \sum_{n=0}^{\infty} c_{mn} R_{mn}(\hat{r}) e^{im\phi + i\Lambda\hat{z}}$$

is a superposition of TEM_{mn}-modes fulfilling the equation

$$\left[\frac{d^2}{d\hat{r}^2} + \frac{1}{\hat{r}}\frac{d}{d\hat{r}} + \mu_{mn}^2 - \frac{m^2}{\hat{r}^2}\right] R_{mn}(\hat{r}) = 0 \tag{2.112}$$

to determine the eigenvalue $\mu_{mn} \equiv \mu(\Lambda_{mn})$ and eigenfunction R_{mn}.
For physically reasonable solutions the radiation field must drop to zero in the limit $\hat{r} \to \infty$ and must be free of any singularity at the origin $\hat{r} = 0$.
The simplest problem is the case of a stepped beam profile with $f_0 = f_0(\eta)\Theta(1 - \hat{r})$, where $\Theta(1 - \hat{r})$ is the step or Heaviside function. For this profile μ is piecewise constant and the general solutions are Bessel or modified Bessel functions. Outside the electron beam ($\hat{r} > 1$) the modified Bessel function of the second kind $K_n(g\hat{r})$ vanishes for large values of $\hat{r}$. The scaling factor $g = \sqrt{2B\Lambda}$ with $\Re\mathrm{e}(g) > 0$ is applied to obtain a solution of Eq. 2.112.
Within the electron beam ($\hat{r} \leq 1$) the Bessel function $J_m(\mu\hat{r})$ solves the differential equation. As demanded the radiation field has no singularities at the origin. The two regions are connected by the constraints of continuity of $R_{mn}(\hat{r})$ and its derivative at $\hat{r} = 1$. Using basic relations of the Bessel functions [58] the continuity conditions are transferred into the dispersion relation

$$\mu J_{m+1}(\mu) K_m(g) = g J_m(\mu) K_{m+1}(g) \quad . \tag{2.113}$$

Arising from the oscillatory behavior of the Bessel function $J_m(\mu\hat{r})$ more than one possible solution exist. The different eigenvalues of Λ are identified by the index n.
The dispersion equation is solved numerically to obtain the dependence on the beam properties, which are completely defined by three parameters: the detuning δ, the diffraction parameter B, the space charge parameter $\hat{\sigma}^2$ as well as the distribution $f_0(\eta)$ in the electron energy. For simplicity the calculations to find the eigenvalues Λ are done for no space charge forces and zero energy spread. Their impact is very similar to that encountered for the 1D model, where $\hat{\sigma}^2 \neq 0$ or $\Delta \neq 0$ degrades the growth rate of the FEL.

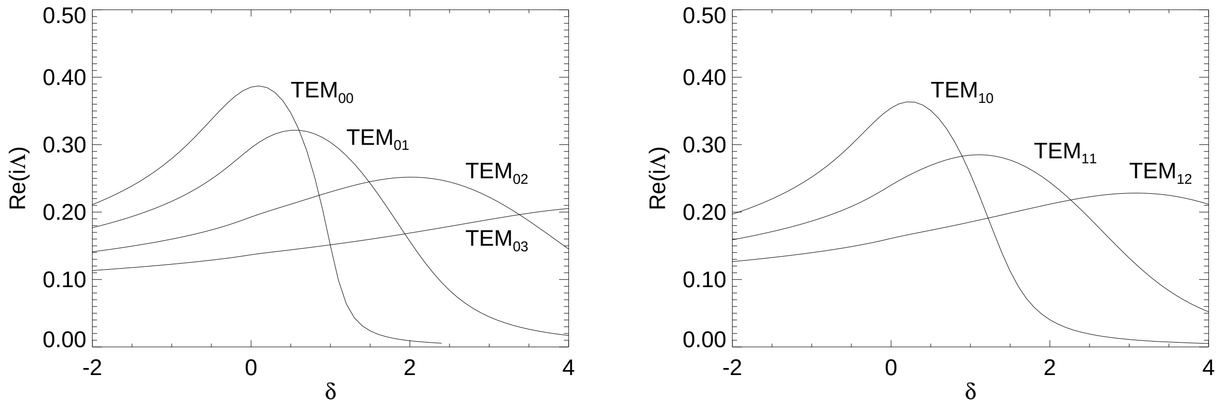


Figure 2.9: *Growth rate versus detuning δ for different TEM modes ($B = 10$, $\hat{\sigma}^2 = 0$, $\Delta = 0$).*

For a diffraction parameter of $B = 10$ the dependence of the growth rate on the detuning parameter δ is plotted in Fig. 2.9. The coexistence of several radial and azimuthal modes is the main difference to the single growing mode of the 1D model.
The general shape of the gain curve is identical for all modes with an edge at $\delta > 0$ and a long tail towards negative values of δ. The higher radial or azimuthal modes exhibit a reduced maximum growth rate, broader bandwidth and a shifting of the curve towards positive values of δ.

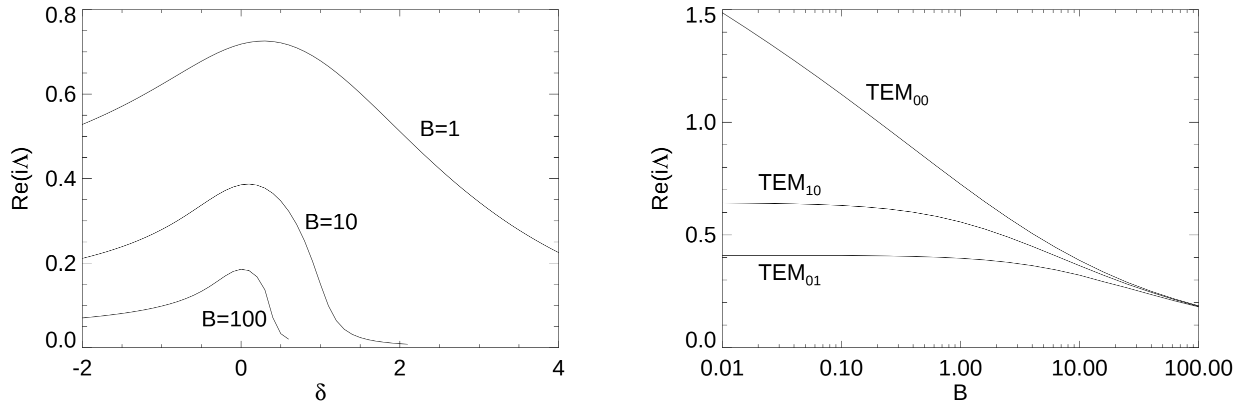


Figure 2.10: *Gain curve of the ground mode TEM_{00} (left) and maximum growth rate of the lowest modes (right) depending on the diffraction parameter B ($\hat{\sigma}^2 = 0$, $\Delta = 0$).*

For stronger diffraction, thus a smaller diffraction parameter, the growth rate increases while the bandwidth gets wider. The maximum growth shifts towards larger value of δ. The left plot of Fig. 2.10 shows the gain curves of the ground mode for different settings of the diffraction

parameter. Although difficult to see in the plot the bandwidth narrows down if the diffraction parameter is increased.
Assuming that the optimum detuning for maximum gain is found the growth rate depends on the diffraction parameter B as shown for the three lowest modes in the right plot of Fig. 2.10. Except for the ground mode all modes do not exceed a certain threshold value. Only the TEM_{00} would exhibit a singularity for $B \to 0$ if the value of B is unlimited towards zero. This is not the case because the approximation has been made that the electrostatic field depends only on the local distribution of the electron beam (Eq. 2.97). With the definition of the diffraction parameter in Eq. 2.100 the lower limit of B is

$$B \gg \hat{\rho} \tag{2.114}$$

for valid results of this analytic model.
In the limit $B \to \infty$ the gain curves coincide. The maximum growth rate is independent of the mode and drops as $\Re\mathrm{e}(i\Lambda) \approx B^{-1/3}$. The growth rate of the radiation field is $2k_U\hat{\rho}B^{-1/3}$. Due to Eq. 2.101 it agrees with the 1D FEL model.
Another method to verify the limit towards 1D model is given by the dispersion relation. For $B \to \infty$ the Bessel function $K_n(g)$ and $K_{n+1}(g)$ are identical in the asymptotic behavior for large arguments [59]. To solve the dispersion equation $\mu J_{n+1}(\mu) = gJ_n(\mu)$ the eigenvalue μ must be close to a root of the Bessel function J_n. The second term of μ^2 in Eq. 2.110 is g^2 and must be compensated by the first term for any finite root of the Bessel function. Although the individual terms have large values μ itself is small and can be set to zero. If space charge and energy spread are neglected the dispersion equation becomes $B\Lambda(\Lambda+\delta)^2 + 1 = 0$. Both parameters, Λ and δ, are normalized quantities. The extra factor B can be transferred into these parameters by replacing the normalizing factor $\hat{\rho}$ with $\hat{\rho}B^{-1/3}$, the 1D FEL parameter.

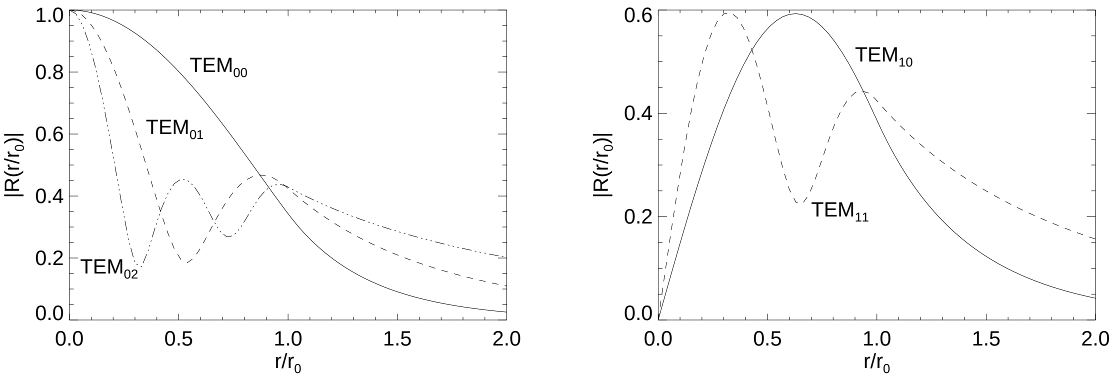


Figure 2.11: *Radial radiation profile R_{mn} of different TEM-modes. Each mode has been optimized for maximum growth rate*

Using the obtained eigenvalue Λ_{mn} the radiation profile of the TEM_{mn} mode is given by

$$R_{mn}(\hat{r}) = \Theta(1-\hat{r})J_m(\mu_{mn}\hat{r}) + \Theta(\hat{r}-1)\frac{J_m(\mu_{mn})}{K_m(g_{mn})}K_m(g_{mn}\hat{r}) \quad ,$$

where $\Theta(\hat{r})$ is the Heaviside function. Unfortunately the differential operator $H = \hat{\nabla}_\perp^2 + \mu^2$ of Eq. 2.109 is not Hermitian and therefore the eigenfunctions R_{mn} are not orthonormal in n. The initial amplitude of each eigenmode is proportional to the overlap integral [62] as shown in Appendix B. This initial value problem is solved by using Laplace transformation.
The general profile of the TEM radiation fields is given by the azimuthal and radial mode number, m and n, respectively. Each profile has n minima and can be approximated for $\hat{r} \ll 1$ by $R_{mn}(\hat{r}) = \hat{r}^m$. Higher radial modes exhibit larger diffraction and converge towards zero for $\hat{r} \to \infty$ less rapidly than lower modes. Fig. 2.11 shows the radiation profiles of different eigenfunctions $R_m n$ for the stepped electron beam profile.
For any arbitrary electron distribution $f_0(\eta, \hat{r}) = f_\eta(\eta) S(\hat{r})$ as a product of the distributions in energy and transverse position the general solution can be approximated by piecewise constant electron density of a multi-layer model [55]. The solution of each layer is a linear combination of Bessel and Neumann function $R(\hat{r}) = a_j J_n(\mu_j \hat{r}) + b_j Y_n(\mu_j \hat{r})$, where $\hat{r}$ is restricted to the jth layer $\hat{r}_{j-1} < \hat{r} < \hat{r}_j$. The eigenvalues μ_j are evaluated at the center of each layer according to Eq. 2.110. The eigenvalue problem is reduced to solve the continuity condition for each layer and the amplitudes for each layer can be calculated iteratively, starting from the inner most layer ($j = 1$) with $b_1 = 0$ to avoid a singularity. The new amplitudes are found by a simple linear transformation

$$\begin{pmatrix} a_{j+1} \\ b_{j+1} \end{pmatrix} = T_j \begin{pmatrix} a_j \\ b_j \end{pmatrix} \tag{2.115}$$

with the matrix elements

$$\begin{aligned}
(T_j)_{11} &= \frac{\pi}{2}\hat{r}_j[\mu_j J_{n+1}(\mu_j \hat{r}_j) Y_n(\mu_{j+1}\hat{r}_j) - \mu_{j+1} J_n(\mu_j \hat{r}_j) Y_{n+1}(\mu_{j+1}\hat{r}_j)] \quad , \\
(T_j)_{12} &= \frac{\pi}{2}\hat{r}_j[\mu_j Y_{n+1}(\mu_j \hat{r}_j) Y_n(\mu_{j+1}\hat{r}_j) - \mu_{j+1} Y_n(\mu_j \hat{r}_j) Y_{n+1}(\mu_{j+1}\hat{r}_j)] \quad , \\
(T_j)_{21} &= -\frac{\pi}{2}\hat{r}_j[\mu_j J_{n+1}(\mu_j \hat{r}_j) J_n(\mu_{j+1}\hat{r}_j) - \mu_{j+1} J_n(\mu_j \hat{r}_j) J_{n+1}(\mu_{j+1}\hat{r}_j)] \quad , \\
(T_j)_{22} &= -\frac{\pi}{2}\hat{r}_j[\mu_j Y_{n+1}(\mu_j \hat{r}_j) J_n(\mu_{j+1}\hat{r}_j) - \mu_{j+1} Y_n(\mu_j \hat{r}_j) J_{n+1}(\mu_{j+1}\hat{r}_j)] \quad .
\end{aligned}$$

The solution of the last layer $j = j_0$ is matched in amplitude and derivate outside the beam at $\hat{r}_{j_0} = 1$ with the modified Bessel function $cK_n(g\hat{r})$. The parameter g is independent of the radial distribution and identical with the stepped profile. The complete boundary condition can be expressed by the matrix equation

$$\begin{pmatrix} J_n(\mu_{j_0}) & Y_n(\mu_{j_0}) \\ \mu_{j_0} J_{n+1}(\mu_{j_0}) & \mu_{j_0} Y_{n+1}(\mu_{j_0}) \end{pmatrix} \times T_{j_0 - 1} \times \ldots \times T_2 \times T_1 \begin{pmatrix} a_1 \\ 0 \end{pmatrix} = \begin{pmatrix} cK_n(g) \\ cgK_{n+1}(g) \end{pmatrix} \quad . \tag{2.116}$$

If T_j is the identity matrix, Eq. 2.116 is reduced to the dispersion equation of the stepped profile. Again certain values of Λ solve the equation defining the different radial modes. The radiation field of a parabolic distribution ($S(\hat{r}) = 2(1 - \hat{r}^2)$) and Gaussian distribution ($S(\hat{r}) = 9\exp[-9\hat{r}^2]$) of the electron beam is shown in Fig. 2.12. The two plots differ because the

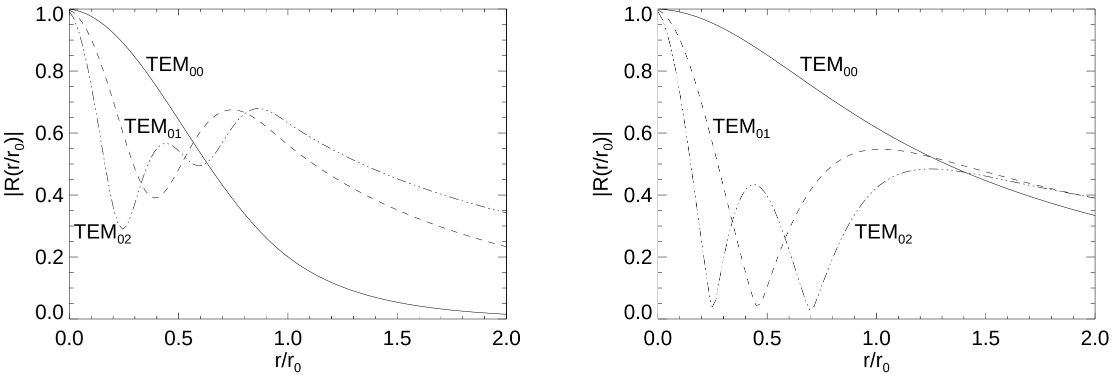


Figure 2.12: *Radiation field profile for a parabolic (left) and Gaussian (right) electron beam distribution ($B = 10$). Only the solutions for the parabolic distribution have been optimized for maximum growth rate.*

root-mean-square size of the Gaussian distribution is much smaller than for the parabolic distribution. The diffraction is larger and the radiation field extends further in the space outside the electron beam. Another difference is that the detuning δ has not been optimized for maximum growth in the case of the Gaussian profile. The imaginary part of μ is smaller and the argument of the Bessel functions is dominated by the real part. Thus the oscillating of J_m and Y_m is marked and the minima in the radiation profile are more distinct.
At the end of this section the impact of the transverse betatron oscillation of the electron motion is discussed. The derivation of the eigenvalue equation Eq. 2.109 has been done under the assumption of no transverse motion of the electrons. The result is the decomposition of the radiation field evolution into independent modes. Each radiation mode corresponds to a similar radially and azimuthally dependent bunching factor. Including transverse motion this bunching profile will be 'carried' away by the electrons. The fundamental mode is less sensitive than higher modes because it has no sharp contours or minima, which can be smeared out by the electron motion. Beside the reduction of the growth rate it yields a coupling of the modes. Another source of mode coupling is a non axi-symmetric electron beam as it is the case for strong focusing by a superimposed lattice of alternating quadrupoles. The beam spot tends to be elliptical and the diffraction parameter is different for the two transverse axis causing astigmatic radiation modes.

2.7 Time Dependence

The whole discussion of the previous sections was done under the assumption that the radiation pulse and the electron bunch is infinitely long. In this model slippage, when the radiation field advances one radiation length relative to the electron bunch while propagating over one undulator period length, is negligible.

Regarding an arbitrarily chosen position $\hat{z}$, the radiation field properties remain unchanged and the time derivative can be dropped in the wave equation Eq. 2.57. Due to the infinite, coherent radiation pulse an observer would measure only a single frequency in his spectrum. This model is called single frequency or steady state model.
Before the time dependence is analyzed it is necessary to have a closer look on the two variables $\hat{z}$ and t. If is convenient to chose the position $\hat{z}$ within the undulator as the independent variable instead of the time t. For the electron motion the time becomes a new canonical variable replacing $\hat{z}$ in the Hamilton function. The switching of the independent variable is formally done by a Legendre transformation resulting in an artificial minus sign of the longitudinal position. This is the mathematical expression for the obvious fact that the head of the bunch arrives earlier at a certain position than the tail.
For simplification the moving frame of the ponderomotive wave is preferred. Here t is transformed into a slow varying variable already introduced as the electron phase $\theta = (k+k_U)z - kct$. In this new frame the time derivative has to be modified [62]. Some straight forward calculations show that the derivative operator $\partial / c\partial t$ is scaled by $(\beta_0 - 1)^{-1}$.
To understand this normalization the differential operators are integrated over a small distance. For an interval of $\Delta z = \lambda_U$ of the independent variable the corresponding time interval is $\Delta t = (\beta_0 - 1)\lambda_U / c = -\lambda / c$. The time derivative changes the radiation field by $\Delta A(t_0) \approx A(t_0 - \lambda/c) - A(t_0)$. This is nothing else than the effect of the slippage, where the radiation field advances one radiation wavelength in this moving frame over one undulator period length.
An important aspect, described only by a time-dependent FEL model, is the phenomenon of self-amplified spontaneous emission. As mentioned in the previous sections the FEL radiation can either be started by an initial seeding field or a modulated electron beam. Due to the finite number of electrons per radiation wavelength the initial bunching factor may be small but not necessarily zero. The amplitude of the fluctuation depends on the total number of electrons per bucket.
The basic assumption of the following analysis is that the electron position is not correlated with energy but that it is random. These random positions t_k yield the sum of random phasors $\exp[-i\omega t_k]$ for the source term of the Maxwell's equation.
The random nature of the source term implies that the physics of a Self-Amplified Spontaneous Emission Free-Electron Laser (SASE FEL) is regarded as a stochastic problem.
Unless the bunch length is comparable to the radiation wavelength the phase distribution of the phasor sum is uniform while the absolute square of the bunching factor $b(\omega) = < \exp[-i\omega t_k] >$ — the mean value is taken over the whole electron bunch — follows a negative exponential probability distribution [63]

$$p(|b(\omega)|^2) = \frac{1}{< |b(\omega)|^2 >} \exp\left[- \frac{|b(\omega)|^2}{< |b(\omega)|^2 >} \right] \quad , \tag{2.117}$$

with

$$< |b(\omega)|^2 > = \frac{1}{N}. \tag{2.118}$$

and N as the total number of electrons. Eq. 2.117 remains the same if the averaging of the

bunching factor is done only over one radiation wavelength. In this case N has to be replace with the average number N_B of electrons per bucket.
For the following discussion the problem is reduced to the 1D model in the limit of negligible space charge forces and a mono-energetic electron beam. The solution in the steady state model with no seeding field but initial bunching b_0 is

$$A(\hat{z}) = \frac{b_0}{i}\frac{\Lambda_1 \exp(i\Lambda_1 \hat{z})}{(\Lambda_1-\Lambda_2)(\Lambda_1-\Lambda_3)} = \frac{b_0}{3i}\exp\left(\frac{\sqrt{3}}{2}\left[1-\frac{1}{9}\delta^2\right]\hat{z} + i\left[\frac{1}{2}-\frac{2}{3}\delta+\frac{1}{18}\delta^2\right]\hat{z}\right) \tag{2.119}$$

for the fastest growing mode. In the last step the roots Λ_i of the cubic equation $(\Lambda+\delta)^2\Lambda+1=0$ have been expanded into Taylor series around the optimum tuning of maximum growth at $\delta = 0$. In the high gain limit the oscillating and exponentially decaying mode can be neglected.
This result is very similar for the Helmholtz equation, the field equation in the frequency domain. The time derivative in Eq. 2.57 is simply replaced by the factor $-i\omega/c$ in the paraxial approximation. Note that the steady state solution $A(\hat{z})$ implicitly depends on ω by the detuning δ. A shift in the wavelength would change the resonant condition according to the approximation

$$\delta = -\frac{\omega-\omega_0}{2\rho\omega_0}$$

with $\omega_0 = 2ck_U\gamma_0^2/(1+K^2)$. The detuning is defined slightly differently for the SASE FEL. In an FEL amplifier only one certain frequency is present and the energy of the electron beam might be more or less on resonance. Due to the broad bandwidth of the spontaneous undulator radiation [40], which seeds the SASE FEL, the electron beam is automatically in resonance with a certain frequency. A different electron energy would not change the growth rate but change the frequency which exhibits maximum growth. Therefore δ depends rather on ω for SASE FELs than on γ_0 for FEL amplifiers.
The Helmholtz equation is solved by adding a phase factor to the solution of the steady state model ($A(\hat{z})\exp[i(\omega/2ck_u\rho)\hat{z}] \rightarrow \tilde{A}(\omega,\hat{z})$). Compared to Eq. 2.119 the Fourier amplitude $\tilde{A}(\omega,\hat{z})$ is not dimensionless but has a 'time'-unit instead. In the remaining part of this section a rectangular pulse profile with an averaged beam current I_0 is assumed. Under this assumption the bunching factor b_0 can be replaced with the normalized Fourier component of the beam current $I(\omega)/I_0$ in order to get the right unit for $\tilde{A}(\omega,\hat{z})$.
In SI units (Eqs. 2.67 and 2.55, respectively), the Fourier amplitude of the electric field becomes

$$\tilde{E}(\omega,\hat{z}) = \frac{E_0}{3i}\exp\left(\frac{\sqrt{3}}{2}\left[1-\frac{1}{9}\delta^2\right]\hat{z} + i\left[\frac{\omega}{2ck_U\rho}+\frac{1}{2}-\frac{2}{3}\delta+\frac{1}{18}\delta^2\right]\hat{z}\right)\frac{\tilde{I}(\omega)}{I_0} \tag{2.120}$$

with

$$E_0 = \frac{4mc^2}{e}\cdot\frac{k_U\gamma_0^2\rho^2}{K} \quad . \tag{2.121}$$

The ratio between electric field and seeding bunching in the electron beam current is a factor in the frequency domain, depending on $\hat{z}$ and ω. Formally it can be identified with a Green's function [47, 64].
The absolute square of the radiation spectrum [65]

$$|E(\omega,\hat{z})|^2 = A(\hat{z}) \exp\left(-\frac{(\omega-\omega_0)^2}{2\sigma_A(\hat{z})^2}\right) |\tilde{I}(\omega)|^2 \quad , \tag{2.122}$$

with

$$A(\hat{z}) = \left(\frac{E_0}{3I_0}\right)^2 \exp(\sqrt{3}\hat{z}) \quad , \qquad \sigma_A(\hat{z}) = \sqrt{\frac{18}{\sqrt{3}\hat{z}}}\rho\omega_0$$

exhibits a Gaussian-like amplification around the resonant frequency ω_0 with the root-mean-square bandwidth σ_A for the exponential regime [65, 67]. In the limit $\hat{z} \rightarrow 0$, not carried out here, the radiation spectrum is identical with the spontaneous undulator radiation.
For the validity of Eq. 2.122 the pulse length T must be large so that the bandwidth of the FEL amplification σ_A encloses the full bandwidth of the radiation pulse. This constraint is expressed by the condition $\sigma_A T \gg 1$ and is discussed in detail later in this section.
According to Eq. 2.120 the radiation field $\tilde{E}(\omega,\hat{z})$ fluctuates in the same manner as the Fourier component of the electron current. The power spectrum density (Eq. 2.122) follows a negative exponential probability for any arbitrarily chosen frequency. The distribution changes significantly if the power spectrum is partially integrated over the frequency. The frequency range $\Delta\omega$, within the negative exponential probability distribution for the partially integrated spectrum remains unchanged, is called the spectral coherence. It is derived later in this section.
In the saturation regime the radiation field reaches its maximum value which is independent of the initial conditions and less dependent on the detuning parameter δ. A more detailed analysis can be carried out only by numerical simulation [66, 68]. These results are presented at the end of this section.
A major issue of Self-Amplified Spontaneous Emission FEL radiation is that the radiation bandwidth decreases as $1/\sqrt{\hat{z}}$ compared to the $1/\hat{z}$ reduction of the unamplified spontaneous undulator radiation or the low gain FEL (Eq. 2.87).
The dependence of the latter can be explained by the slippage, where each electron has emitted a radiation pulse of the length $L = N_U\lambda$ after N_U undulator periods. The superposition of all emissions will not change this characteristic length. In the frequency domain this length corresponds to a characteristic frequency width, which is proportional to the inverse of L.
The main difference of high gain FEL is that the emitted radiation pulse has further interaction with other electrons. Work is done by the radiation field to bunch the electrons at the radiation phase and the head of the emitted radiation pulse is weakened. The linear growth in the pulse length of the independently emitted pulses is inhibited and the frequency bandwidth drops slower than $1/\hat{z}$.
Due to the random nature of self-amplified spontaneous emission it is necessary to use a statistical approach [65]. Important quantities are the spectral and time correlation functions of first and second order.
The first order spectral correlation function is given by

$$
\begin{aligned}
S(\omega,\omega',\hat{z}) &= \frac{<\tilde{E}(\omega,\hat{z})\tilde{E}^*(\omega',\hat{z})>}{\sqrt{<|\tilde{E}(\omega,\hat{z})|^2><|\tilde{E}(\omega',\hat{z})|^2>}} \\
&\approx \exp\left[i\frac{\omega-\omega'}{2ck_U\rho}\hat{z}\right]\frac{<\tilde{I}(\omega)\hat{I}^*(\omega')>}{\sqrt{<|\tilde{I}(\omega)|^2><|\tilde{I}(\omega')|^2>}} \quad , \qquad (2.123)
\end{aligned}
$$

where the weak influence of the detuning factor δ can be neglected compared to the fast oscillating term $\exp[i(\omega/2ck_U\rho)\hat{z}]$
The temporal structure of the current is a sum of Dirac-functions $I(t) = -e\sum\delta(t-t_k)$, where t_k is the position of the kth electron. Within the small interval $t, t+dt$ the number of electrons, averaged over many independent bunches, can be expressed by a distribution function $NF(t)dt$. This distribution is related to the current and its Fourier Transform by

$$
\begin{aligned}
<I(t)> &= -eNF(t) \quad , \\
<\tilde{I}(\omega)> &= <-e\sum e^{i\omega t_j}> = -eN\int_{-\infty}^{\infty}F(t)e^{i\omega t}dt = -eN\tilde{F}(\omega) \quad .
\end{aligned}
$$

The calculation of the first order spectral correlation function is straight forward and gives

$$
\frac{<I(\tilde{\omega})\tilde{I}^*(\omega')>}{\sqrt{<|\tilde{I}(\omega)|^2><|\tilde{I}(\omega')|^2>}} = \tilde{F}(\omega-\omega') + (N-1)\tilde{F}(\omega)F^*(\omega') \qquad (2.124)
$$

The last term correspond to the enhanced coherent emission based on the longitudinal profile. Nevertheless it is negligible for the frequencies in the narrow bandwidth of the FEL amplification, because the resonant frequency ω_0 is large compared to the characteristic frequency of the bunch profile. As an example for a Gaussian bunch with an energy of 20 MeV and a root-mean-square length of 1 mm ($\lambda_0 \approx 0.1$ mm) the coherent enhancement is suppressed by the factor $\exp(-100)$. Even for a high current beam $N \approx 10^{11}$ this term is small compared to $\tilde{F}(0)$.
The spectral correlation function of the electric field simplifies to $S(\omega,\omega',z) \equiv S(\omega-\omega',\hat{z}) = \exp[i(\omega-\omega')\hat{z}]\tilde{F}(\omega-\omega')$.
A measure of correlation of the radiation power for different frequencies is the spectral coherence $\Delta\omega_c$ [69] and is defined by

$$
(\Delta\omega_c)^2 = \frac{\int_0^{\infty}\omega^2|S(\omega,\hat{z})|^2d\omega}{\int_0^{\infty}|S(\omega,\hat{z})|^2d\omega} \quad . \qquad (2.125)
$$

Unfortunately the integration of Eq. 2.125 does not converge for all profiles, in particular for the uniform rectangular profile used in this discussion. To avoid this problem a different definition of the spectral coherence $\Delta\omega_c$ is often used [65], where the denominator of Eq. 2.125 alone is chosen. Despite this inconsistency both definitions lead to a dependence of the spectral coherence inversely proportional to the bunch length.

The second order correlation function

$$S_2(\omega,\omega') = \frac{<|\tilde{E}(\omega,\hat{z})|^2|\tilde{E}(\omega',\hat{z})|^2>}{<|\tilde{E}(\omega,\hat{z})|^2><|\tilde{E}(\omega',\hat{z})|^2>} = \frac{<|\tilde{I}(\omega)|^2|\tilde{I}(\omega')|^2>}{<|\tilde{I}(\omega)|^2><|\tilde{I}(\omega')|^2>} \tag{2.126}$$

is calculated in an analogous way as S. Keeping the leading terms the second order correlation function is related to the first order one by the identity $S_2(\omega-\omega') = 1+|S(\omega-\omega',\hat{z})|^2$. This relation is only valid if the positions of the electrons are purely random. This is assumed for the SASE FEL.

Based on these two spectral correlation functions the energy fluctuation of the full radiation pulse can be calculated.

The total energy of one single radiation pulse is

$$W = \frac{\Sigma}{Z_0}\int_{-\infty}^{\infty}|E(t,\hat{z})|^2 dt = \frac{2\Sigma}{Z_0}\int_0^{\infty}|\tilde{E}(\omega,\hat{z})|^2 d\omega \quad , \tag{2.127}$$

where Σ is the cross section of the radiation pulse and $Z_0 = \sqrt{\mu_0/\epsilon_0} \approx 377\Omega$ is the vacuum impedance.

The identity of the integration in the time and frequency domain is known as Parseval's theorem [70], where the integration over negative frequencies is the same as for positive frequencies due to the relation $\tilde{E}(-\omega,\hat{z}) = \tilde{E}^*(\omega,\hat{z})$ and can be replaced by weighting the integration over $[0,\infty]$ by the factor 2 in Eq. 2.127.

Averaging over many bunches the mean energy is

$$<W> = \frac{2\Sigma}{Z_0}\int_0^{\infty} <|\tilde{E}(\omega,\hat{z})|^2> d\omega \quad . \tag{2.128}$$

The evaluation of the standard deviation is proportional to the integration over the second order spectral correlation function. Normalizing with the mean value $<W>$ and replacing S_2 by the identity $S_2 = 1+|S|^2$ the variance is given by

$$\begin{aligned}\sigma_W^2 &= \frac{\int_0^{\infty}\int_0^{\infty}<|\tilde{E}(\omega,\hat{z})|^2|\tilde{E}(\omega',\hat{z})|^2> d\omega d\omega'}{\left[\int_0^{\infty}<|\tilde{E}(\omega,\hat{z})|^2> d\omega\right]^2} - 1 \\ &= \frac{\int_0^{\infty}\int_0^{\infty}<|\tilde{E}(\omega,\hat{z})|^2><|\tilde{E}(\omega',\hat{z})|^2>|S(\omega-\omega')|^2 d\omega d\omega'}{\left[\int_0^{\infty}<|\tilde{E}(\omega,\hat{z})|^2> d\omega\right]^2}\end{aligned} \tag{2.129}$$

As derived before the first order correlation function S depends only on the difference of the frequency and is independent of $\hat{z}$ if the absolute value is taken.

For the evaluation of σ_W two cases are considered.

The first case is when the bandwidth of the radiation is much smaller than the spectral coherence $\Delta\omega_c$. Due to definition of the spectral coherence (Eq. 2.125), which uses the simplifications that the first order spectral correlation function S depends only on the frequency difference,

the analytic model is only valid for $\sigma_A \gg \Delta\omega_c$. To achieve a narrow bandwidth in this analytic model the radiation is filtered by a monochromator with a frequency bandwidth σ_M and a central frequency ω_M. The Green's function of the monochromator is approximated by $H_M(\omega) \propto \exp[-(\omega - \omega_M)^2/2\sigma_M^2]$ and multiplied to the Green's function of the FEL amplification in Eq. 2.122. With $\sigma_M < \Delta\omega_c$ the integration of Eq. 2.129 has only a non-negligible argument for $\omega = \omega_M$ and $\omega' = \omega_M$. The fluctuation of the energy is $\sigma_W \approx |S(0)|^2 = 1$.
In the opposite limit for an FEL bandwidth much larger than the spectral coherence and without a monochromator the correlation function narrows the integration to $\omega \approx \omega'$ similar to a Dirac-function. The fluctuation of the radiation energy σ_W is $\sqrt{\pi}/\sigma_A T$ for a rectangular bunch profile with the pulse length T.
Based on the assumption that the amplified spontaneous radiation is completely random the probability of the radiation energy can be approximated by the gamma distribution

$$p(W) = \frac{M^M}{\Gamma(M)} \frac{W^{M-1}}{< W >^M} \exp\left(-M \frac{W}{< W >}\right) \quad , \tag{2.130}$$

with $M = \sigma_W^{-2}$ and $\Gamma(M)$ as the gamma function. The gamma distribution is identical with a negative exponential distribution if the independent parameter M is equal to unity.
An obvious representation of M is the number of spikes in the power spectrum of the radiation. If the SASE spectrum is filtered by a monochromator with $\sigma_M < \Delta\omega_c$ the spectrum exhibits only one spike and the most likely radiation energy is zero according to a negative exponential distribution ($M = 1$). This is in agreement with the probability of the spectral power (Eq. 2.122), because the width of a single spike is about $\Delta\omega_c$ and therefore almost coherent.
Without a monochromator the spectrum has many spikes ($M \gg 1$) The integration over the spectrum is comparable to taking many random samples of a single spike. The distribution peaks at the mean value of the negative exponential distribution and the relative root-mean-square width narrows as $1/\sqrt{M}$. In the limit $M \to \infty$ the probability distribution becomes a Gaussian distribution
For the analysis in the time domain the inverse Fourier transformation is applied to Eq. 2.120 yielding

$$E(t, \hat{z}) = \sqrt{\frac{8\pi}{\sqrt{3}\hat{z}}} \frac{e\omega_0}{I_0} \rho E_0 \exp\left[\frac{i + \sqrt{3}}{2}\hat{z}\right] \sum_k \exp\left[-\sigma_A^2 (t - t_z - t_k)^2 + i\omega_0 t_k\right] \quad , \tag{2.131}$$

with

$$t_z = \frac{z}{c}\left(1 + \frac{1 + K^2}{3\gamma_0^2}\right) \quad . \tag{2.132}$$

Eq. 2.131 inhibits a special feature of SASE radiation. Beside the growth of the amplitude, which depends only on $\hat{z}$, the radiation profile moves with a velocity slower than the speed of light but faster than the electron bunch. The FEL amplification depends on the frequency ω and thus the radiation field propagation is dispersive. To compare the result of Eq. 2.131 the group velocity of the radiation pulse is calculated. Combining the dominant oscillating term kz

of the radiation field and the slow phase variation of the amplitude A (Eq. 2.120) the artificial wavenumber $\tilde{k}$ is

$$\hat{k}(\omega) = \frac{\omega}{c} + \left[\frac{1}{2} - \frac{2}{3}\delta + \frac{1}{18}\delta^2\right] 2\rho k_U \tag{2.133}$$

The group velocity v_G [43] is given by

$$\frac{1}{v_G} = \left.\frac{d\hat{k}(\omega)}{d\omega}\right|_{\omega=\omega_0} = \frac{1}{c}\left(1 + \frac{1+K^2}{3\gamma_0^2}\right) \quad . \tag{2.134}$$

In the moving frame of the group velocity ($z = v_G t$) t and t_z in Eq. 2.131 cancel each other.
In most cases the width of the Gaussian distribution in Eq. 2.131 is large ($\propto 1/\rho\omega_0$) compared to the complex phase $i\omega_0 t$. Thus the evaluation of the sum over a single period of ω_0 is dominated by the random phases $\exp(i\omega_0 t_k)$ while the modulation of the amplitude is weak. The statistic of the radiation power $\propto |E(t,\hat{z})|^2$ is in this approximation identical with the spectral power and follows a negative exponential distribution.
The first order correlation function in time

$$\Gamma(t,t',\hat{z}) = \frac{< E(t)E^*(t') >}{\sqrt{< |E(t)|^2 >< |E(t')^2 >}} \tag{2.135}$$

can be related in the limit of a long electron bunch ($\Delta\omega_c \ll \sigma_A$) to the spectral power. Some lengthly but straight forward calculation gives

$$\Gamma(t-t') = \frac{\int_{-\infty}^{\infty} < |E(\omega)|^2 > e^{-i(\omega-\omega_0)(t-t')} d\omega}{\int_{-\infty}^{\infty} < |E(\omega)|^2 > d\omega} = \exp\left[-\frac{\sigma_A^2(t-t')^2}{2}\right] \tag{2.136}$$

for the rectangular bunch profile, where the average spectral power has been replaced by Eq. 2.122 to obtain the last identity.
Similar to the spectral coherence the coherence time τ_c is defined as the root-mean-square value of $|\Gamma(t-t')|^2$. In the model of the high gain exponential regime, discussed here, the coherence time is

$$\tau_c = \frac{1}{\sqrt{2}\sigma_A} \quad . \tag{2.137}$$

While the width of the spikes in the power spectrum is given by the pulse length, the spikes in the time domain are related to the coherence time. Due to the combined effects of slippage and amplification the width grows as $\sqrt{\hat{z}}$ and the total number of spikes decrease. In the frequency domain the root-mean-square width of the spectrum σ_A narrows down also decreasing the number of spikes.
The fluctuation of the energy follows in good approximation a gamma distribution and an explicit calculation shows that the parameter M tends to be unity for an integration time $\Delta T < \sqrt{2\pi}\tau_c$. With the interpretation that M is the number of spikes the explicit length of one spike is typically $\sqrt{2\pi}\tau_c$. Integrating the radiation power over the full length $T \gg \tau_c$ of an

uniform bunch gives the maximum number of spikes $\sqrt{\pi}/\sigma_A T$, being identical with the number of spikes in the spectrum.
All these results of this chapter have been derived for the exponential regime. The statistical behavior changes in the saturation regime. As in the case of the FEL amplifier this regime can only be analyzed by numerical simulations using macro particles. The explicit method to perform time-dependent simulation, such as a SASE FEL, is discussed in detail in Chapter 3. For this chapter a 1D steady state code, based on normalized variables, has been extended to cover longitudinal variation of the radiation field or bunching factor.
Based on the parameter set of the 100 nm SASE FEL of the TESLA Test Facility, 120 independent runs have been performed.

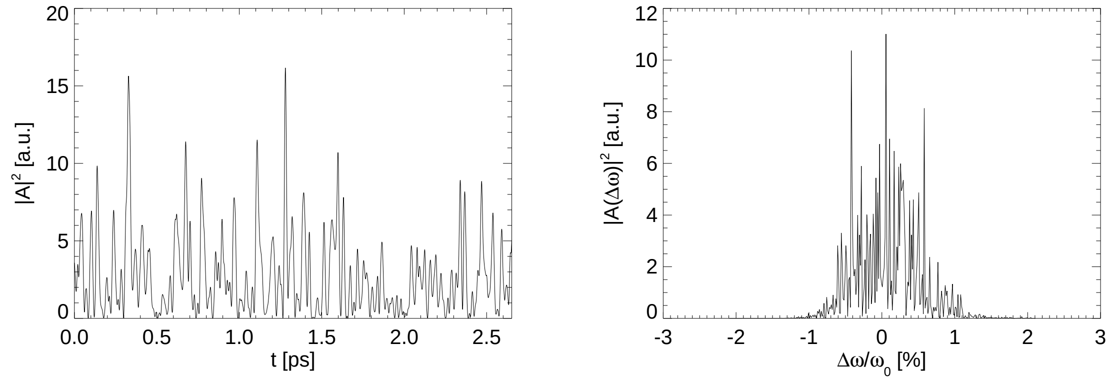


Figure 2.13: *Radiation pulse and spectrum of* $|A|^2$ *at* $\hat{z} = 10$.

Fig. 2.13 shows a typical radiation pulse profile and spectrum in the linear regime ($\hat{z} = 10$). For simplicity an uniform electron bunch profile has been assumed, because the obtained results do not claim to be a complete parameter study of the SASE-FEL of the TESLA Test Facility, which would be beyond the scope of this chapter.
The mean value of the normalized radiation power $|A|^2$ is plotted in Fig. 2.14. In comparison to an FEL amplifier the radiation power grows further after saturation at $\hat{z} = 18$. The SASE FEL provides additional amplification called superradiance [71].
Compared to the exponential growth of the radiation field, the gain of the superradiance regime is rather poor. The limitation of an FEL amplifier is given by a maximum bunching of the electron beam and an energy spread which fills out the full bandwidth of the amplification (see Section 2.5). Superradiance differs in that point that due to slippage the superradiant spike covers new regions of the electron beam, which have not provided much amplification so far and which parameters are still close to the initial conditions. The superradiant spike can further be amplified.
Reaching saturation the statistics of the radiation power is changed. The transition from the linear regime to saturation cuts the tail of the negative exponential distribution because spikes, corresponding to those energies, saturate first. The lower part of the distribution is not inhibited by saturation and shifts due to amplification towards larger energy. In Fig. 2.15 the probability

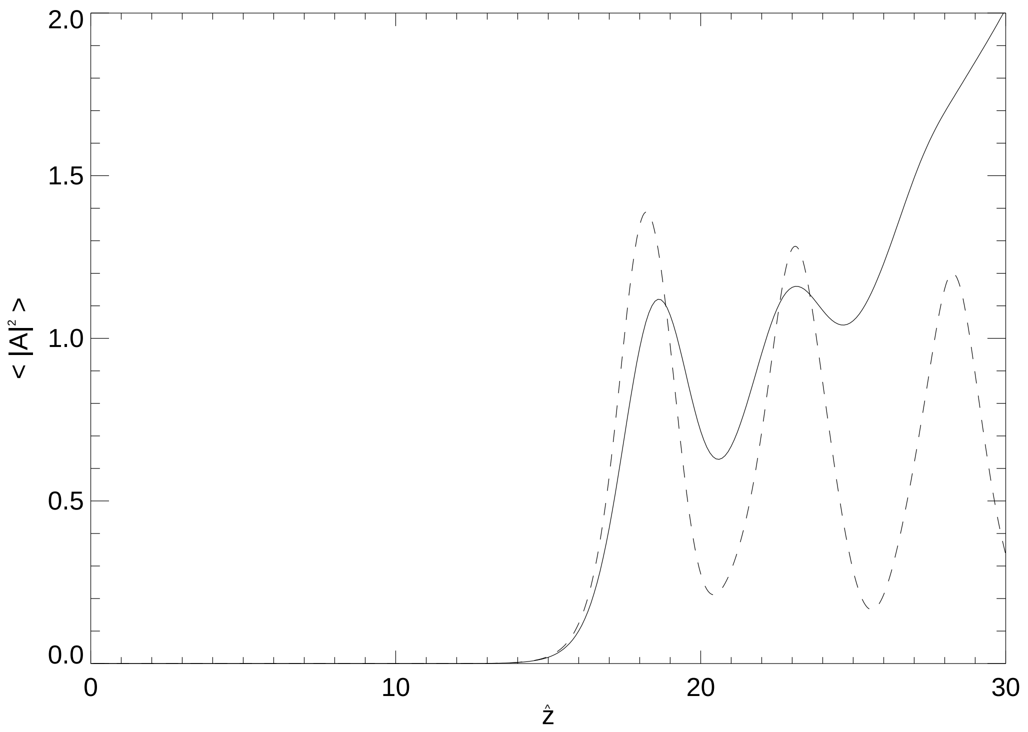


Figure 2.14: *Normalized radiation power $|A|^2$ versus normalized position $\hat{z}$ of a SASE FEL, averaged over the bunch length. For comparison the results for an FEL amplifier are drawn by a dashed line.*

distribution for the linear regime ($\hat{z} = 10$) and saturation ($\hat{z} = 18$) is plotted, left and right respectively. It is an important fact that at saturation the radiation power distribution deviates from a negative exponential.

In the transition from saturation to superradiance the probability distribution turns almost back to a negative exponential. The fluctuation of the relative radiation power versus normalized position $\hat{z}$ is plotted in Fig. 2.16 (left). A variation of unity corresponds to a negative exponential distribution. In the superradiance regime the distribution seems to be stable with no or only a weak dependence on $\hat{z}$ as in the linear regime, although the probability of no or small radiation power is enhanced compared to the negative exponential distribution in the linear regime. The reason is that the superradiant spikes, still being amplified, become shorter. This is reflected by the degradation of the coherence time, clearly visible in the right plot of Fig. 2.16 for $\hat{z} > 25$. The region of small radiation field between two spikes gets wider and thus the probability of no or small radiation power is enhanced.

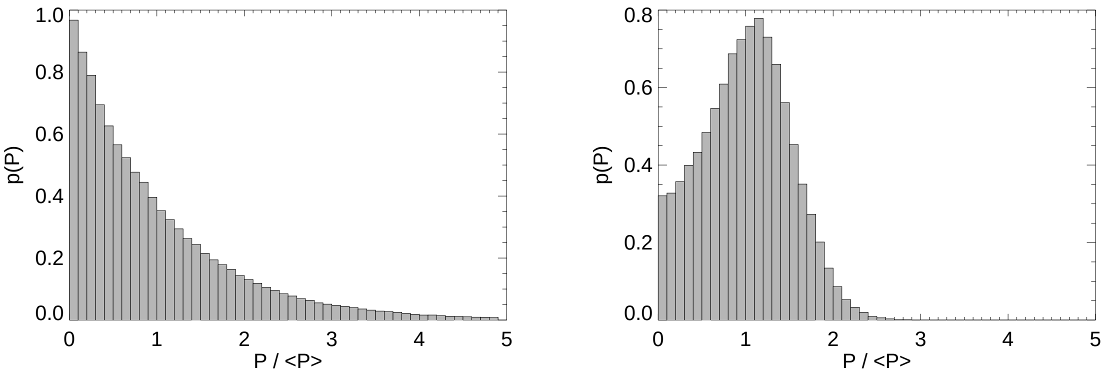


Figure 2.15: *Radiation power distribution in the linear regime ($\hat{z} = 10$) and saturation ($\hat{z} = 18$).*

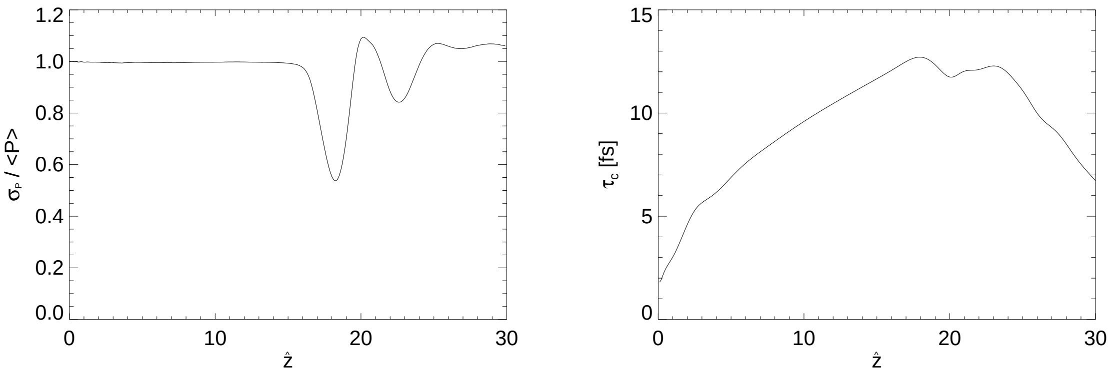


Figure 2.16: *Power fluctuation (left) and coherence time of the radiation pulse (right) versus normalized position $\hat{z}$.*

Chapter 3

Numerical Methods for FEL Simulations

In Chapter 2 a set of self-consistent equations has been derived, based on the fundamental Maxwell's and Hamilton equations which describe the physics of a Free-Electron Laser (FEL). These coupled differential equations, where the radiation and electrostatic field equations are expressed as partial differential equations, cannot be solved analytically. Therefore various simulation codes have been written to obtain the desired results numerically [22, 23, 34, 73, 74, 75, 76, 77, 78]. These codes are more or less optimized for certain aspects of a particular problem. During the development phase the performance of the computers at that time was – and still is – a major issue for the structure and capacity of these codes. With the on-going improvement of computational power these programs became more complex and were able to use less approximations of the basic FEL equations [79].

These codes are classified by three criteria, which give a rough indication of the underlying model.

The first criterion is given by the description of the electron beam parameters, either by collective variables such as the bunching factor or by macro particles, representing an ensemble of electrons with the same properties. Although a code, based on collective variables, is significantly faster, it works correctly only in the linear regime of the Free-Electron Laser. The saturation regime is connected with a rather complicated distribution in longitudinal phase space, where a few collective variables are not sufficient for an adequate representation.

Another important issue is the model of the radiation field. The difference between a one dimensional and a two dimensional model is significant, because the assumption of an infinite extension of the radiation field (1D code) ignores the important effect of diffraction of the radiation field. The simplest model, including diffraction, assumes radial symmetry. This 2D model can be extended by a Fourier decomposition of the radiation field and the electron beam in the azimuthal angle to cover radially symmetric problems like an elliptical beam cross section. This decomposition becomes inefficient if the beam distribution is irregular such as beam halos. A lot of higher Fourier components have to be included to describe the beam profile sufficiently, which results in a time consuming calculation of the Fourier transformation. A fully 3D code, based on the Cartesian coordinate system for the electron motion and radiation

field, avoids this problem. Although the number of parameters to describe the radiation field is much larger in a 3D code the execution time can be comparable to or faster than that of a 2D code. Because memory demands for 3D codes are not a critical issue for modern computers, they should be preferred due to two reasons. No conversion and mode decomposition from the Cartesian system of the electrons to the cylindrical system of the radiation field is needed and the partial differential equations have no singularity as in cylindrical coordinates (Section 3.4).
The last criterion refers to the longitudinal dependence of the electron beam and radiation field. Steady-state codes assume a periodicity of all parameters and need to define them only within one radiation wavelength. They are much faster than time-dependent codes but lack the ability to describe the radiation field slippage correctly. Time-dependent codes are unavoidable if a Self-Amplified Spontaneous Emission Free-Electron Laser (SASE FEL) is simulated.
This chapter presents several numerical methods to solve the set of self-consistent FEL equations (Section 3.1), which have been incorporated into various existing codes. All these methods are guided by the idea of a highly efficient numerical code based on the fewest possible approximations and assumptions.
Any FEL code has to solve four major problems:

1. Generating the initial phase space distribution of the electron beam (Section 3.2),
2. Solving ordinary differential equations of the electron beam variables (Section 3.3),
3. Solving partial differential equations of the radiation and electrostatic field (Section 3.4),
4. Bookkeeping of the radiation field and electron beam parameters and efficient use of the computer resources for time-dependent simulation (Section 3.5).

If the simulation runs under certain assumptions, as it is mentioned above, some of these problems may not occur. For one dimensional steady-state FEL simulations, operating in the linear regime, only the second problem remains, namely to solve an ordinary third order differential equation.

3.1 The Underlying Differential Equations

For convenience a coordinate system is chosen, which moves with the electron bunch along the undulator axis with the mean velocity $c\beta_0$. In this frame the transverse variables are x, y, p_x and p_y. For simplicity the transverse momenta have been normalized to mc, where m is the electron mass and c is the speed of light. The longitudinal position has been transformed to the electron phase $\theta = (k + k_U)z - kct$ in the ponderomotive wave with k as the wavenumber of the radiation field and k_U as the undulator wavenumber. The corresponding canonical momentum is γ, the electron energy normalized to the rest energy mc^2. The independent variable is the longitudinal position z within the undulator.
The FEL equations (see Section 2.3) for a single electron are

$$
\begin{aligned}
\theta' &= \frac{k_U}{\beta_0} - k\frac{1+p_x^2+p_y^2+K^2+f_cK(ue^{i\theta}-\text{c.c.})}{2\beta_0\gamma^2} \quad , &(3.1)\\
\gamma' &= -k\frac{f_cK}{2\beta_0\gamma}(ue^{i\theta}+\text{c.c.}) + \frac{e}{mc^2}\sum_{l=1}^{\infty}[\tilde{E}_l e^{il\theta}+c.c] \quad , &(3.2)\\
x' &= \frac{p_x}{\gamma} \quad , &(3.3)\\
y' &= \frac{p_y}{\gamma} \quad , &(3.4)\\
p_x' &= -\left(\frac{k_x^2K^2}{\gamma}+\frac{eg}{mc}\right)x \quad , &(3.5)\\
p_y' &= -\left(\frac{k_y^2K^2}{\gamma}-\frac{eg}{mc}\right)y \quad , &(3.6)
\end{aligned}
$$

with u as the complex amplitude of the radiation field, $\tilde{E}_l$ as the Fourier components of the longitudinal electrostatic field, f_c as the coupling factor and g as the gradient of a superimposed quadrupole field. The derivatives are taken with respect to z. The dimensionless undulator field amplitude K exhibits a transverse dependence up to second order in x and y as

$$
K(x,y,z) = K_0(z)\left(1+\frac{k_x^2}{2}x^2+\frac{k_y^2}{2}y^2\right) \quad , \tag{3.7}
$$

where $K_0 = e\hat{B}_0/mck_U$ is called the undulator parameter with $\hat{B}_0$ as the root-mean-square magnetic field on the undulator axis ($x=0$, $y=0$). The values of k_x and k_y depend on the undulator type and the curvature of the pole faces. For a helical undulator the squares of both values are identical with $k_U^2/2$, while in the case of a planar undulator the values depend on the explicit curvature of the pole faces fulfilling the constraint $k_x^2+k_y^2=k_U^2$. Additional transverse focusing is supplied by an external quadrupole field gradient g [30]. The undulator amplitude K_0 may also depend on z if drift sections of field tapering are included in the simulations.
A longitudinal oscillation of the electron exists only for a planar undulator. This desynchronisation of the electron position with the radiation field is covered by the coupling factor f_c, defined for the fundamental frequency as

$$
f_c = \begin{cases} J_0\left(\frac{K_0^2}{2(1+K_0^2)}\right) - J_1\left(\frac{K_0^2}{2(1+K_0^2)}\right) & \text{planar undulator} \\ 1 & \text{helical undulator} \end{cases} \quad . \tag{3.8}
$$

Analogous to the undulator field the dimensionless but complex value of the radiation field amplitude is given by

$$
u(x,y,t,z) = \frac{e\hat{E}(x,y,t,z)}{imc^2k}e^{i\Psi(x,y,t,z)}. \tag{3.9}
$$

The dependence of the root-mean-square amplitude $\hat{E}$ and the phase Ψ on x and y defines the transverse profile, while the variables t and z describe the change in the radiation field amplitude along the radiation pulse and the undulator, respectively.

The longitudinal electrostatic field is expanded into a Fourier series of the phase of the ponderomotive wave. The amplitudes for negative frequencies ($\hat{l} < 0$) can be calculated by the complex conjugate coefficients of the corresponding positive frequencies, using the identity $\tilde{E}^*_{-\hat{l}} = \tilde{E}_{\hat{l}}$. The non-resonant constant term $\tilde{E}_0$ can be neglected.
The paraxial approximation of Maxwell's equation for the radiation field (Section 2.3) is

$$\left[\vec{\nabla}_\perp^2 + 2ik\left(\frac{\partial}{\partial z} + \frac{\partial}{c\partial t}\right)\right] u = \frac{ie^2\mu_0}{m}\sum_j \delta(\vec{r} - \vec{r}_j)\frac{f_c K}{\gamma_j}e^{-i\theta_j} \quad , \tag{3.10}$$

where μ_0 is the magnetic permeability.
The expansion of the electrostatic field into a Fourier series modifies Maxwell's equation of the static problem to

$$\left[\nabla_\perp^2 - \frac{\hat{l}^2 k^2(1+K^2)}{\gamma_R^2}\right]\tilde{E}_{\hat{l}} = i\frac{ec^2\mu_0 \hat{l} k(1+K^2)}{\gamma_R^2}\sum_j \delta(\vec{r} - \vec{r}_j)e^{-i\hat{l}\theta_j} \tag{3.11}$$

The resonant energy is defined as $\gamma_R = \sqrt{k(1+K_0^2)/2k_u}$.
The underlying approximations, which determine the applicable range of the simulation, are rather modest and demand

- a relativistic electron beam,
- a small transverse beam size, compared to the undulator period length,
- a variation of the radiation field amplitude u slower than the dominant oscillation $\exp[ik(z-ct)]$, yielding the paraxial representation of Maxwell's equation, and
- a resonant interaction between electron beam and radiation field.

The last two demands can be combined to the constraint $\rho \ll 1$ for the FEL parameter (Eq. 2.69).
These FEL-equations differ slightly from those in Chapter 2. In the frame of normalized variables the normalized amplitude A (Eq. 2.67) reaches saturation at about unity and is thus comparable to K_0. According to the normalization, the value of u is proportional to ρ^2. For a relativistic electron beam and a typical undulator the FEL parameter ρ is smaller than unity. The source term proportional to u in Eq. 2.57 as well as the term proportional to $|u|^2$ in Eq. 2.59 are overshadowed by those linear in K and can be omitted.
The equations Eqs. 3.1 – 3.6, 3.10 and 3.11 can be simplified to reduce the amount of calculation. Approximations are made, which are typically

- the description of the electron bunch by a few collective variables such as the bunching factor or energy spread (linear model),
- the extrapolation of the radiation field and bunch profile in longitudinal direction (steady-state model),

- the assumption of a symmetric transverse profiles of radiation field and electron beam (1D or 2D models).

Although these codes take into account major aspects of the Free-Electron Lasers they cannot cover them all. In particular they are not quite sufficient for the final design of an FEL, where details have to be studied, or as an aid in the analysis and interpretation of measured data.
Modern computers provide the opportunity to simulate more complicated problems. For steady-state simulations, where the longitudinal variation of all parameters is ignored, the remaining self-consistent FEL equations can be solved numerically within minutes or less. Only a full time-dependent simulation would be beyond the capability of present computers due to the large amount of memory needed. Under rather less restricting assumptions the simulation can be split into several subprocesses with a moderate memory consumption as it is described in Section 3.5.

3.2 Generating the Electron Phase Space Distribution

Due to the large number of electrons per bunch, in the range of typically 10^9 – 10^{11} electrons, modern computers are still unable to simulate all electrons independently. Therefore all electrons are either represented by sample particles, called macro particles, to reduce the memory demands by several orders of magnitude or by collective variables such as mean energy, energy spread, root-mean-square beam size, bunching factor and any correlation between these variables. To cover the complete FEL amplification process, from the start-up regime to saturation, macro particles should be used. An alternative method would include higher moments of the 6D phase space distribution in the FEL equations, but the high number of moments needed makes this approach impractical for non-linear simulations.
Out of all phase space coordinates the longitudinal coordinates, the electron phase θ and the energy γ, are the most critical ones, because the FEL amplification is driven by the correlation of these two variables. If the initial values of the electron phases θ are calculated by a random number generator, the resulting distribution is not suitable for FEL simulation, regardless how excellent the generator is. For typically 10^4 macro particles per ponderomotive bucket the average bunching factor b has a root-mean-square amplitude of $\sqrt{<|b^2|>} = 0.01$, according to Eq. 2.117. This value is closer to the situation at saturation than to typical initial conditions. The same argument is valid for the energy because any random distribution in γ yields an undetermined variation in the electron phases over time and thus a fluctuation in the bunching factor. The amplitude of this fluctuation is much too large due to the insufficient number of macro particles. A simulation is less sensitive to the distributions of the transverse variables x, y, p_x and p_y. Randomly generated distributions might be used in this case although more advanced methods exist (see below).
This problem is solved by ‘mirroring’ a macro particle by another macro particle at the different phase $\theta_{j+1} = \theta_j + \pi$. Loading half of the macro particles between $-\pi < \theta < 0$ and copying them into the remaining phase range while keeping the other coordinates constant yields a bunching factor solely given by the numerical precision of the computer. Unfortunately the source term of the electrostatic field (Eq. 3.11) would peak at the second harmonics of θ. To avoid this

unphysical high field a macro particle is loaded between $-\pi < \theta < -\pi/2$ and then mirrored three times [38] with phase offsets of $\pi/2$, π and $3\pi/2$. The expansion of the electrostatic field up to the third harmonic in the ponderomotive phase is more than sufficient. If it is intended to extend the code to include higher harmonics of the resonant radiation wavelength the number of mirror particles has to be larger than the highest harmonic that the simulation can resolve.
The problem of filling the 6D phase space of the macro particles can be reduced to the problem of filling a 6D unit-cube homogeneously. For any projection of the 6D phase space on a subspace, in particular to the one dimensional subspace, the particle distribution must be uniform. The problem is solved by a sequence f_n, which generates this one dimensional uniform distribution. The values of f_n should lie within zero and one, including the upper and excluding the lower boundary (semi-open area). This restriction is convenient for further transformation of the distribution.
In general more than one possible sequence exists. The simplest type is the symmetric loading f_n^S. The 6D cube is superimposed on a lattice and a macro particle is assigned to each grid point. For the 1D case the sequence is given by

$$f_n^S = \frac{n}{N} \qquad 1 \leq n \leq N \quad , \tag{3.12}$$

where N is the total number of macro particles. Although this method is often used for codes in plasma physics [80], it is only partially useful for FEL simulation. Regarding a 6D cube, which is mirrored three times in phase, even a coarse lattice with 10 grid points in each dimension requires 4 million macro particles. The actual problem is that 10 grid points per dimension artificially excite higher modes (see Section 2.6) because the transverse beam profile is not smooth.
An alternative method is to fill each dimension independently, where the different sequences for each dimension must be completely uncorrelated. For this purpose a random number generator f_n^R is sufficient. Most of the random number generators are based on a recursive algorithm to ensure no correlation between the sequence values [81]. All random number generators have in common that the initial value f_0^R is defined by an argument, supplied at the first call of f_n^R. This initialization of the random number generator is called 'seeding'.
Although the random loading of the phase space fulfills the requirements it has the disadvantage of the statistical fluctuation in the particle position. In the left plot of Fig. 3.1 a two dimensional unit-square is loaded with 1000 macro particles using a random number generator. As it can be seen the macro particles tend to form local clusters while other areas remain unpopulated. In addition succeeding values of the random number sequence are uncorrelated. This exceeds the goal of generating the phase space distribution in one dimension, because an explicit order of the macro particles is irrelevant. The simulation regards the distribution as a whole. The only important constraint is that, for any arbitrarily chosen order of the macro particles, the coordinates in the six dimensional phase space are not correlated with respect to each other.
Therefore a much more sophisticated method of phase space filling is the 'quasi-random' loading [82], which minimizes the formation of local clusters. The sequences used are called Halton or Hammersley sequences [83], which is a generalization of the 'bit-reverse' technique. The algorithm decodes the value of the sequence index n into the representation of the base b with

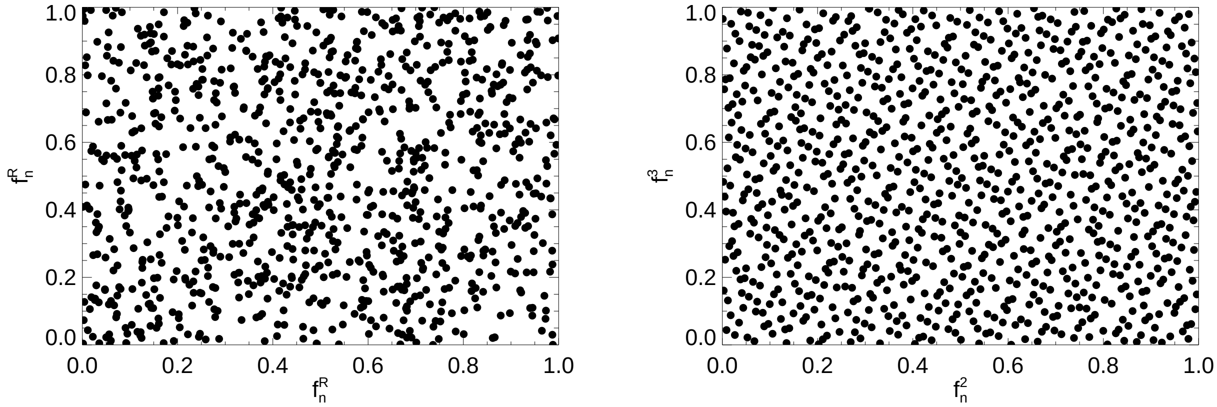


Figure 3.1: *Filling of the unit-square by two independent sequences, using a random number generator f_n^R with different seeding (left) or two Hammersley sequences f_n^b with the bases $b = 2$ and $b = 3$ (right).*

$$n = \sum_{j=0}^{\infty} a_j(n) b^j \quad .$$

The 'digits' a_j are always integer numbers with $0 \leq a_j(n) < b$. As an example the decimal number '17' is identical with $a_0 = 2$, $a_1 = 2$, $a_2 = 1$ and $a_j = 0$ for $j > 2$ for a system with the base $b = 3$. In analogy to the decimal system or the binary system ($b = 2$), commonly used in computer science, this number would be written as '122' for $b = 3$.
The Hammersley sequence for the base b is then given by

$$f_n^b = \sum_{j=0}^{\infty} a_j(n) b^{-(j+1)} \quad . \tag{3.13}$$

Fig. 3.1 shows the difference between loading the phase space randomly or based on Hammersley sequences. The latter loading produces a much smoother distribution, where the macro particles are almost equally spaced. The correlation between two Hammersley sequences with different bases is negligible unless one base is a multiple of the other or the chosen base is as large as the number of sample values. To avoid this it is advisable to use only the smallest prime numbers as bases.
The algorithm of the Hammersley sequence guarantees that the most important digit $a_0(n)$ is fast changing and sweeps over the entire range between zero and one within b cycles.
The final step is, once the unit-cube has been filled, to transform it to the desired distribution. Tilting and shifting the unit-cube in phase space are trivial and can be directly applied, while it is more difficult to change the shape of the distribution itself.
Using the identity

$$\int_0^1 dx = \int_{x^{-1}(0)}^{x^{-1}(1)} \left| \frac{dx(y)}{dy} \right| dy = \int_{y_0}^{y_1} p(y)dy \qquad (3.14)$$

the generating transformation for the distribution $p(y)$ is

$$y(x) = P^{-1}(x) \quad , \qquad (3.15)$$

where $P(y)$ is the indefinite integral of $p(y)$ and P^{-1} is the inverse function of P. As an example, the function $y = -\ln(x)$ transforms a uniform distribution into a negative exponential. Unfortunately the most common distributions such as Gaussian or parabolic distributions as well as a homogeneously distributed unit-sphere cannot be derived directly without integrating and inverting the distribution $p(y)$ numerically.

Certain distributions can be generated by joint probability distribution $p(x_1, x_2, \ldots, x_n)$, which is the generalization of Eq. 3.14 for n dimensions. The derivative $|dx(y)/dy|$ is replaced by the Jacobian determinant $|\partial(x_1, x_2, \ldots)/\partial(y_1, y_2, \ldots)|$ [59]. For independent random numbers y_j the transformed distribution $p(y_1, y_2, \ldots)$, and thus the Jacobian determinant, must be the product of single distributions $\hat{p}_i(y_i)$ with

$$p(y_1, y_2, \ldots, y_n) = \prod_{i=1}^{n} \hat{p}_i(y_i) \quad .$$

Using these methods by combining two uniform distributions, the Gaussian distribution and the uniform, two dimensional sphere, can be easily generated [84], both frequently used for FEL simulations. The generating transformations and the final distribution are

$$\begin{aligned} y_1 &= \sqrt{-2\ln x_1}\cos(2\pi x_2) \\ y_2 &= \sqrt{-2\ln x_1}\sin(2\pi x_2) \\ p(y_1, y_2) &= \left[\frac{1}{\sqrt{2\pi}} e^{-y_1^2/2}\right]\left[\frac{1}{\sqrt{2\pi}} e^{-y_2^2/2}\right] \end{aligned}$$

for the Gaussian distribution and

$$\begin{aligned} y_1 &= \sqrt{x_1}\cos(2\pi x_2) \\ y_2 &= \sqrt{x_1}\sin(2\pi x_2) \\ p(y_1, y_2) &= \frac{1}{\pi} \end{aligned}$$

for the homogeneously filled circle. A parabolic distribution can be obtained by projecting the macro particles to one axis of a uniformly filled 3D sphere.

The phase space filling is much simplified if the FEL code assumes no transverse betatron motion. The 2D subspace of the transverse momenta is omitted and the transverse beam

profile is modeled by weighting the charge of the macro particles. The momentum spread is indirectly expressed by an effective energy spread (Eq. 2.96).
The remaining part of this section concerns the simulation of the initial fluctuation of the electron positions, the driving source for a Self-Amplified Spontaneous Emission FEL. The goal is to generate a random distribution in the electron phase θ so that the variance of the bunching factor amplitude is inversely proportional to the total number N of electrons and thus agrees with the modeled electron beam. The phase of the bunching factor should be arbitrary, which is equivalent to a vanishing expectation value of the bunching factor over many bunches.
The algorithm starts with a preloaded distribution of n macro particles with a zero bunching factor as described above. Each macro particle has an initial phase θ_j. To generate a fluctuation a random phase offset is added to each particle. This offset is uniformly distributed within the limits $-\Theta < \Delta\theta_j < \Theta$. The expectation value and variance of $\exp[i(\theta_j + \Delta\theta_j)]$ for the jth macro particle are, over many samples,

$$\begin{aligned} < e^{-i\theta_j} > &= \frac{1}{2\Theta}\int_{\theta_j-\Theta}^{\theta_j+\Theta} e^{-i\theta'}d\theta' \\ &= -e^{-i\theta_j}\frac{\sin(\Theta)}{\Theta} \end{aligned} \tag{3.16}$$

and

$$< \Delta|e^{-i\theta_j}|^2 >= 1 - \frac{\sin^2(\Theta)}{\Theta^2} \quad , \tag{3.17}$$

respectively. Averaging over all macro particles the expectation value is zero. This agrees with the demand of an arbitrary phase per single sample. In contrast the variance of the bunching factor amplitude for the complete ensemble of macro particles remains larger than zero, being the result of a single macro particle (Eq. 3.17) divided by the number of macro particles n. The right choice for Θ matches the variance with

$$\frac{\sin^2(\Theta)}{\Theta^2} = 1 - \frac{n}{N} \tag{3.18}$$

to agree with the statistics of the simulated electron beam [85]. Normally the number of macro particles is much smaller than N and the random phase fluctuation is approximately $\Theta \approx \sqrt{3n/N}$. In the case that each macro particle represents only one single electron the random phase offset is within $\pm\pi/2$. Together with the mirror particle, which compensates the initial phase θ_j, the fluctuation of both phases uniformly cover the complete phase range between $-\pi$ and π. The positions are totally random as they should be in this case.
An alternative way of simulating the initial fluctuation of the particle position is the direct calculation of the phases [77]. Assuming that a group of 4 macro particles corresponds to $\tilde{N} = 4N/n$ electrons in the beam, a uniform random number generator provides two independent samples x_n and y_n. According to the statistic of the bunching factor, where the absolute square of the bunching factor b follows a negative exponential distribution and the phase a uniform distribution, a random sample of its amplitude and phase is obtained by $b_n = \sqrt{-\ln(x_n)/\tilde{N}}$

(see Eq. 3.15) and $\phi_n = 2\pi y_n$. If this sample of amplitude and phase is expressed in polar coordinates, the resulting vector can be constructed by 4 unit vectors, corresponding to the 4 macro particles. The polar angles are the resulting ponderomotive phases ϕ_n, $\phi_n - \varphi_n + \pi/2$, $\phi_n + \pi$ and $\phi_n + \varphi_n + 3\pi/2$ with $\varphi_n = \arcsin(2b_n)$ of the 4 macro particles.
Both methods are equivalent, as far as the correct statistic they produce is concerned, but differ slightly in their computational time, because the calculation of the arc sine is rather time consuming. Considering that most of the computational time for FEL simulation is used for solving the differential equations, this numerical criterion becomes insignificant to decide which algorithm should be chosen.

3.3 Coupled Differential Equations and the Integration of the Particle Equations of Motion

For an FEL simulation, using N macro particles, $6N + 2$ differential equations have to be solved, 6 ordinary differential equations for each macro particle and the two partial differential equations for the radiation and electrostatic field. The treatment of the partial differential equations is explicitly described in Section 3.4.
The equations are coupled, so that they cannot be solved independently. In particular the particle phases and energies affect the source term of the field equations and vice versa. The transverse betatron motion is almost independent except for the change in the particle energy. The efficiency of Free-Electron Lasers is roughly ρ (Eq. 2.92), where ρ is the FEL parameter. For high gain FELs of relativistic beams typical values are $\rho \ll 1$ and the overall energy change is small in particular in the linear regime. This allows at least the electron trajectory to be pre-calculated to a rather good approximation.
The general approach for FEL simulation, based on macro particles, is the separation of electron motion and radiation field propagation. The goal is to advance both one integration step of the length Δz. Because each of them forms the source term of the other differential equation(s) the integration might not be stable if field and particles are evaluated at the same position z. To exclude this instability the integration step length must be chosen sufficiently small and the overall effort to advance the particles one integration step is high. Advancing the particles first they would always see the 'old' field, while the source term of the field equation is given by the 'new' positions of the particles. The accuracy of the integration is improved by one order ($\mathcal{O}(\Delta z^2)$) when the source term is defined at the middle of each integration step. The result is an offset of $\Delta z/2$ between the particle variables and the radiation field. This alternating integration is the working principle of the 'leapfrog' method.
Fig. 3.2 shows a slightly modified leapfrog method, which is applied when the transverse betatron motion is calculated analytically. The integration of the radiation field and the longitudinal variables of the macro particles still reensemble the basic 'leapfrog'. The difference lies in pushing the transverse variables half an integration step (step 1 in Fig. 3.2), where the transverse positions and momenta are used to evaluate the source terms of Eq. 3.1 and 3.2. After the integration of the longitudinal variables (step 2) the transverse variables are advanced the remaining half of the integration step (step 3), using the new values for the energy. To

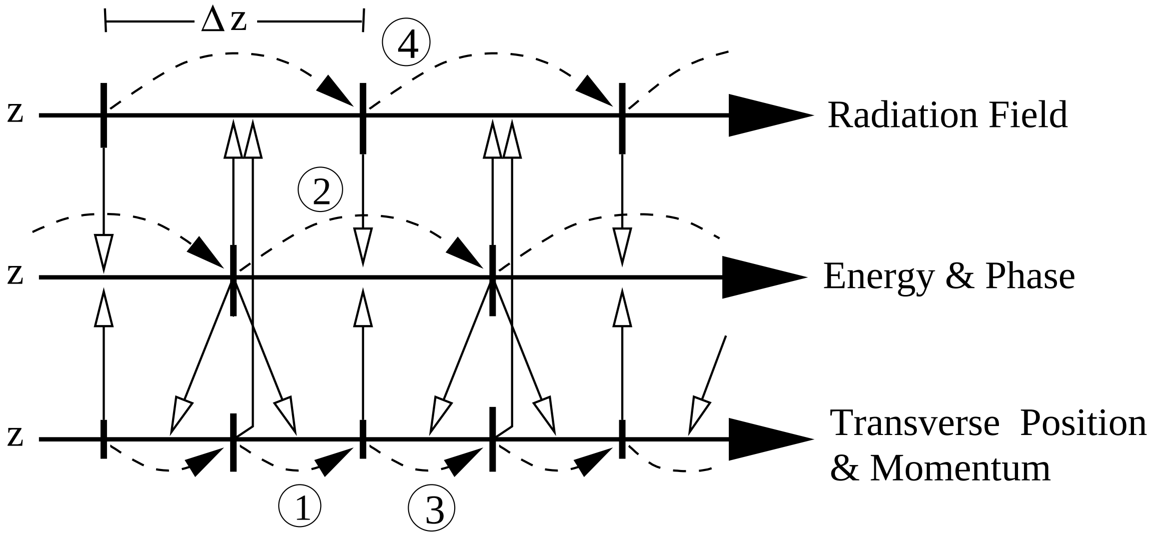


Figure 3.2: *Leapfrog integration scheme. The marks on the z-axis indicate the positions where the radiation field and electron variables are evaluated. A dashed arrow represents the integration step, a hollow arrow the usage of the variable at that position as a source term for a different integration. Four integration steps are needed to advance particles and field about Δz. The order of these steps are indicated by the numbers.*

complete the integration the field equation is solved (step 4) before the procedure is started over again with step 1.

Beside the general integration scheme, described above, the accuracy of the simulation depends critically on the method used to solve the differential equations. The best result are obtained, of course, if the differential equation is analytically solved and the solutions are matched to the initial conditions.

If the integration step Δz is chosen in such a way, that the transverse focusing strength in Eqs. 3.5 and 3.6, for either natural or strong focusing, can be assumed to be constant over the integration step, the differential equations Eqs. 3.3 – 3.6 have well known solutions [28]. The transverse variables are advanced by two simple matrix operations

$$\begin{aligned} \begin{pmatrix} x \\ p_x \end{pmatrix}_{n+1/2} &= M_x(\gamma_n, q_x(z_n)) \begin{pmatrix} x \\ p_x \end{pmatrix}_n \\ \begin{pmatrix} x \\ p_x \end{pmatrix}_{n+1} &= M_x(\gamma_{n+1}, q_x(z_n)) \begin{pmatrix} x \\ p_x \end{pmatrix}_{n+1/2} \end{aligned}$$

for the x-plane and similar for the y-plane, where the index n denotes the nth integration step. Due to the modified leapfrog method the 'updated' value of the energy γ_{n+1} is used for the second matrix multiplication. The 2×2 matrix M depends on the focusing strengths $q_x(z_n) = K^2k_x^2/\gamma_n^2 + q_0(z_n)/\gamma_n$ or $q_y(z_n) = K^2k_y^2/\gamma_n^2 - q_0(z_n)/\gamma_n$ for the x- and y-direction,

respectively. The opposite sign of the superimposed quadrupole field gradient q_0 in the definition of q_x and q_y accounts for the fact that any quadrupole focuses only in one plane, while it defocuses with the same strength in the perpendicular plane.
Three cases have to be considered for the transport matrix M_x with

$$M_x = \begin{pmatrix} \cos\left(\sqrt{q_x}\frac{\Delta z}{2}\right) & \frac{1}{\sqrt{q_x}}\sin\left(\sqrt{q_x}\frac{\Delta z}{2}\right) \\ -\sqrt{q_x}\sin\left(\sqrt{q_x}\frac{\Delta z}{2}\right) & \cos\left(\sqrt{q_x}\frac{\Delta z}{2}\right) \end{pmatrix} \tag{3.19}$$

for a focusing section ($q_x > 0$),

$$M_x = \begin{pmatrix} \cosh\left(\sqrt{q_x}\frac{\Delta z}{2}\right) & \frac{1}{\sqrt{q_x}}\sinh\left(\sqrt{q_x}\frac{\Delta z}{2}\right) \\ \sqrt{q_x}\sinh\left(\sqrt{q_x}\frac{\Delta z}{2}\right) & \cosh\left(\sqrt{q_x}\frac{\Delta z}{2}\right) \end{pmatrix} \tag{3.20}$$

for a defocusing section ($q_x < 0$) and

$$M_x = \begin{pmatrix} 1 & \frac{\Delta z}{2\gamma} \\ 0 & 1 \end{pmatrix} \tag{3.21}$$

for a drift section ($q_x = 0$).
Replacing q_x with q_y yields the solution for the y-plane. The integration step size is half as long as the step size for the radiation field or longitudinal variables of the macro particles, because the transverse variables are advanced twice per complete 'leapfrog' integration step.
A special problem is the influence of undulator field errors on the electron trajectory. Because the field strength varies for each pole the phase advance of the transverse oscillation per half undulator period is not exactly π and the angle of the trajectory relative to the undulator axis does not match the following half period of the oscillation. Under the constraint that the step size Δz is restricted to $\Delta z = \lambda_U/2$ the change in the electron trajectory is modeled by a random kick Δp, added to the transverse momentum of the macro particle.
With the continuity condition of the transverse momentum at the nth integration step

$$p_x(z_n) = -[\sqrt{2}]K_n\cos(nk_u\Delta z) = -[\sqrt{2}]K_{n+1}\cos(nk_u\Delta z) + \Delta p_x$$

and $\Delta z = \lambda_u/2$ the change of the momentum has the magnitude

$$\Delta p_x = (-1)^n[\sqrt{2}](K_n - K_{n+1}) \quad . \tag{3.22}$$

The factor $\sqrt{2}$ in the square brackets is only applied for the case of the planar undulator because the definition of the undulator field amplitude K is based on its root-mean-square value (Eq. 2.11). As long as the change in the magnetic field is small this momentum kick can be treated as a perturbation and can be added directly to the momentum right after each integration step.

Unlike the transverse motion the longitudinal differential equations Eqs. 3.1 and 3.2 for the electron energy and phase have no general analytic solution and must be integrated numerically. For any ordinary, first order differential equation $y' = f(z, y)$ the simplest method to advance the variable y_n is the linear extrapolation $y_{n+1} = y_n + f(z_n, y_n)\Delta z$. The major draw-back is the extremely poor accuracy of this first order method because it does not follow the curvature of the analytic solution very well. The accuracy is improved by including higher orders in the integration. The most commonly used is the Runge-Kutta fourth order formula [80, 81]

$$y_{n+1} = y_n + \left(\frac{k_1}{6} + \frac{k_2}{3} + \frac{k_3}{3} + \frac{k_4}{6}\right)\Delta z \tag{3.23}$$

with

$$\begin{aligned}
k_1 &= f(z_n, y_n) \\
k_2 &= f(z_n + \frac{\Delta z}{2}, y_n + \frac{k_1}{2}) \\
k_3 &= f(z_n + \frac{\Delta z}{2}, y_n + \frac{k_2}{2}) \\
k_4 &= f(z_n + \Delta z, y_n + k3) \quad .
\end{aligned}$$

Fig. 3.3 indicates how this method works. Starting from the initial point three other points are calculated, always based on the position and slope of the previous point. For the final point the values of the differential equation $y' = f(z, y)$ at each point are weighted and added up.
The Runge-Kutta method is a robust and valid approach to solve any differential equation although the accuracy – not always identical with the order [81]– of this method might be poor if the integration step size becomes too large. Another problem is that the Runge-Kutta method is an explicit integration scheme with an inherent instability for certain sets of differential equations and large steps sizes [81]. By choosing an integration step length Δz smaller than the gain length, which is the characteristic length for the dependence of the electron energy and phase on z, this problem can be avoided. The major advantage is that this method is very reliable even for non-continuous terms in the differential equations, such as the jump in the undulator field amplitude at the entrance and exit of an undulator module. Therefore the Runge-Kutta method is found in most FEL simulation codes.
Somewhere between the Runge-Kutta and the highly sophisticated Bulirsch-Stoer integration, discussed below, the 'Predictor-Corrector' method is placed [86]. The algorithm is based on the extrapolation of the previously calculated function values as an initial guess of an iterative approximation. Unless this method uses a low order integration scheme the efficiency is comparable to or better than the Runge-Kutta method but the memory needed to store the history of all macro particles makes the Predictor-Corrector integration less favorable for FEL simulations.
In a more theoretical approach to the physics of an FEL it is common to normalize the variables and differential equations. The integration step size is typically a fractional part of the gain length and depends rather on the electron beam properties than on the undulator parameters.

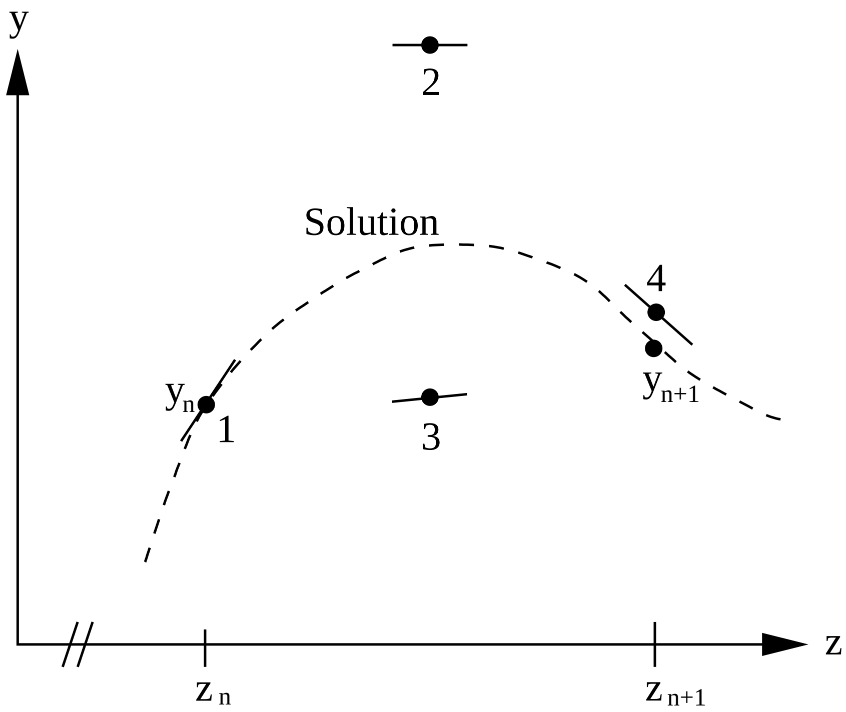


Figure 3.3: *4th order Runge-Kutta method to solve an ordinary differential equation. The order of the four evaluated points in the (z, y)-plane is indicated by the numbers.*

Using this model the differential equations are rather smooth and a more sophisticated solver can be used, such as the Bulirsch-Stoer method [87]. An extended discussion of this method would be beyond the scope of this section and only the outline of the algorithm is given here.
The basic idea is to divide the integration from z to $z + \Delta z$ into m steps. For coarse steps the extrapolated solution y_m might be far from the analytic solution y but converges towards it when the number of steps is increased. It is not desirable to increase m till the desired accuracy is reached, but the convergence point is predicted by fitting rational or polynomial functions to the extrapolated integration points y_m.
To be more explicit, a monotonically increasing sequence m_k defines the number of steps for the kth iteration. The step length is $h_k = \Delta z / m_k$. The extrapolation from z to $z + \Delta z$ within m_k steps can be regarded as a general function $y(h)$. After k iterations several sample points of this function are obtained with $y_k = y(h_k)$. As mentioned above the analytic solution at the new integration point $z + \Delta z$ is identical with $y(0)$. For k sample points a polynomial of the order $k - 1$ exists, going through all points.
The most critical question of this integration method is: "When should the iteration stop and the extrapolated value $y(0)$ be accepted as the new value ?". The answer is that the resulting error must be below the tolerance while the computational work is as small as possible. The error estimate $\epsilon_{k,k+1}$ is the change of the extrapolated value $y(0)$ for the next higher iteration.
The termination criterion is simply

$$\epsilon_{k,k+1} < \epsilon \quad , \tag{3.24}$$

where ϵ is the tolerance level determining the accuracy of the integration. It is a trivial fact that for large steps, Δz, more iterations are needed than for smaller steps but it might be faster to perform two complete integrations with fewer iterations and half the integration step length than one integration over the full distance Δz. To optimize the overall performance the step length Δz should be adaptive. For the sequence $m_k = 2k$ of the Bulirsch-Stoer method the estimated error [88] of the kth iteration is proportional to $\epsilon_{k,k+1} \propto \Delta z^{2k+1}$. Therefore the optimum step length for k integrations would be

$$\Delta z_k = \Delta z \left(\frac{\epsilon}{\epsilon_{k,k+1}} \right)^{\frac{1}{2k+1}} \quad . \tag{3.25}$$

The choice of the optimum step length Δz_k, and thus the number of iteration, has to be balanced against the computational work $A_{k+1} = A_k + m_{k+1}$ with the starting value $A_0 = 1$. For the best performance the ratio $W_k = A_{k+1}/\Delta z_k$ should be minimal. If the k_fth iteration fulfills the termination criterion of Eq. 3.24 the next integration chooses the adapted step length Δz_k, where the quantity W_k $(1 \leq k \leq k_f)$ has a minimum.
The algorithm might run into two kinds of problem, where Δz is either too small or too large. In the first problem the adaptive step size Δz_k will always be the same $(k = k_f)$ although the algorithm might be more efficient for larger step sizes. When the integration step becomes too large, too many iterations have to be carried out before the convergence is acceptable. If the integration runs into one of these problems, which can easily be monitored, the calculation of $\epsilon_{k,k+1}$ and Δz_k are approximated by an empirical function of the 'average expected convergence behavior' [88].
The major advantage of the Bulirsch-Stoer integration is its capability to vary the step size for best performance. If the curvature of the solution is large the step size is reduced to resolve it in a reasonable way. In the opposite case, where the function behaves almost linearly, the position z is advanced by large steps. Pitfalls exist when the curvature is very large, approaching the limit of a discontinuity in the derivate of the solution, such as a singularity. The Bulirsch-Stoer method tries to resolve these bends and discontinuities by increasing the number of iteration and decreasing the step length Δz. In these cases this method is highly inefficient and a simpler but more robust solver would be more convenient.
To conclude this section the Bulirsch-Stoer algorithm, as it is implemented in the FEL simulation code, is presented step by step.

1. Initialize the integration $(k := 1)$
2. If allowed number of iterations is exceeded $(k > k_f)$, restart integration with smaller step length Δz (step 1)
3. Integrate ordinary differential equation from z to $z + \Delta z$ by performing m_k steps.
4. Extrapolate to zero step size by a $(k-1)$th order polynomial.

5. Calculate error $\epsilon_{k,k+1}$, work function A_k and the work per unit step W_k.

6. If error $\epsilon_{k,k+1}$ does not lie within the tolerance ϵ, start next iteration (step 2 with $k := k+1$).

7. Compute most efficient order k of iteration, where W_k has its minimum.

8. Advance position $z \to z + \Delta z$, adapt step size for best performance $\Delta z = \Delta z_k$ and start new integration.

3.4 Integration of the Radiation and the Electrostatic Field Equation

In comparison to ordinary differential equations additional problems arise in the treatment of partial differential equations such as the radiation or electrostatic field equations Eqs. 3.10 and 3.11, respectively. While for a particle only 6 values are sufficient to precisely define its state of motion, the radiation and electrostatic field are defined at every point of the whole transverse plane. A complete description of the field is therefore impossible in a numeric code and has to be approximated by a finite number of parameters. There are various approaches for defining these parameters. The most common methods use a discretization of the field on a regular grid ('finite differences') or on simply shaped elements, describing an irregularly shaped domain, where the field is defined ('finite elements'), the expansion into orthonormal functions ('finite modes') or the description by the moments in the field distribution ('method of moments'). The decomposition of the radiation field into Hermite-Gaussian modes for free space propagation (see Appendix A) is an example for the 'finite modes' method but is less suitable for the FEL interaction. Only the 'finite difference method' will be discussed in this section.

Using finite differences the problem of the boundary condition arises because the grid cannot be extended to infinity in order to keep the number of grid points limited. The most commonly used boundary condition in FEL simulations is the Dirichlet boundary condition, where the field has a certain amplitude on all outer points of the grid. If the Dirichlet boundary conditions forces the radiation field to vanish at the boundary of the grid, the underlying model represents the field propagation within a waveguide, where the radiation field is reflected by the perfectly conducting walls of a waveguide. The situation of a less than perfectly conducting material of the waveguide is given by the Neumann boundary condition with a given value for the normal derivative of the radiation field. For a large grid and a localized radiation field with low diffraction both conditions are very close to the free space propagation. The constraint of a large grid is the most significant drawback of this method which cannot be avoided.

The transverse position of the macro particles is normally not identical with a grid point and a way has to be found to assign the macro particles to the grid points and to interpolate the radiation field to the particle position. The simplest way, where a macro particle sees only the radiation field of the nearest grid point, has the disadvantage of exhibiting the largest fluctuation in the number of assigned macro particles. For N_g macro particles in average at a certain grid point the root-mean-square fluctuation is $\sigma_N = \sqrt{N_g}$. The relative fluctuation

σ_N/N_g can be reduced either by using a coarser grid, by more macro particles or by a different assignment such as linear interpolation. Linear interpolation reduces the fluctuation by a factor of two [89]. This improvement can be exceeded by more sophisticated assignment functions. Examples are the parabolic or Gaussian interpolations.
Special care has to be taken for the transverse tails of the electron beam distribution. There the number of macro particles is small and, consequently, the fluctuation is large. Fortunately these tails do not contribute significantly to the FEL amplification and the larger fluctuation has less impact.
Linear interpolation also modifies the statistic of the local bunching factor. To illustrate this problem the case is considered, where one macro particle and its three mirror particles (see Section 3.2) are placed in the middle of a cell. The resulting bunching factor of these four particles follows the statistics for a certain number of electrons. In the linear interpolation one fourth of this bunching factor is assigned to each grid point of this cell as the source term in the field equation 3.10. Averaging over many samples an enhancement of a factor of 16 is observed, if the total number of electrons is extracted from the statistics. To provide the correct statistics the nearest grid point assignment should be chosen for simulation of SASE FELs, at least for constructing the source terms of Eqs. 3.10 and 3.11.
For a general discretization of the field distribution a mesh is superimposed on the area of interest G – here the transverse plane. To limit the number of grid points the mesh is finite and the field amplitude u at the edge ∂G of the mesh has to fulfill the boundary conditions. For simplicity the Dirichlet boundary condition is applied with $u|_{\partial G} = 0$. Each grid point $\vec{r}_j$ is enclosed by an associated cell $\mathcal{V}_j$, bounded by normal planes at the midpoints between two adjoining grid points. The union of all non-overlapping cells reensembles the entire domain

$$G = \bigcup_j \mathcal{V}_j \quad . \tag{3.26}$$

The volume of the jth cell is V_j, which has the dimension of an area in the case of a discretized transverse plane. There is no special need to order the grid points symmetrically although the field equations as well as the efficiency of the memory storage on a computer can be optimized for a symmetric layout of the mesh.
Using this definition the partial differential equations are averaged over each cell $\mathcal{V}_j$. The transverse Laplace operator can be simplified with the help of Green's identity

$$\frac{1}{V_j}\int_{\mathcal{V}_j} \nabla^2_\perp u\, dV = \frac{1}{V_j}\int_{\partial\mathcal{V}_j} \vec{\nabla}_\perp u\, d\vec{A} \tag{3.27}$$

For FEL simulations the most common discretizations are either the equidistant two dimensional Cartesian mesh $\vec{r}_j \equiv \vec{r}_{n(j),m(j)} = \vec{r}_0 + (n(j)\Delta)\vec{e}_x + (m(j)\Delta)\vec{e}_y$, where $\vec{e}_{x,y}$ are unit vectors in the x- and y-direction respectively, $\vec{r}_0$ is the origin of the mesh, Δ is the separation of two adjoining grid points and $(n(j), m(j))$ are the grid point indices ordered by j with $j > 0$, or the one dimensional radial mesh $\vec{r}_j = r_j\vec{e}_r$, where $\vec{e}_r$ is a unit vector in the radial direction and r_j is a monotonically increasing sequence with $r_1 = 0$. Fig. 3.4 shows these two meshes as well as the integration path, indicated by arrows, of the surface integration in Eq. 3.27.

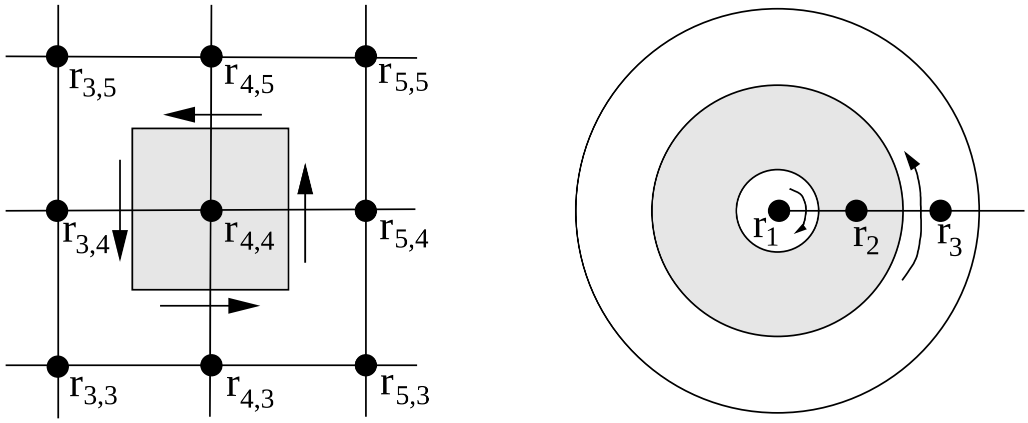


Figure 3.4: *Discretization on a Cartesian and radial mesh, left and right, respectively. The arrows indicate the orientation of the integration path enclosing the grey shaded area of interest.*

With the special notation $u_{n,m} = u(\vec{r}_{n,m})$ for the discretized Cartesian mesh, where the ordering index j has been omitted, the Laplace operator becomes

$$\nabla_\perp^2 u_{n,m} = \frac{u_{n,m+1} + u_{n,m-1} + u_{n+1,m} + u_{n-1,m} - 4u_{n,m}}{\Delta^2} \quad . \tag{3.28}$$

Field elements, which lie outside the boundary ∂G are set to zero to force the Dirichlet boundary condition. The Laplace operator could also have been found by the intuitive guess of approximating the differential operation by $u'_{n,m+1/2} \approx (u_{n,m+1} - u_{n,m})/\Delta$ and $u''_{n,m} = (u'_{n,m+1/2} - u'_{n,m-1/2})/\Delta$. For a limited Cartesian mesh there is always the freedom to order all mesh points as a sequence. With a given, but not further defined order of the Cartesian grid points $(n(j), m(j))$, the Laplace operator in Eq. 3.28 can be expressed by a simple matrix operation $\boldsymbol{\mathcal{L}u}$ of the vector $\boldsymbol{u} = (u_{n(1),m(1)}, u_{n(2),m(2)} \dots)$ [81]. The matrix $\boldsymbol{\mathcal{L}}$ is sparse and has only up to 5 non-zero elements per row or column. The main diagonal elements are all the same with a value of $-4/\Delta^2$.
The one dimensional radial mesh is convenient if the field is axi-symmetric. The azimuthal dependence on φ is removed by the expansion of the field into a Fourier series, where $\hat{m}$ indicates the $\hat{m}$th azimuthal mode ($\propto \exp[i\hat{m}\varphi]$). The derivative with respect to φ in the Laplace operator is replaced by $-\hat{m}^2/r^2$. The resulting matrix elements of the discretized Laplace operator are

$$\mathcal{L}_{j,j-1} = \frac{1}{\pi(r_{j+1/2}^2 - r_{j-1/2}^2)} \frac{2\pi r_{j-1/2}}{r_j - r_{j-1}} \tag{3.29}$$

$$\mathcal{L}_{j,j+1} = \frac{1}{\pi(r_{j+1/2}^2 - r_{j-1/2}^2)} \frac{2\pi r_{j+1/2}}{r_{j+1} - r_j} \tag{3.30}$$

$$\mathcal{L}_{j,j} = -(\mathcal{L}_{j,j-1} + \mathcal{L}_{j,j+1}) - \frac{2\pi\hat{m}^2}{\pi(r_{j+1/2}^2 - r_{j-1/2}^2)} \log\left[\frac{r_{j+1/2}}{r_{j-1/2}}\right] \tag{3.31}$$

and zero otherwise with $r_{j+1/2} = (r_j + r_{j+1})/2$ and $r_{1/2} = 0$ for the innermost point [90]. Problems arise at the origin for field amplitudes of higher modes ($\hat{m} \neq 0$), where the logarithmic function has a singularity. Mathematically a dependence on φ cannot exist at the origin. For this reason the logarithmic term in $\mathcal{L}_{1,1}$ is dropped. Another possible solution is to introduce an arbitrarily chosen shielding radius replacing the singularity of $r_{1/2}$.
The electrostatic problem (Eq. 3.11) has the freedom of matching the grid to the electron beam because it is recalculated after each integration step without the reuse of the previous results. If the electron beam is off the undulator axis the grid can be placed in the way that the beam center is always at the origin. For cases where the beam is almost round or elliptical all modes except for the monopole and quadrupole of a multipole expansion into cylindrical coordinates are significantly suppressed unlike the expansion of an off-center beam. Expanding the electrostatic field into a Fourier series in the azimuthal angle and longitudinal position the field equation becomes

$$\left[\boldsymbol{\mathcal{L}} - \frac{\hat{l}^2 k^2 (1+K^2)}{\gamma_R^2} \boldsymbol{\mathcal{I}}\right] \boldsymbol{E}_{\hat{l},\hat{m}} = \boldsymbol{\rho}_{\hat{l},\hat{m}} \quad , \tag{3.32}$$

where $\boldsymbol{\mathcal{L}}$ is given by Eqs. 3.29 – 3.31, $\boldsymbol{\mathcal{I}}$ is the unit matrix and the jth element of the source term vector is

$$(\rho_{\hat{l},\hat{m}})_j = i \frac{ec^2 \mu_0 \hat{l} k (1+K^2)}{\pi (r_{j+1/2}^2 - r_{j-1/2}^2) \gamma_R^2} \sum_{\vec{\rho}_n \in \mathcal{V}_j} e^{-i\hat{l}\theta_n - i\hat{m}\varphi_n} \tag{3.33}$$

The sum in Eq. 3.33 runs over all macro particles, which positions $\vec{\rho}_n$ lie within the cell $\mathcal{V}_j$, enclosing the jth grid point. This assignment corresponds to the nearest grid point method.
The electrostatic problem is reduced to find the solution of a set of linear equations or, which is equivalent, to the inverse of the matrix $\tilde{\boldsymbol{\mathcal{L}}} = \boldsymbol{\mathcal{L}} - (\hat{l}^2 k^2 (1+K^2)/\gamma_R^2)\boldsymbol{\mathcal{I}}$. General methods of solving Eq. 3.32, even for a different discretization scheme, which are not biased by the special shape of $\tilde{\boldsymbol{\mathcal{L}}}$ as a sparse matrix, are presented later in this section.
Due to the symmetry of the matrix $\tilde{\boldsymbol{\mathcal{L}}}$, which has besides the main diagonal only non-vanishing array elements above and below the main diagonal, an algorithm exists to solve Eq. 3.32 without storing the complete, sparse matrix or its non-sparse inverse in memory. Going through all rows of the matrix except for the last and starting with the first, all three components of the jth row of $\tilde{\boldsymbol{\mathcal{L}}}$ and the jth element of the vector $\boldsymbol{\rho}_{\hat{l},\hat{m}}$ are normalized to $\tilde{\mathcal{L}}_{j,j}/\tilde{\mathcal{L}}_{j+1,j}$ and subtracted from the next row. The matrix is reduced to the main diagonal and the diagonal above. The last row depends only on the unknown value of the outermost electric field and can be directly calculated. The problem is solved by going backwards and inserting the values obtained for the electrostatic field in the previous row. For a $J \times J$ matrix only $7J$ arithmetic operations have to be performed with a memory storage of $4J$ values. With typical values of $J = 100$ grid points the calculation of the electrostatic field is one or two order of magnitudes faster than integrating the differential equations of 10^4 macro particles. Another benefit is that this method of solving a tridiagonal system of equations is unconditionally stable, because $\tilde{\boldsymbol{\mathcal{L}}}$ always has a 'diagonal dominance' with $|\tilde{\mathcal{L}}_{j,j}| > |\tilde{\mathcal{L}}_{j,j-1}| + |\tilde{\mathcal{L}}_{j,j+1}|$ for all rows [81].

This straightforward method for a one dimensional radial grid and the decomposition into multipoles becomes more and more inefficient when the source terms are not axi-symmetric. It can be compensated by including higher multipoles but suffers under the conversion of the radial mesh of the electric field and the Cartesian coordinate system of the macro particles. For each azimuthal mode the source term has to be recalculated and the benefit of a small number of grid points compared to the large number of macro particles is lost. If the simulation code shall cover arbitrary particle and field distributions it is advisable to base the discretization on a Cartesian mesh. Regarding the radiation field equation Eq. 3.10 the fast solution of a tridiagonal system can be applied at least partially, being discussed later.
The steady-state radiation field equation

$$\left[\nabla_\perp^2 + 2ik\frac{\partial}{\partial z}\right] u = S \tag{3.34}$$

is classified as a parabolic partial differential equation in mathematics [59]. The source term S is identical with the right hand side of Eq. 3.10. In the Free-Electron Laser amplification process this type of partial differential equation is typical for diffusion problems and, in particular due to the imaginary constant of the time derivative, for Schrödinger's time-dependent quantum mechanics. In addition to the transverse discretization the radiation field is only calculated at certain longitudinal positions z_n. The index n of the longitudinal integration step is used as an upper index of the radiation field u_j^n to distinguish the longitudinal position from the transverse position, indicated by the lower index. The distance between two succeeding longitudinal positions is Δz. With the definition of the z-derivative $(u_j^{n+1} - u_j^n)/\Delta z$ the question arises at which position z between z_n and $z_{n+1} = z_n + \Delta z$ the Laplace operator $\boldsymbol{\mathcal{L}}$ has to be evaluated. As a general and valid 'ansatz' the weighted sum $\alpha u_j^{n+1} + (1-\alpha)u_j^n$ is used with $0 \leq \alpha \leq 1$. The discretized, parabolic equation becomes in the matrix-vector notation

$$\boldsymbol{u}^{n+1} = \boldsymbol{u}^n + i\frac{\Delta z}{2k}\boldsymbol{\mathcal{L}}[\alpha \boldsymbol{u}^{n+1} + (1-\alpha)\boldsymbol{u}^n] + \boldsymbol{s}^{n+1/2}\Delta z \quad . \tag{3.35}$$

The source term $\boldsymbol{s} \equiv \boldsymbol{S}/2ik$ is defined by the macro particles in the middle of the integration step Δz according to the leapfrog integration scheme. Unfortunately the matrix $\boldsymbol{\mathcal{L}}$ is not tridiagonal if the field $\boldsymbol{u}$ is defined on a Cartesian mesh or any other 2D mesh because it has up to five non-zero array elements per column or row. The fast algorithm of the electrostatic problem cannot be applied.
For numerical reasons it would be convenient to set $\alpha = 0$. No matrix needs to be inverted and the field equation is solved by just multiplying the field vector $\boldsymbol{u}^n$ with a sparse matrix and adding the source term. Because the evolution of the radiation field in z depends only on the 'old' field values the integration is referred to as the full explicit 'Forward Time' integration [91].
For $\alpha = 1$, in contrast to the explicit integration, $\boldsymbol{\mathcal{L}}$ operates solely on the 'new' variables, still to be calculated. This method is called 'full implicit'. The choice of $\alpha = 1/2$ yields the Crank-Nicholson integration, being partially implicit and explicit by the equal weighting of $\boldsymbol{u}^n$ and $\boldsymbol{u}^{n+1}$. The Crank-Nicholson is of importance for solving the time-dependent Schrödinger

equation, because this choice of α is the only one, which corresponds to an unitary Hamilton operator [92].
Assuming that a method has been found to invert the matrix $\tilde{\mathcal{L}} = \mathcal{I} - i(\alpha\Delta z/2k)\mathcal{L}$ the integration of the radiation field over z is a recursive process. This method inherits a possible error source by carrying numerical errors through all calculations. Under special circumstances these errors might add up and the integration scheme would become unstable. To check the applicable range of this method the 'von Neumann stability analysis' estimates the evolution of the error $\boldsymbol{\epsilon}$ in z [93]. Because the parabolic field equation is linear any arbitrary field distribution can be decomposed into the exact solution $\boldsymbol{u}$, which solves Eq. 3.35, and the residual error $\boldsymbol{\epsilon}$. Inserting this decomposition into the discretized field equations gives a homogeneous parabolic equation for the error

$$[\mathcal{I} - i(\alpha\Delta z/2k)\mathcal{L}]\,\boldsymbol{\epsilon}^{n+1} = [\mathcal{I} + i((1-\alpha)\Delta z/2k)\mathcal{L}]\,\boldsymbol{\epsilon}^{n} \tag{3.36}$$

The error is decomposed into eigenmodes of the Cartesian mesh with

$$\epsilon^{n}_{j_x,j_y} = \xi^n e^{i(j_x k_x \Delta + j_y k_y \Delta)} \quad , \tag{3.37}$$

where k_x and k_y can have any arbitrary value. For the stability analysis it is only necessary that the eigenfunction ξ^n exists. The explicit solution is not of importance. Inserting Eq. 3.37 into the matrix equation 3.36 yields, after some basic algebra,

$$\xi^{n+1} = \frac{1 - i(1-\alpha)\frac{2\Delta z}{k\Delta^2}\left[\sin^2\left(\frac{k_x}{2}\Delta\right) + \sin^2\left(\frac{k_y}{2}\Delta\right)\right]}{1 + i\alpha\frac{2\Delta z}{k\Delta^2}\left[\sin^2\left(\frac{k_x}{2}\Delta\right) + \sin^2\left(\frac{k_y}{2}\Delta\right)\right]}\xi^n \quad . \tag{3.38}$$

At each integration step the eigenfunction is scaled by a complex number. To guarantee stability the eigenmode must not grow in amplitude and requires an absolute value of the scaling factor equal or less than unity. The weighting parameter α, so far undefined, has therefore to fulfill the constraint $\alpha \geq 1/2$. Comparing to the Crank-Nicholson integration scheme the full implicit method is numerically stable and damps the eigenmodes of the error. The reason why the Crank-Nicholson integration is often preferred, lies in the accuracy, which is one order better ($\mathcal{O}(\Delta z^2)$) than the full implicit method ($\mathcal{O}(\Delta z)$) [81].
The stability analysis above has been made under the assumption that the inversion of the matrix $\tilde{\mathcal{I}}$ is completed within a single step. Unfortunately this is not always possible on modern computers. For a Cartesian mesh with 100 mesh point in each direction the matrix has 10^8 array elements and the exact inversion would need enormous memory resources. Therefore it is unavoidable either to use approximation methods, discussed at the end of this section, or to modify the problem to obtain a more convenient matrix such as a tridiagonal matrix. The latter case is the topic of the following discussion.
The matrix operator $\mathcal{L}$ consists of two independent parts $\mathcal{L}_x$ and $\mathcal{L}_y$, describing the second derivative of the Laplace operator in x and y respectively. Depending on the order of the grid points, one of these matrices is tridiagonal. The fact that the matrix $\mathcal{L} = \mathcal{L}_x + \mathcal{L}_y$ can be split suggests that the integration can be divided into two steps with a step length of $\Delta z/2$ each. Only one operator is applied to the new field values, yielding a full implicit integration in this

direction. In the following integration step, the operators are exchanged. This is known as the alternating-direction implicit (ADI) method [94, 95], which can easily be generalized for any mesh of arbitrary dimension. The two integrations are

$$\boldsymbol{u}^{n+1/2} = \boldsymbol{u}^n + i\frac{\Delta z}{4k}[\boldsymbol{\mathcal{L}}_x \boldsymbol{u}^{n+1/2} + \boldsymbol{\mathcal{L}}_y \boldsymbol{u}^n] + \boldsymbol{s}^{n+1/2}\frac{\Delta z}{2} \quad , \tag{3.39}$$

$$\boldsymbol{u}^{n+1} = \boldsymbol{u}^{n+1/2} + i\frac{\Delta z}{4k}[\boldsymbol{\mathcal{L}}_x \boldsymbol{u}^{n+1/2} + \boldsymbol{\mathcal{L}}_y \boldsymbol{u}^{n+1}] + \boldsymbol{s}^{n+1/2}\frac{\Delta z}{2} \quad . \tag{3.40}$$

If the grid points are reorganized after each step, the matrix, corresponding to the implicit integration ($\boldsymbol{\mathcal{L}}_x$ in Eq. 3.39 and $\boldsymbol{\mathcal{L}}_y$ in Eq. 3.40), can always be transformed into a tridiagonal matrix. The matrix $\boldsymbol{\mathcal{I}} - (i\Delta z/4k)\boldsymbol{\mathcal{L}}_{x/y}$, which must be inverted, always has diagonal dominance and the fast and efficient algorithm, presented for the electrostatic problem, can be applied. Because the full integration contains one implicit step in each direction it is stable overall. The von Neumann stability analysis yields

$$\xi^{n+1} = \frac{\left[1 - i\frac{\Delta z}{k\Delta^2}\sin^2\left(\frac{k_x}{2}\Delta\right)\right]\left[1 - i\frac{\Delta z}{k\Delta^2}\sin^2\left(\frac{k_y}{2}\Delta\right)\right]}{\left[1 + i\frac{\Delta z}{k\Delta^2}\sin^2\left(\frac{k_x}{2}\Delta\right)\right]\left[1 + i\frac{\Delta z}{k\Delta^2}\sin^2\left(\frac{k_y}{2}\Delta\right)\right]}\xi^n \quad . \tag{3.41}$$

The eigenfunction ξ^n is scaled by a complex factor with an absolute value of unity. The ADI methods provides a stable integration, although the numerical errors are not damped.

Any discretization scheme converts the field equation into a system of linear equations. The Laplace operator is transformed into the matrix $\boldsymbol{\mathcal{L}}$. The size of this matrix is the square of the number of grid points and a two or higher dimensional discretization will often be beyond the possibility of modern computers to store the complete matrix or its inverse in memory. The methods presented so far are restricted to a one dimensional grid for the hyperbolic (electrostatic) problem and to an n dimensional grid for the parabolic equation of the radiation field using an alternating-direction implicit integration scheme. Although these methods are sufficient to simulate the FEL process correctly and reasonably fast, more sophisticated integration schemes exist.

In particular the simulation of SASE Free-Electron Lasers is rather inefficient when the ADI integration scheme is used. The mesh must be large to simulate the fast diffracting modes and to reduce the impact of the boundary conditions at the beginning of the SASE FEL amplification process. At an advanced stage of the FEL amplification the fundamental mode dominates and the outer parts of the grid are filled with negligible field amplitudes. To optimize the simulation performance it would be desirable to cut out these areas if they do not contribute significantly to the radiation field profile.

Adaptive multigrid integration is a method, which fulfills all these requirements, by still being competitive with the straightforward integration of the alternating-direction implicit integration. These two methods are related to each other in the same way as the Bulirsch-Stoer and Runge-Kutta integration for ordinary differential equations (Section 3.3). Indeed the adaptive multigrid integration is based on the same principle as the Bulirsch-Stoer method, to find the optimum balance between computational work and achievable accuracy. To understand the

adaptive multigrid method – a detailed description would be beyond the scope of this thesis – other methods have to be explained first, which are basic elements of the final algorithm [96]. Unless the matrix equations 3.32 or 3.35 cannot be solved directly the solutions must be found by an iterative method starting with an initial guess. The iteration stops when the difference between the iterative and exact solution is below the tolerance level. Because the exact solution is unknown the error has to be estimated, e.g. by the convergence rate of the iteration.
To avoid confusion and reduce the number of new indices or parameters the notation is redefined. The general problem is to solve the discretized field equation

$$\boldsymbol{\mathcal{M}}\boldsymbol{u} = \boldsymbol{\rho} \quad . \tag{3.42}$$

If the field $\boldsymbol{u}$ and the source term $\boldsymbol{\rho}$ are printed in bold letters the discussion refers to the complete vector of these values, discretized on the Cartesian mesh. A lower index indicates a single element of the vector.
The parabolic equation Eq. 3.35 contains the radiation field values at two succeeding position z_n and z_{n+1}. Regardless whether the Crank-Nicholson, the full implicit method or any other interpolation ($1/2 < \alpha < 1$) is used, the known field at z_n is incorporated into the source term $\boldsymbol{\rho}$ and $\boldsymbol{u}$ refers solely to the field at z_{n+1}.
The matrix $\boldsymbol{\mathcal{M}}$ has diagonal dominance for both the electrostatic and the radiation field problem and it is split into three matrices $\boldsymbol{\mathcal{M}}$=$\boldsymbol{\mathcal{D}}$−$\boldsymbol{\mathcal{A}}$−$\boldsymbol{\mathcal{B}}$, where $\boldsymbol{\mathcal{D}}$ contains only the diagonal elements, $\boldsymbol{\mathcal{A}}$ and $\boldsymbol{\mathcal{B}}$ all elements above and below, respectively.
A formal solution is $\boldsymbol{u} = \boldsymbol{\mathcal{D}}^{-1}(\boldsymbol{\mathcal{A}} + \boldsymbol{\mathcal{B}})\boldsymbol{u} + \boldsymbol{\mathcal{D}}^{-1}\boldsymbol{\rho}$. Although this equation is not simpler to solve it is the base of a recursive relaxation algorithm

$$\boldsymbol{u}^{l+1} = \boldsymbol{\mathcal{D}}^{-1}(\boldsymbol{\mathcal{A}} + \boldsymbol{\mathcal{B}})\boldsymbol{u}^{l} + \boldsymbol{\mathcal{D}}^{-1}\boldsymbol{\rho} \quad , \tag{3.43}$$

where the approximation $\boldsymbol{u}^l$ converges towards the exact solution. The upper index l has a different definition for the discussion of the adaptive multigrid method and indicates the iteration steps. The former upper index n for the longitudinal position is omitted for the remaining part of this section because only one integration step in z is regarded. Therefore n is fixed and it is needless to indicate this step by an index.
Using this definition of recursion, a single field amplitude is calculated by

$$u_j^{l+1} = -\frac{1}{\mathcal{M}_{jj}}\left(\sum_{i\neq j}\mathcal{M}_{ji}u_i^l - \rho_j\right) \tag{3.44}$$

and depends only on the field values of the previous iteration. This is the definition of the Jacobi relaxation.
Summing over i in Eq. 3.44 new iteration values are known for $i < j$. If these 'updated' values are already used in the current summation the iteration is modified to the Gauss-Seidel relaxation corresponding to

$$\boldsymbol{u}^{l+1} = (\boldsymbol{\mathcal{D}} - \boldsymbol{\mathcal{B}})^{-1}(\boldsymbol{\mathcal{A}}\boldsymbol{u}^{l} + \boldsymbol{\rho}) \quad . \tag{3.45}$$

These methods to solve a matrix equation by relaxation are highly inefficient if the grid is large. It takes many iterations before the information at a certain grid point propagates to all others because the iteration connects only five grid points per iteration in the Jacobi method. This is only slightly enhanced for Gauss-Seidel.
Defining the numerical error as the difference between the lth iteration and the asymptotic solution with $\boldsymbol{\epsilon}^l = \boldsymbol{u}^\infty - \boldsymbol{u}^l$, it can be expanded in a set of orthonormal functions on the grid, namely sine and cosine functions. As the relaxation has the effect of averaging over a few grid points, all close to each other, higher frequencies are damped faster than lower frequencies and the error after many iteration consists of only low frequencies.
The efficiency is further improved by the multigrid method. It always starts with several iterations to damp the higher frequencies, where the line between higher and lower frequencies is drawn at $(4\Delta)^{-1}$ with Δ as the grid spacing. The convergence rate $\bar{\mu}$ is defined as the poorest damping rate of all high frequencies and lies theoretically roughly at $\bar{\mu} \approx 0.5$ [97]. Three iterations would reduce the amplitudes of all high frequencies by one order of magnitude. The exact value of $\bar{\mu}$ depends on the size of the mesh, the kind of matrix $\boldsymbol{\mathcal{M}}$ and the relaxation scheme, where the Gauss-Seidel and Jacobi are only two out of a variety of different methods [96].
The relaxation follows an assignment of the field $\boldsymbol{u}$ to a new, coarser grid. The spacing of the grid points of the coarser is twice as large as that of the finer grid. Higher frequencies on the finer mesh cannot be resolved by the coarser grid, while the lower frequencies are partially converted to higher frequencies on the coarser grid. The assignment can either be done by direct injection, where the coarser grid point has the amplitude of the coincident grid point of the finer mesh, or a weighted sum including the surrounding points of the finer grid. Relaxation on this coarser grid damps the former low frequencies faster than on the finer grid and the overall damping rate is increased. To reduce even lower frequencies the step of relaxation and the restriction on a coarser grid is done several times till the coarsest grid allows a damping of all frequencies or even the discrete but exact solution with low expense in computational time and memory. The coarsest grid has typically less than 10 discrete grid points per dimension.
The second part of the multigrid integration inverts the direction of the assignment from the coarse grid to the next finer grid. The solution on the coarser grid is used as the correction to the field on the finer grid. The field amplitudes are interpolated to the intermediate grid points of the finer grid. The most practical interpolation is the bilinear interpolation. As this assignment might introduce new high frequency errors on the finer grid, each transfer to a finer grid should be concluded by another relaxation step. The excitation of these frequencies can be reduced by a higher order interpolation scheme such as cubic interpolation but will take longer.
The overall convergence of any iterative method depends significantly on the initial guess $\boldsymbol{u}^0$. A bad initial distribution needs longer till the error is damped. To avoid any initial guess on the finest grid and the resulting bad performance the full multigrid algorithm (FMG) starts with the coarsest grid for the initial guess and interpolates it to the next finer level. After one sweep of the multigrid method, including only these two grids, the solution is interpolated to the next finer grid as its initial guess. This is repeated till the initial field distribution of the finest grid is obtained. One final multigrid integration through all grids yields the desired solution. The calculation of the initial guess takes about 33% of a full multigrid integration step including all

levels of grids.
The multigrid and full multigrid methods can be used to obtain the solution of a matrix equation in a reasonable time simply by following a given procedure independent of the distribution itself. The algorithm extends the finest grid over the entire domain. This is highly inefficient if the solution is smooth and flat. The finest grid will provide no essential different result as it can be obtained by interpolating the solution from a coarser grid. An improved solver for matrix equations must be able to detect possible unneeded calculations and exclude them. These demands are realized by the adaptive multigrid method.
Similar to the Bulirsch-Stoer integration the depth of layers of multigrids is controlled by the optimization of the error relative to the total number of arithmetic operations referred to as 'work'. Both the error and the work depend on two parameters. One parameter is the mesh size $h(x,y)$, which depends for a general approach on x and y to allow for irregular meshes. In the following it is assumed that the ratio of two succeeding grid spacing is 2:1. Another more intrinsic parameter is the order of approximation for the Laplace operator p. A multigrid method, where this parameter can be varied, is extremely difficult to realize and therefore p is kept constant. If $w(p)$ is a measure of the computational work for a single grid point the total work is

$$W = \int_G \frac{w(p)}{h(x,y)^2} dxdy \quad , \tag{3.46}$$

where G is the full domain, which includes the grid. Assuming a computer without any limitation of precision and no errors in the calculation, the error $\boldsymbol{\tau}(h,p)$ depends only on the discretization,

$$\boldsymbol{\tau}(h,p) = \boldsymbol{\mathcal{M}}(h,p)\boldsymbol{\mathcal{I}}^h\tilde{\boldsymbol{u}} - \boldsymbol{\mathcal{I}}^h\tilde{\boldsymbol{\mathcal{M}}}\tilde{\boldsymbol{u}} \quad . \tag{3.47}$$

In Eq. 3.47 the operator $\boldsymbol{\mathcal{I}}^h$ represents the discretization on a grid with the mesh size h. The solution and the operator of the partial differential equations for the undiscretized, continuous domain ($h \to 0$) are $\tilde{\boldsymbol{u}}$ and $\tilde{\boldsymbol{\mathcal{M}}}$, respectively. The error $\boldsymbol{\tau}(h,p)$ is namely the difference in the solutions of a field equation defined in a continuous or a discretized space. The function values of $\boldsymbol{\tau}(h,p)$ might depend on the position within the grid and is therefore called local discretization error. The overall global error is

$$E = \int_G g(x,y)\tau(x,y,h(x,y),p)dxdy \quad . \tag{3.48}$$

An arbitrary function $g(x,y)$ has been introduced in Eq. 3.48 to weight the local error [97], because its influence might be different to the overall solution. A large error close to the edge of a grid modifies the boundary conditions more strongly than a point in the center of the grid. The optimization problem of the adaptive multigrid method is described by the Euler-Lagrange equation

$$\frac{\partial}{\partial h}\left[g(x,y)\tau(x,y,h(x,y),p) + \lambda_{opt}\frac{w(p)}{h(x,y)^2}\right] = 0 \quad . \tag{3.49}$$

The Lagrange multiplier λ_{opt} [37] is the control parameter of the optimization process. The derivative of the discretization error is approximated by the truncation error $\boldsymbol{\tau}_h^H$ with

$$h\frac{\partial \boldsymbol{\tau}}{\partial h} \approx \boldsymbol{\tau}(2h,p) - \boldsymbol{\tau}(h,p) = \boldsymbol{\tau}_h^{2h} = (\boldsymbol{\mathcal{M}}_{2h}\boldsymbol{\mathcal{I}}_h^{2h} - \boldsymbol{\mathcal{I}}_h^{2h}\boldsymbol{\mathcal{M}}_h)\boldsymbol{u}_h \quad , \tag{3.50}$$

where $\boldsymbol{u}_h$ is the discretized field on the h-size grid, $\boldsymbol{\mathcal{M}}_h \equiv \boldsymbol{\mathcal{M}}(h,p)$ and $\boldsymbol{\mathcal{I}}_h^{2h}$ is the operator to assign a given discretization to a coarser grid with a mesh size twice as large. Using this approximation the quantity

$$Q(x,y) = g(x,y)\tau_h^{2h}(x,y)h(x,y)^2 \tag{3.51}$$

can be calculated by the full multigrid (FMG) method with almost no extra computational time. This quantity, which is proportional to the first term of Eq. 3.49, can be interpreted as the local improvement of the accuracy by interpolating the radiation field to a finer grid. The values of Q at each grid point might deviate from an optimal value $Q_{opt} = 2w(p)\lambda_{opt}$, which fulfills the Euler-Lagrange equation for the best relation between acceptable error and computational work. Theoretical investigation of the multigrid algorithm estimates a quadratic dependence of the truncation error $\boldsymbol{\tau}_h^{2h}$ on h. Due to the 'average' dependence of Q [97] on h as $Q \propto h^4$ the mesh is locally refined around a certain grid point if the benefit of the higher accuracy is larger than the extra work to be done, namely $Q > 16Q_{opt}$. For a much larger value $Q > 256Q_{opt}$ the calculation would benefit from refining the grid twice.
Special care has to be taken while interpolating the boundary grid points of the local refined mesh. Because the finer grid is a sub-domain of the coarser grid, any solution must provide the correct amplitudes (Dirichlet boundary condition), derivatives (Neumann boundary conditions) and, if necessary, higher derivatives. The assignment of the boundary grid points must be based on a higher order interpolation scheme, typically a cubic interpolation.
In conclusion, the adaptive multigrid method is a highly sophisticated way to solve the FEL equations for the electrostatic and radiation fields. Compared to the straightforward calculation by inverting a tridiagonal matrix or the alternating-direction implicit integration, this method is faster if the simulation demands a large grid while the evolution in z restricts the radiation field to a fractional part of that grid. A typical example is the growing transverse coherence (see Chapter 5) of SASE FEL simulations.

3.5 Time-Dependent Simulation

So far computers have not reached a point in performance and memory to be a suitable platform for a fully self-consistent simulation of the time-dependent Free-Electron Laser process. The main problem is the huge amount of macro particles necessary to resolve the radiation wavelength sufficiently over the whole bunch length. With 10^4 macro particles resolving one radiation wavelength the simulation of a 300 μm electron bunch, lasing at 6 nm radiation wavelength demands a total number of $5 \cdot 10^8$ macro particles. This number is beyond the capability of recent computers.

A way out of this dilemma is the fact that FEL radiation is highly directional towards the head of the electron bunch. Indeed the paraxial approximation, which leads to the FEL equations Eqs. 3.1 - 3.6, 3.10 and 3.11, synchronizes the longitudinal velocity of electron beam and radiation field so that after one undulator period the radiation field is advanced one radiation wavelength with respect to the electron beam. In the frame of the moving electron beam the time derivative in Eq. 3.10 must be weighted to keep the FEL equations consistent in this preferable frame. According to Section 2.7 the factor applied is $-\lambda_U/\lambda$, where λ_U is the undulator period length and λ the radiation wave length. For an integration step of $\Delta z = \lambda_U$ the corresponding time step in the moving frame is $\Delta t = -\lambda/c$.
A discretization of the radiation field in z and t, separated by Δz and $(\lambda/c\lambda_U)\Delta z$ and indicated by the indices n and m, respectively, yields the modified field equation in the Crank-Nicholson integration scheme

$$\boldsymbol{u}^{n+1,m-1} - \boldsymbol{u}^{n,m} = \frac{i\Delta z}{4k}\boldsymbol{\mathcal{L}}(\boldsymbol{u}^{n,m} + \boldsymbol{u}^{n+1,m-1}) + \boldsymbol{\rho}^{n,m} \quad . \tag{3.52}$$

In this form the simulation code must keep all beam slices – the discretization in t – in memory at the same time, encountering the memory problem described above.
This problem is avoided by the basic assumption that, for a small integration step, the interaction time between the electron beam and a slice of the radiation field is too short to build up a significant collective instability before the field is further advanced. In practice the integration is split into two steps, the interaction and the slippage. For the interaction step the new radiation field $\boldsymbol{u}^{n+1,m-1}$ is replaced by the intermediate field $\tilde{\boldsymbol{u}}^{n+1,m}$ in Eq. 3.52. The resulting equation can be solved by the standard methods, described in the previous section. The second step is simply advancing the radiation field by replacing it with the field from the time slice behind it $\boldsymbol{u}^{n+1,m} = \tilde{\boldsymbol{u}}^{n+1,m-1}$.
This approximation allows that the entire electron beam does not need to be considered. If the integration step Δz is larger than one undulator period length, the slippage is larger than the length of the bucket of the electron ponderomotive wave, which is a good approximation for the radiation wavelength. The resulting discretization in t defines the positions of the slices, which are separated by $\lambda\Delta z/c\lambda_U$. Because the modified radiation field equation omits the time derivative, the field amplitude is extrapolated to the next discretization point (bucket) and periodicity is assumed before and after the bucket. This is identical to steady-state FEL simulations, where the extrapolation is extended to infinity. A discretization of t with a grid point spacing less than λ yields an unphysical overlap of the buckets. To handle smaller integration steps Δz less than λ_U the modified field equation is integrated several times before the electron beam is advanced at least one undulator period and the radiation field is replaced by the values of the previous grid point in t.
For the further discussion each discrete grid point in t refers to a sample slice (bucket) of the radiation field or electron beam. Each slice has a thickness of λ. Due to the freedom to choose any arbitrary value for Δz the slices do not necessarily cover the complete bunch.
The underlying model of this approach is that the radiation field successively passes the electron slices, each interacting with the radiation field over the distance Δz. This series of short amplifications is stopped if the end of the undulator is reached or the field escapes through the

first slice at the head of the electron beam.
The amount of memory needed is significantly reduced because at no time during the simulation are all the bunch slices represented in the computer. There are two optional ways to perform the simulation, where either a sample slice of the radiation field or of the electron beam is tracked over the whole undulator length. The latter method starts with the last slice at the tail of the electron beam. For each integration step the radiation field of this last slice is temporarily stored into a record. Advancing to the next slice the radiation field is taken from this record before each integration step, integrated and written back in the next record to store it for the electron slices of the following integration. Using this approach the memory demands are still rather high but are manageable for a modern computer. For N integration steps along the undulator axis the simulation needs N records to store the radiation field and one single record of the electron beam slice. The other method is similar; where a slice of radiation is held in memory and interacts with N succeeding sample slices of the electron beam. Starting at the head of the electron beam the algorithm goes backwards through the electron beam. After a radiation slice has been completely calculated for the whole undulator length, the last record is deleted and all other records are shifted up one position. The deleted record corresponds to the foremost slice of the electron beam. As the integration of the next radiation slice starts further backwards, this electron slice will not interact with the radiation field anymore. The first record, which has become free, is filled with the parameters of the electron beam, where the new integration cycle starts. This method has the disadvantage of an enhanced bookkeeping but the set of records can be much smaller, compared to the first method, if a large grid for the field discretization is used. The order of integration of both methods is schematically drawn in Fig. 3.5.
The remaining part of this section is concerned with the appropriate choice of the integration step length Δz and the total number of sample slices M, covering the electron beam. The whole algorithm of the time-dependent simulation is based on the assumption that over one integration step no significant collective instabilities occur. A measure of this collective phenomena is the gain length L_g. According to Chapter 2 the step length Δz is limited to

$$\Delta z \ll \frac{\lambda_U}{4\pi\rho} \quad , \tag{3.53}$$

where ρ is the FEL parameter. The restriction of 'much less' is not obligatory and even step sizes of one forth of the gain length provide reasonable results [98]. The extreme but rare case, when the gain length is comparable to the undulator period length, the whole approach of time-dependent simulation is not suitable any longer due to the conflicting limitation of Δz with $\lambda_U \leq \Delta z \ll L_g \approx \lambda_U$. Because the FEL saturates within a dozen undulator periods and, thus, the total slippage covers the same number of ponderomotive buckets the code is modified in such a way that it can keep the complete radiation field and all macro particles over a full slippage length. The memory expensive set of electron beam or radiation field records is significantly reduced in this extreme case of a short integration length. The integration yields at least one radiation field slice, which has seen the 'full' interaction with the electron beam over the whole undulator length. A 'full' interaction excludes slices, where the radiation field slips out of or into the sample range of the electron beam before the end of the undulator is

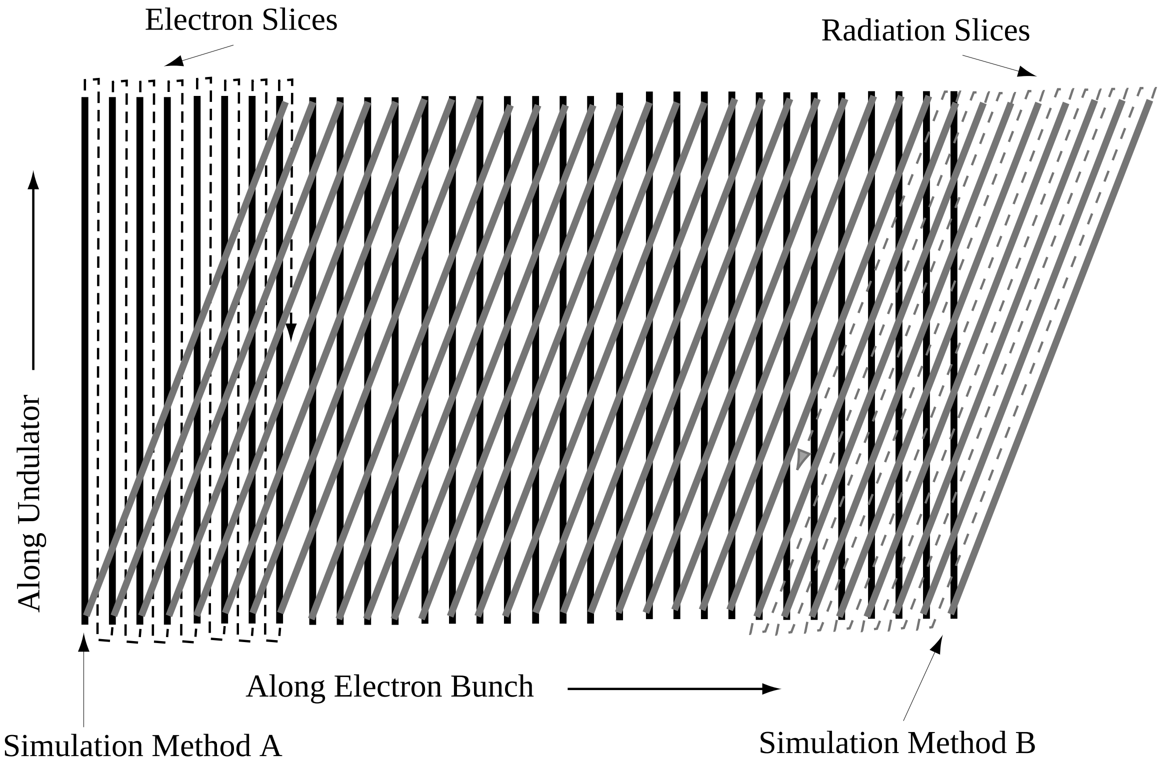


Figure 3.5: *Schematic order for time-dependent simulations. Due to the slippage the radiation field slices (tilted grey bars) propagate in the forward direction with respect to the electron beam slices (black bars). The integration can either be performed by starting from the end of the bunch and keeping one electron beam slice in memory or from the bunch head with a radiation slice in memory (Method A and B, respectively).*

reached. This integration is repeated for different positions within the electron beam to obtain a reasonable radiation profile.
Another important parameter is the total length of the simulated electron beam. As the radiation wavelength becomes longer, the total slippage of the radiation field can be comparable to the bunch length. An appropriate simulation should cover the whole bunch length including the radiation field escaping through the head of the electron bunch. Because the bucket size of the ponderomotive wave is rather large the total number of slices M is limited and the simulation can be performed within a reasonable time. For high beam energy and short radiation wavelengths the simulation time exceeds the acceptable limit due to the enormous number of radiation wavelengths fitting within the bunch length. Even for a rather large integration step size only a subsection of the electron beam can be considered in the simulation. Using this approach the simulation code cannot calculate the correct radiation field which would slip through the back of the subsection of the beam during the FEL interaction. After the complete integration over the total undulator length $L_u = N \cdot \Delta z$ the radiation field of the last N slices are physically not correct and must be ignored. For this reason the 'time' window defined (sub-

section of the electron beam) is always reduced by the full slippage length. The total number of simulated slices M must therefore fulfill the condition

$$M \geq \frac{L_u}{\Delta z} \quad . \tag{3.54}$$

This guarantees that at least one radiation slice remains in the simulated subsection of the electron beam.
Although the result of a single slice might be reasonable, it cannot provide a sufficient analysis in the frequency domain. This analysis can only be done as a post-processing step, because it would demand that the complete radiation field is kept in memory for any position z within the undulator. It cannot be realized in the simulation code due to the memory problems discussed above.
For a record set of the radiation field the number of useful entries depends on the simulation and is

$$\tilde{M} = \begin{cases} M & \text{Simulation covers complete bunch} \\ M - \frac{L_U}{\Delta z} & \text{Simulation covers subsection} \end{cases} \quad . \tag{3.55}$$

According to Nyquist's theorem the sampling frequency of $f_s = c\lambda_U/\Delta z\lambda$ limits the observed frequency range to $\pm f_s/2$ around the center frequency $f_0 = c/\lambda$. Higher frequency components cannot be resolved and are converted to lower frequency values. This phenomena is called 'aliasing'. The typical FEL bandwidth is roughly $2\rho f_0$ which should lie within the frequency range to avoid the aliasing. If the integration step length Δz is properly chosen by Eq. 3.53 the frequency range is always on the safe side due to $f_s \gg 4\pi\rho f_0$. Problems arise only in the start up phase, where the spectrum is dominated by the broad-band spontaneous radiation – the seeding of the SASE FEL.
Another issue is the frequency resolution. With a total sample time of $\tilde{M}\lambda\Delta z/c\lambda_U$ the resolution of the normalized frequency is $\Delta f/f_0 = \lambda_U/\tilde{M}\Delta z$. If the FEL amplification bandwidth $\propto 2\rho$ should be resolved by at least four points the number of slices must be set to

$$\tilde{M} \geq \frac{2\lambda_U}{\rho\Delta z} \quad . \tag{3.56}$$

The lower limit of this number can be roughly estimated. By using Eq. 3.53 the minimum number must be at least $\tilde{M} \gg 8\pi \approx 25$. Although 100 slices are sufficient to resolve the spectrum there is often necessary to increase the number. One reason might be to resolve the spikes in the spectrum (see Section 2.7). Their width is defined by the electron bunch length and for a sufficient resolution the total number of slices $\tilde{M}$ should cover the complete bunch.

Chapter 4

GENESIS 1.3 - A 3D Time-Dependent Simulation Code

One of the major aspects of the work for this thesis has been the development of a Free-Electron Laser (FEL) simulation code including all dimension and time-dependent effects such as Self-Amplified Spontaneous Emission Free-Electron Laser (SASE FEL) radiation. The code is named GENESIS 1.3.

GENESIS 1.3 is based on the self-consistent FEL equations Eqs. 3.1 – 3.6, 3.10 and 3.11 of Chapter 3 with no further approximations or assumptions. The electron beam is represented by macro particles and the radiation field is discretized on a Cartesian mesh, using the alternating-direction implicit integration method to solve Maxwell's equations. The electrostatic field is evaluated on a secondary, radial mesh, centered on the electron beam. GENESIS 1.3 has its origin in TDA3D [22], a three dimensional axi-symmetric steady-state simulation code, but there is nothing in common except for the memory efficient 4th order Runge-Kutta integration of the macro particle differential equations.

The user has the option to generate the initial distribution of the radiation field and macro particles as well as the magnetic field of the undulator internally, or to supply the explicit description of these parameters by additional input files. This feature can easily be extended to an interface to codes tracking the electron beam through the linear accelerator to the entrance of the undulator. In addition complicated undulator designs such as an arbitrary tapering of the undulator field or non-periodic focusing structures can be covered.

Any results of numerical calculations have to be critically reviewed because the algorithm used might have some intrinsic weakness or errors, in particular if the problem to be solved is as complex as the FEL process. To estimate the accuracy or correctness of the code, special runs have to be performed to check the results against either the theoretical analysis, other existing and tested codes or experimental results. These benchmarks for the code GENESIS 1.3 are the topic of this chapter.

This chapter is divided into three sections regarding the level of benchmarks. Section 4.1 excludes the FEL interaction between radiation field and electron beam to test parts of the basic FEL algorithm independently. Namely they are the simulation of the initial phase fluctuation of the macro particles and the integration of the radiation field equation for the free space

	TTF-FEL (Phase I)	UCLA
Electron Beam		
Energy [MeV]	230	18
Energy spread [keV]	500	45
Charge [nC]	1	0.3-2.2
Bunch length [μm]	240	900 – 1650
Peak current [A]	500	40 – 170
Beam Size [μm]	66	115 – 145
Undulator		
Period [mm]	27.3	20.5
Peak Field [T]	0.50	0.54
Undulator Parameter	0.894	0.743
Modules	3	1
Length per module [m]	4.5	2.0
Focusing	Quadrupole lattice	Canted poles
FEL-Radiation		
Resonant wavelength [μm]	0.12	13
1D FEL parameter [10^{-3}]	4.0	17.7
Diffraction parameter	1.1	0.11

Table 4.1: *Parameters for the Free-Electron Laser at the TESLA Test Facility (Phase I) and of the UCLA/LANL/RRCKI/SSRL experiment, used for the simulations.*

propagation. Section 4.2 discusses the dependence on the simulation control parameters such as number of macro particle, mesh size and integration step size on the FEL simulation results. Under certain approximations and slight modifications of GENESIS 1.3 the results obtained can be cross-checked with theoretical results. This comparison as well as the results of simulating the UCLA/LANL/RRCKI/SSRL experiment [16] on high gain SASE FEL are found in Section 4.3. This proof-of-principle experiment of starting an FEL from the initial fluctuation of the electron position is briefly described in this section.
Beside the UCLA experiment the simulations are closely related to the Free-Electron Laser at the TESLA Test Facility [99]. All relevant parameters for the benchmarks are listed in Tab. 4.1. A more detailed description of the TTF-FEL is given in the next chapter.

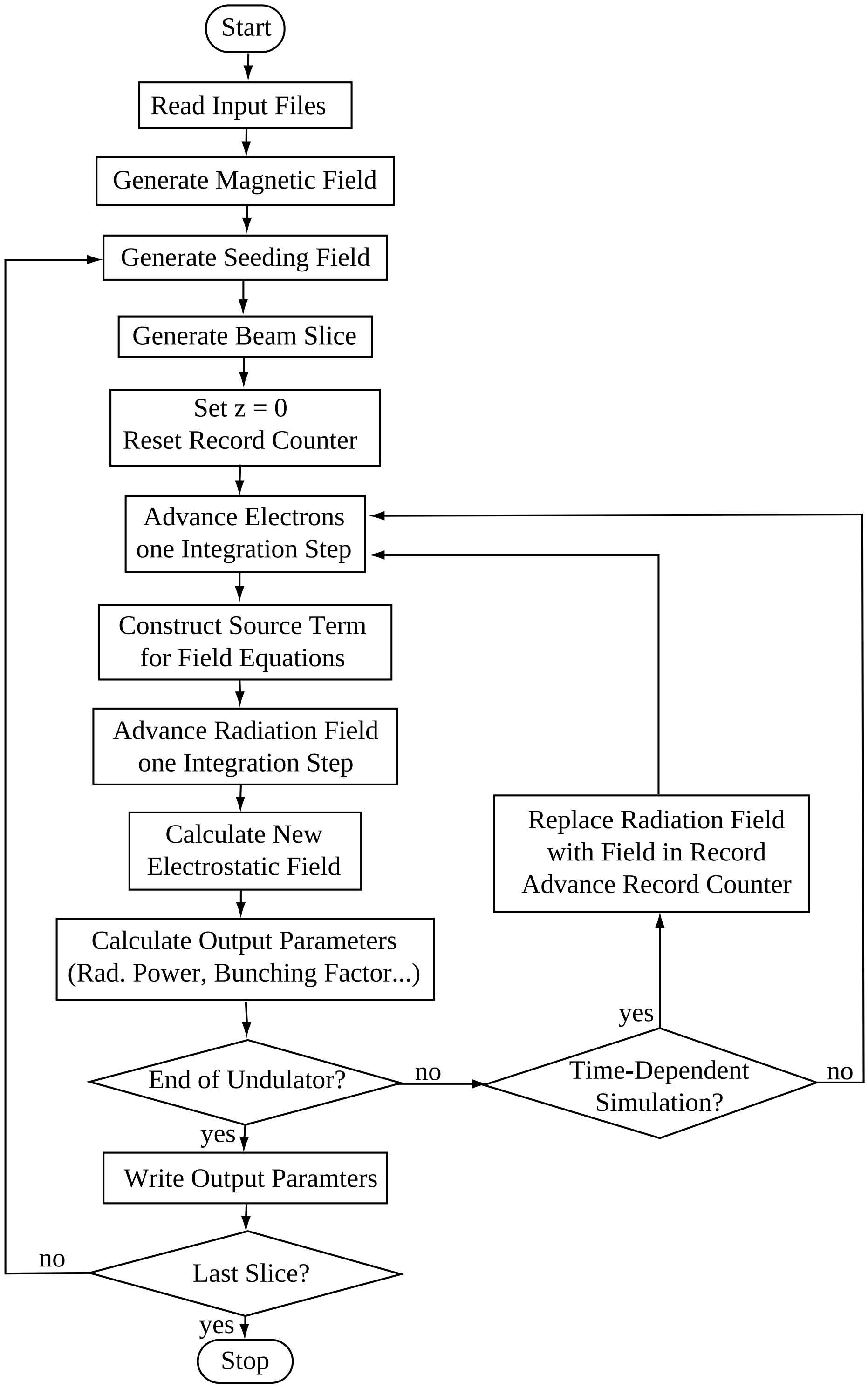


Figure 4.1: *Basic flow chart of GENESIS 1.3.*

4.1 The Basic Parts of the GENESIS 1.3 Algorithm

Out of the four major blocks of the algorithm (see Chapter 3), which construct the GENESIS 1.3 simulation code, only two are of particular interest for the benchmark tests. The memory management for time-dependent simulation does not need to be regarded here, as well as the Runge-Kutta integration of ordinary differential equations, which is a standard method, well tested and commonly used in various types of simulation codes. The general flow chart of the code algorithm is shown in Fig. 4.1.

The discussion in this section is focused on the remaining two parts, loading the phase space distribution for time-dependent simulation and the performance of the radiation field integration algorithm.

The loading of the particle phase space uses Hammersley sequences [83] to fill the electron phases uniformly and to generate a Gaussian distribution in energy. Four independent Hammersley sequences fill a four dimensional unit sphere for the transverse variables. The unit-hypersphere is further transformed into the final distribution.

This homogeneous loading provides residual mean values and correlations of the distributions which are in the range of the computational accuracy unless the number of macro particles is not chosen large enough. In particular the correlation between the longitudinal variables, energy and electron phase, is completely removed. The bunching factor remains zero at any transverse position over the entire undulator length if the FEL interaction is artificially disabled. This is the fundamental demand to cover FEL simulations from the start-up regime up to saturation. Special care has to be taken for the modification of the electron phase θ in order to simulate the initial fluctuation in the longitudinal position correctly. A good criterion to either accept or reject the algorithm is the resultant statistics. It must reproduce the underlying assumption of completely random phases. The bunching factor, $b = < \exp(-i\theta) >$, is the mean value of the electron phasors, whose phase has to follow a uniform distribution, while the probability of the absolute square value $|b|^2$ is given by a negative exponential distribution.

Two different algorithms, presented in Section 3.2, have been tested in order to choose the one with the best performance. They differ either by applying a small random phase shift to all macro particles or by the direct rearrangement of groups of four. In Fig. 4.2 the distributions for the phase and the absolute square of the bunching factor are presented. The values of 2000 independent samples have been taken and split into ten equidistant bins. Both models agree with the theoretical prediction, also drawn in Fig. 4.2, within the range of the fluctuation $\sqrt{L}$ for L samples per bin. With typically 200 counts per bin the fluctuation is about 7%.

The requirement for the correct mean value of $|b|^2$ is more stringent. Unlike the shape of the probability distribution, which is a more general demand, the mean value obtained for the absolute square of the bunching factor should be the inverse of the number of electrons to be simulated. If the transverse motion of the electron is included in the simulation, as is the case for GENESIS 1.3, the mean value must be correct not only for the total ensemble of macro particles but also for any subset. The smallest possible subset is four macro particles (one particle and its three mirror particles) as the fundamental base of the phase fluctuation for $\tilde{N}$ corresponding electrons. Any larger subset of $4k$ macro particles convolute the independent negative exponential distributions for 4 macro particles to the correct distribution. The

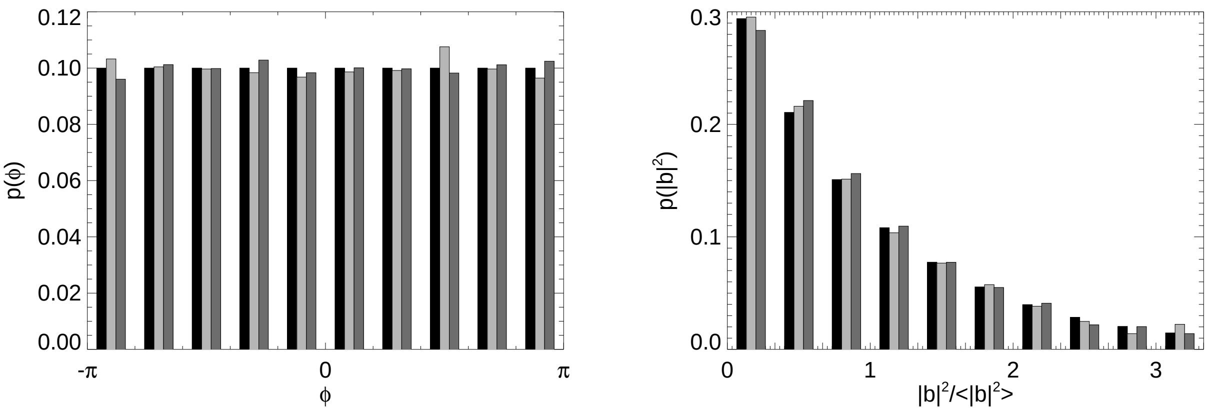


Figure 4.2: *Fluctuation of the initial complex bunching factor $b = |b|\exp[i\phi]$. The probability distribution of the phase (left) and the absolute square (right) are plotted for 2000 independent samples. Two different loading algorithms have been tested: adding a small random phase offset to the macro particles (dark grey bar) and direct arrangement of pairs of four macro particles in phase (light grey bar). The theoretical distributions are plotted by black bars.*

averaged value of the absolute square of the bunching factor is $1/k\tilde{N}$.
For over 200 independently loaded electron beams the macro particles were distributed into radial bins of arbitrarily chosen width. The corresponding numbers of electrons were extracted from the statistic of the bunching factor and compared to the number of macro particles assigned to each bin. The result is shown in Fig. 4.3. Again, both methods have been tested and are in good agreement with the theoretical prediction.
Regarding the test results the two methods are equivalent, even if the computational time is included in the overall efficiency. The execution times needed to apply the fluctuation are almost identical, but short compared to the homogeneously filling of the 6D phase space itself, which takes about 96% of the time. In the final version of GENESIS 1.3 the algorithm of random phase offsets is implemented. The reason is motivated not physically but practically, because the source code is slimmer and easier to understand. The random phase shifts resemble, at least partially, the idea of the pure initial random phase of the electrons.
The most critical part of the simulation is solving the partial differential equation of the radiation field. Although the alternating-direction implicit method is stable (see Section 3.4) the discretization itself has a significant influence on the performance of the integration. A coarser grid yields faster execution but excludes details in the radiation field due to lower resolution.
A standard problem is the free space propagation of the radiation field which can be compared with analytic results. For this reason the radiation field is decomposed into the set of orthonormal Hermite-Gaussian eigenfunctions. The analytic discussion can be found in Appendix A. All modes are completely defined by three parameters: the radiation wavelength λ, the Rayleigh length z_R and the longitudinal position z_0 of the radiation waist. The physical meaning of the Rayleigh length is the characteristic length of the radiation field diffraction. The two indices

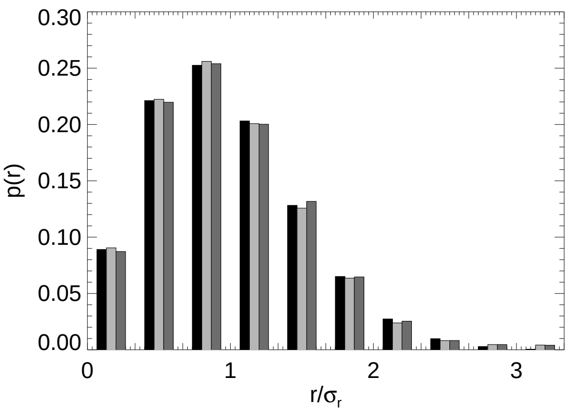


Figure 4.3: *Radial electron density distribution, extracted from the mean value of the absolute square value of the bunching factor. The two different loading algorithms by directly arranging the macro particles or by random electron phase offset are given by light and dark grey bars, respectively. The theoretical distribution is indicated by the black bars.*

n and m for the TEM_{nm}-mode considered indicate the number of nodes of the radiation field amplitude in x- and y-direction, respectively.
In the left plot of Fig. 4.4 the root-mean-square radius of the radiation intensity profile from the lowest TEM_{00}-mode up to the TEM_{33}-mode is plotted. The result for the TEM_{00}-mode, which corresponds to the curve with the smallest diffraction, hardly differs from the analytical prediction

$$\sigma_r = \sqrt{\frac{\lambda z_R}{\pi}}\sqrt{1+\frac{z^2}{z_R^2}} \quad , \tag{4.1}$$

also drawn in this plot. Some of the curves of higher modes coincide, namely for those, where the sum $n+m$ is constant. This is in good agreement with the analysis, predicting the scaling of the root-mean-square radiation radius as

$$\sigma_{nm} = \sqrt{n+m+1}\sigma_{00} \quad , \tag{4.2}$$

where the root-mean-square size σ_{00} of the fundamental mode is given by Eq. 4.1.
The right hand plot of Fig. 4.4 shows the variance of the radiation beam radius at an arbitrarily chosen position z for each mode separately. The dependence of the variance on n and m is linear, being consistent with Eq. 4.2. For all modes the radiation power is conserved.
So far the numerical results for lower modes reach the desired level of accuracy. Problems are only encountered for higher values of n and m, because the diffraction is stronger and the transverse profile is more structured. The first problem is solved by defining a larger grid. If the grid is not large enough a significant part of the radiation field is reflected backwards at the edge due to the Dirichlet-boundary condition of the grid as it is shown in Fig. 4.5. For

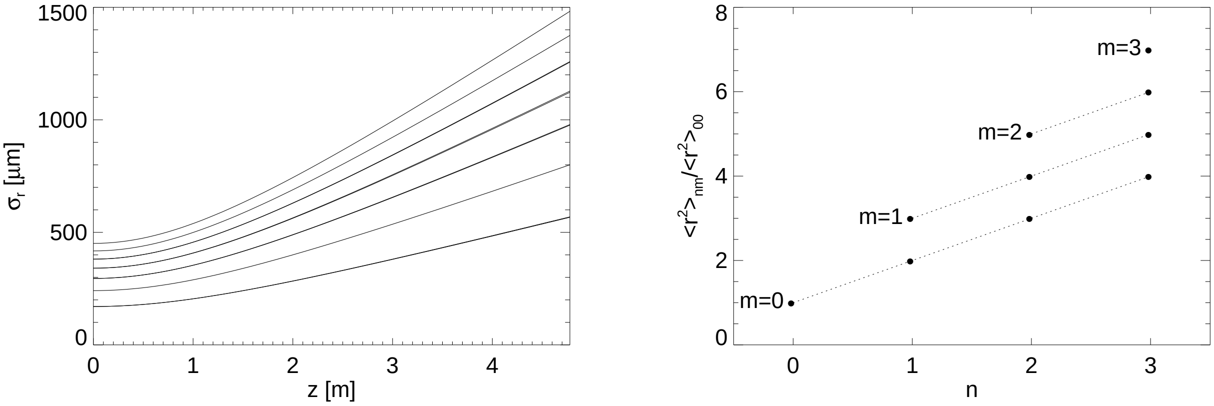


Figure 4.4: *Root-mean-square radius of the radiation intensity profile of the lowest Gauss-Hermite modes (TEM_{00} – TEM_{33}) for free space propagation (left). Curves for the TEM_{nm}-modes with $n + m = const$ coincide in this plot. A coincidence occurs also for the analytical result and the TEM_{00}-curve. To indicate the order the variance $< r^2 >_{nm} \equiv \sigma_r^2$ of the radius is drawn for the different modes (right).*

the simulation the grid has been limited to enclose five standard deviations in positive and negative direction of the initial Gaussian distribution. During the propagation the expanding radiation field reaches the edge of the grid. The reflected radiation field adds up coherently with the remaining field and interference patterns are visible in the intensity profile (right plot of Fig. 4.5). Thus, the results obtained are physically correct, they do not represent free space propagation but rather that in a rectangular waveguide. To avoid boundary problems the mesh size has to be chosen sufficiently large.
The other problem is the resolution of the mesh, which can be roughly estimated. Regarding a mode with many nodes in one direction ($n \gg m$) the size of the radiation field depends on n as $\sigma_{nm} \approx \sqrt{n}\sigma_{00}$. Therefore the average separation of the nodes converges as $\sigma_{00}/\sqrt{n}$ towards zero. Due to the discrete nature of the grid the resolvable frequencies are limited to $1/2\Delta$, where Δ is the distance between two adjoining grid points. The resolution is roughly estimated by

$$n < \frac{\lambda z_R}{4\pi\Delta^2} \quad . \tag{4.3}$$

The problem with cutting off higher modes arises only if a small term evolution of the radiation field is desired. As an example the radiation field propagation of a plane wave after passing a small quadratic aperture is calculated. The intensity profiles immediately after the aperture and much further downstream are plotted in Fig. 4.6. While the distribution further away already approaches the far field distribution, immediately after the aperture the near field distribution has the typical fringes at the edges. The typical size of the fine structure is limited by the resolution of the grid (Eq. 4.3). Although it seems to be paradox, a larger integration length

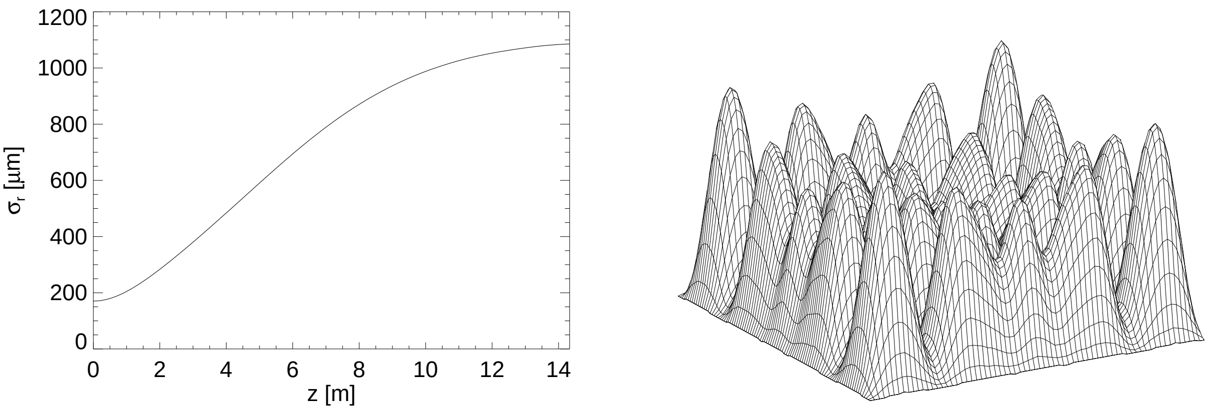


Figure 4.5: *Radiation field for the free space propagation using a too small grid. The dependence of the root-mean-square radius of the radiation intensity profile on the longitudinal position z is shown in left plot, the intensity profile at the end of the integration length in the right plot (arbitrary units).*

exhibits a better accuracy than a shorter length for these kind of problems.
Unlike the number of grid points the question of the optimum choice for Δ cannot be definitely answered and depends on the problem, on the required resolution and on the acceptable computational time.

4.2 The Performance of FEL Simulations

The next step of testing GENESIS 1.3 is the dependence of the simulation on control parameters, which are namely

- the integration step length Δz,
- the grid resolution Δ,
- the total number of grid points N_G,
- the number of macro particles N_M.

Out of these four parameters N_G has to be chosen in such a way that during the FEL simulation no significant amplitude of the radiation field will reach the boundary of the grid and will be reflected backwards. Compared to the free space propagation of the radiation field this constraint is less restricting because the FEL amplification compensates the diffraction by the growth of the radiation field amplitude at the location of the electron beam. Depending on the gain of the FEL the radiation field converges towards a constant transverse profile in the linear regime of the amplification process. It is referred to as 'gain guiding' (see Section 2.6). In the

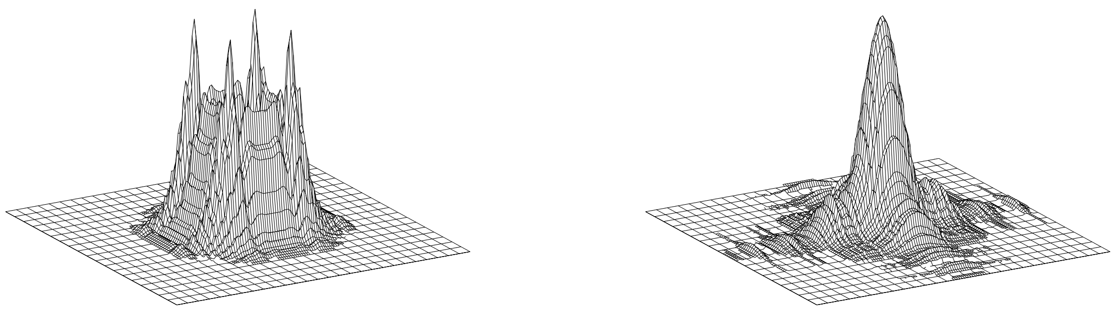


Figure 4.6: *Radiation intensity profile immediately after (left) and further beyond of (right) a small quadratic aperture, being passed by a plane wave.*

saturation regime the strength of this guiding steadily reduces and the diffraction approaches that of the free space propagation. The global size of the grid and thus the total number of grid points N_G can be obtained by a rough estimation of the size of the gain guided modes. A size of $6-8$ times the root-mean-square size of the radiation field in the linear regime of the FEL amplification provides valid results.

The simulations to study the influence of the remaining three parameters are based on the parameter set of the TESLA Test Facility Free-Electron Laser. For simplicity the drift space between the three undulator modules has been ignored for the simulation. For these calculations only the dependence on the simulation parameters is needed and the TTF-Free-Electron Laser, which is intended to operate as a SASE FEL, is approximated by an FEL amplifier. The fluctuation in electron position is replaced by an equivalent seeding field. The initial radiation power for a 230 MeV electron beam is 2.8 W and the root-mean-square radius of the seeding radiation field is $\sigma_r = 140\mu$m. For all simulations the total gain of the radiation power is seven orders of magnitude with a saturation length of roughly 14 m.

The grid resolution is the most sensitive parameter. The results for the overall exponential gain factor $G = \log(P_{out}/P_{in})$ are shown in Fig. 4.7. All values have been normalized to that obtained with the finest resolution. The radiation output power has been evaluated at a fixed position z in the undulator shortly before reaching saturation. Using the saturation power as the output power would flatten the curve in Fig. 4.7, because the saturation level drops only about 9% while the saturation length grows about 10%, going from the finest to the coarsest grid. For a grid resolution, which is finer than one quarter of the radiation size σ_r, the variation of the exponential gain factor is almost negligible. The resolution is set below 0.2 standard deviation of the radiation field for further calculations.

To analyze the impact of the integration step length Δz, the energy conservation is the best parameter to study. Due to the underlying model of self-consistent FEL equations the total energy of the radiation field and the electron beam is conserved. This should be reproduced by

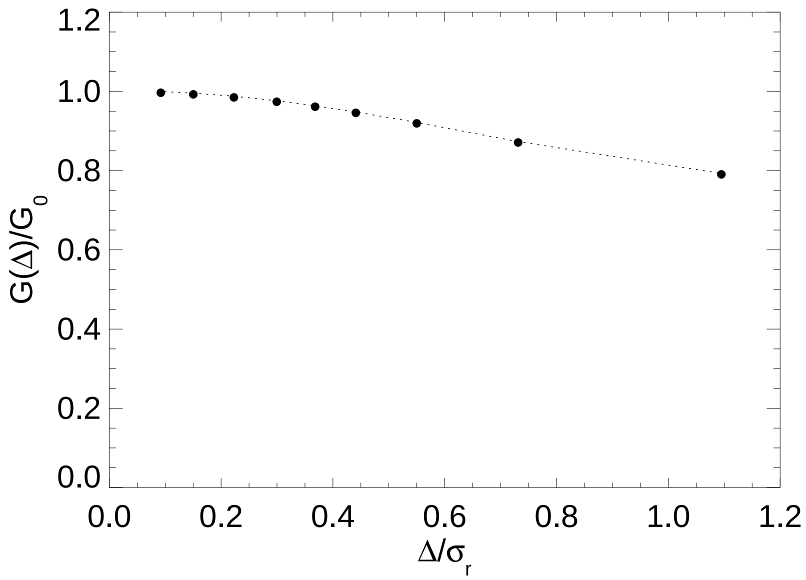
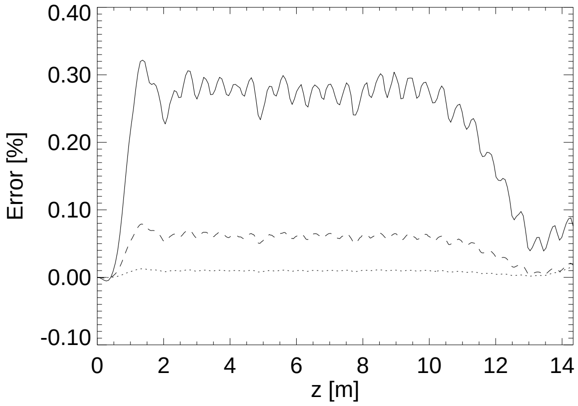


Figure 4.7: *Average exponential gain $G = \log(P_{out}/P_{in})$ of the radiation power for different grid resolution. The values are normalized by the result for the finest resolution with a gain of $G_0 = 5 \times 10^6$. The radiation size is $\sigma_r = 140\mu m$.*

the simulation code, but small numerical errors can occur, which might add up and modify the results. As discussed in Section 3.4, the integration of the radiation field equation will neither amplify any errors in the discrete values of the radiation field, nor will it damp them. Therefore it can be assumed that the error is proportional to the radiation field amplitude. GENESIS 1.3 normalizes the error in the energy conservation as

$$\text{error} \equiv \frac{[P_{rad}(z) + P_{beam}(z)] - [P_{rad}(0) + P_{beam}(0)]}{P_{rad}(z)} \tag{4.4}$$

with P_{rad} as the power of the radiation field, $P_{beam} =< \gamma > Imc^2/e$ as the electron beam power, $< \gamma >$ as the average energy of the electrons and I as the current of the electron beam. The error for different step sizes is shown in the left plot of Fig. 4.8. All curves exhibit the same dependence, indicating an excess of the total energy compared to the initial value. The magnitude of the error depends quadratically on the step size Δz used. The shape of the curves has similarities with that of the growth rate of the radiation field $(1/P)dP/dz$, shown in the right plot of Fig. 4.8. The reason lies in the leapfrog integration algorithm. To calculate the energy conservation GENESIS 1.3 interpolates the radiation power to the position z, where the electron parameters are defined (see Section 3.3). The linear interpolation overestimates the radiation power, which grows exponentially. For rather small step sizes the difference, and thus the error in the energy conservation, is quadratic in Δz. The actual numerical error, excluding this systematic error, is at least one order of magnitude smaller. For the simulations performed



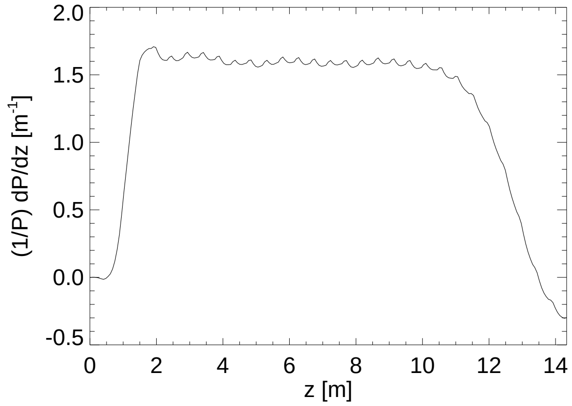


Figure 4.8: *Error in the energy conservation for different integration step sizes Δz of 2.5, 1.25 and 0.5 times the undulator period (solid, dashed and dotted line in the left plot, respectively) and the growth rate of the radiation power (right) along the undulator.*

the chosen step length had no significant impact on the results. The upper limit for the step length is defined by the focusing lattice geometry. A larger value of Δz would no longer resolve the individual quadrupoles.

The last parameter study concerns the number of macro particles used. Again, energy conservation and radiation power are the best performance indicators for the simulation. The error of the energy conservation and the saturation power are shown in Fig. 4.9. Unlike the results of varying the integration step length, the error for the smallest number of macro particles differs significantly from the systematic error, caused by the linear interpolation of the radiation power (see above). The true numerical error dominates and has a value of about 2%. Increasing the number of macro particles the curves converge towards that of the systematic error in Fig. 4.8. For number larger than 2000 particles the influence on the simulation results is rather weak. This can be clearly seen in the plot of the saturation power.

Summarizing, any influence of the control parameters on an FEL simulation is almost excluded if the grid resolves the radiation field with more than 3 points per standard deviation of the transverse size. A total number of 2000 macro particles or more is sufficient. The length of the integration step is limited rather by the undulator model with its focusing structure or gaps between two modules than by the numerical precision.

Concluding this section the relative computational time for the different tasks of a typical run is shown in Fig. 4.10. For the simulation 2^{13} macro particles and 151×151 grid points have been used. For the electrostatic field only the lowest order in the longitudinal and azimuthal Fourier series decomposition has been included. Most of the time is spent in solving the electron equations of motion, followed by the calculation of the radiation and electrostatic fields. Compared to the 22000 grid points of the Cartesian mesh for the radiation field, solving the matrix equation for 40 grid points of the radial grid, used for the electrostatic field, is highly inefficient. The reason lies in the conversion of the Cartesian coordinate system of the

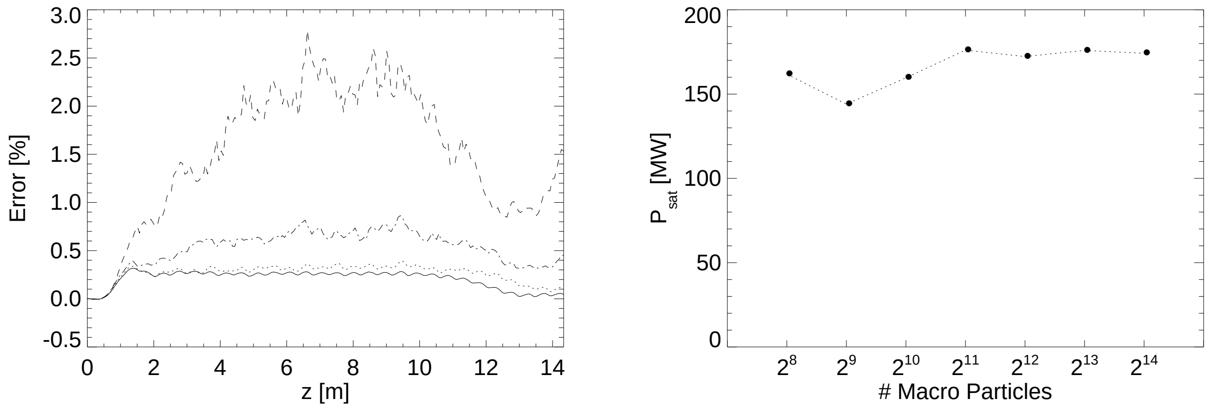


Figure 4.9: *Error in the energy conservation (left) and saturation power (right) for different numbers of macro particles (dashed $\equiv 2^8$, dashed-dotted $\equiv 2^{10}$, dotted $\equiv 2^{12}$ and solid $\equiv 2^{14}$ macro particles in the left plot).*

macro particles to or back from the cylindrical system of the electrostatic field. Therefore the calculation time for this algorithm scales rather with the number of macro particles for the construction of the source term than with the number of grid points to solve the field equation. Two additional source terms have to be calculated for each higher mode with time consuming evaluations of trigonometric functions for each macro particle. The time required for code initialization is almost negligible. This includes the generation of the particle phase space distribution and the discretization of the initial radiation field. The diagnostic part, where output values such as radiation power and size are calculated, does not take too much time relative to the core algorithm of solving the FEL equations.

4.3 Comparison with Analytical and Experimental Results

Compared with the studies of the control parameters in the previous section, a more important aspect is the reliability of GENESIS 1.3, whether the results obtained agree with the analytical solution in the frame of the underlying FEL model. This section is devoted to cross-check GENESIS 1.3 with the theory as well as with the experimental results for the UCLA/LANL/RRCKI/SSRL experiment on high gain SASE FELs.

If transverse betatron motion of the electrons is included it is difficult to assess the results due to the lack of analytic solutions. Therefore any transverse motion has been disabled in GENESIS 1.3 by artificially setting the emittance and any quadrupole field components to zero. This model is identical with that used for the theoretical analysis in Section 2.6. The code FS2R [100] is a well tested numerical tool, providing the analytical results with high accuracy.

The most sensitive parameter is the growth rate Λ of the TEM_{00}-mode. This fundamental

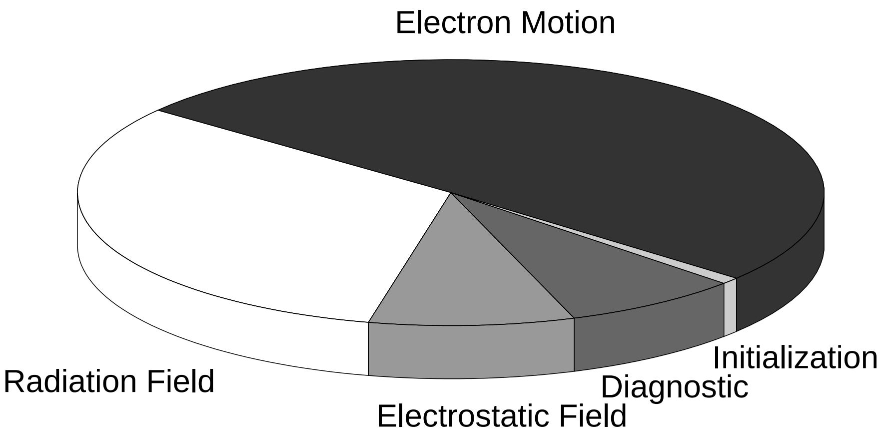


Figure 4.10: *Relative calculation time for the different tasks in GENESIS 1.3.*

mode is independent of the underlying coordinate system of the radiation field in analogy to the Hermite-Gaussian and Laguerre-Gaussian mode decomposition for the free space propagation. GENESIS 1.3, as a simulation code, which is rather related to the explicit design of the FEL, determines the growth rate as the logarithmic increment of the radiation power P. It is related to the theoretical definition in the exponential regime of the FEL amplification by

$$\frac{1}{P}\frac{dP}{dz} = \frac{8\pi\hat{\rho}\Lambda}{\lambda_U} \quad , \tag{4.5}$$

where $\hat{\rho}$ is the 3D FEL parameter (Eq. 2.98) and λ_U is the undulator period. Using the parameters of the TTF-FEL (phase I) with a root-mean-square beam size of $\sigma_{x,y} = 66.6$ μm of the Gaussian distributed electron beam, the results for varying the radiation wavelength or the energy spread are shown in Fig. 4.11.

The overall agreement is good with a maximum deviation of less than 1%. Another important aspect is the size and profile of the radiation field in the Fresnel and Fraunhofer zone (near and far field, respectively). The intensity profiles in both zones give sufficient information about the radial dependence on the radiation amplitude and phase. Again GENESIS 1.3 agrees well with the analytical prediction as is shown in Fig. 4.12.

Extending the code tests to higher modes is rather difficult for two reasons. Firstly, the simulation code does not treat modes independently and the fundamental mode always has a major impact on the radiation field amplification, exhibiting the largest growth rate. Secondly, the transverse motion affects in particular the higher modes of the FEL process. The first problem is addressed by choosing a larger beam size. The reduced diffraction enhances the growth rate of the higher modes and thus the fundamental mode becomes less dominant. Although the TEM_{10} Hermite-Gaussian mode and the FEL eigenmode are not the same, it can be expected that the overlap of these two modes is large, determining the magnitude of the eigenmode (Eq. B.21). Therefore the simulation is seeded by this Hermite-Gaussian mode.

GENESIS 1.3 has calculated the amplification with and without transverse motion. For the latter a zero emittance and no transverse focusing have been chosen. The root-mean-square

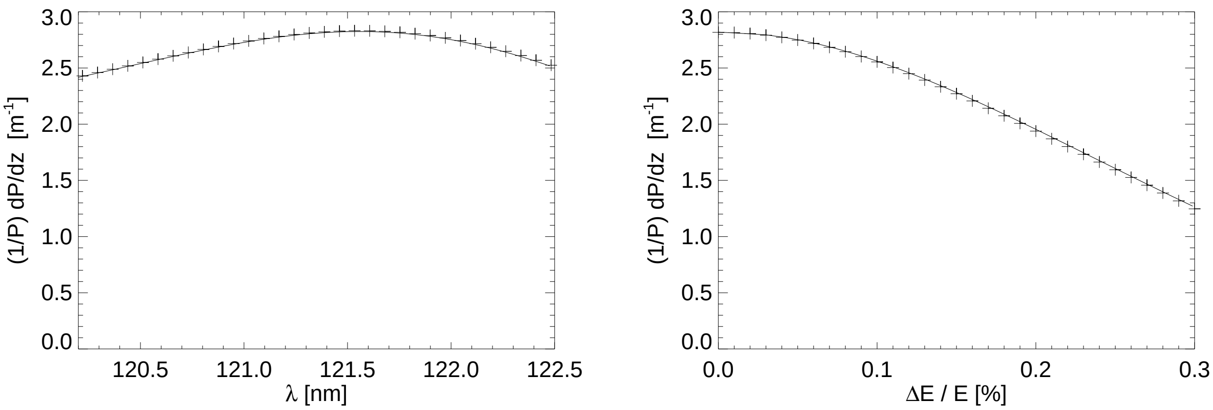


Figure 4.11: *Growth rate of the radiation power for different radiation wavelengths (left) and energy spreads (right), respectively. Results of GENESIS 1.3 are indicated by markers, those of the analytic solution (FS2R) by a solid line.*

size of the electron beam has been increased to $\sigma_{x,y} = 165\ \mu$m, which reduces the strength of diffraction by about a factor of 6. Including the betatron motion with a transverse normalized emittance of 10 π mm-mrad and a quadrupole gradient of 10 T/m nearly the same beam size has been obtained on average. The enhanced axial velocity spread has been adjusted by a smaller energy spread (see Eq. 2.96).

The radiation power as well as the growth rate are plotted in Fig. 4.13. FEL amplification of the fundamental Gaussian mode exhibits the largest growth, almost independent of the betatron motion. In contrast, if the FEL is seeded with the TEM_{10} Hermite-Gaussian mode the results differ as can best be seen in the plot of the power growth rate. While for the case of excluding betatron motion the growth rate is constant for most of the time, it is strongly modulated if the transverse motion is included. The modulation can be correlated to the oscillation of the beam envelope. In addition, the growth rate tends to decrease between $z = 5$ m and $z = 12$ m before it rises again. This can be explained by the electron betatron motion with a period length of roughly 8 m. The TEM_{10} mode is symmetric in the y-direction but antisymmetric in the x-direction. Electrons, which experience a sign change of the transverse position after half a betatron wavelength, are off-resonance relative to the radiation field by a phase difference of π. The amplification would be inverted to a damping of the radiation field, if the field did not adiabatically follow the changing phase of the bunching factor. Even with this self-adjustment of the radiation phase the amplification is inhibited as it is clearly visible in the right plot of Fig. 4.13.

A comparison of the simulation results for higher harmonics with the analytic model (Section 2.6) is difficult because GENESIS 1.3 is not based on a mode decomposition into the FEL eigenfunctions. To study a certain higher mode the initial seeding field has to be modified in such a way that lower modes are suppressed. Otherwise the radiation power, calculated by GENESIS 1.3, would be dominated by these modes. In addition the betatron motion couples

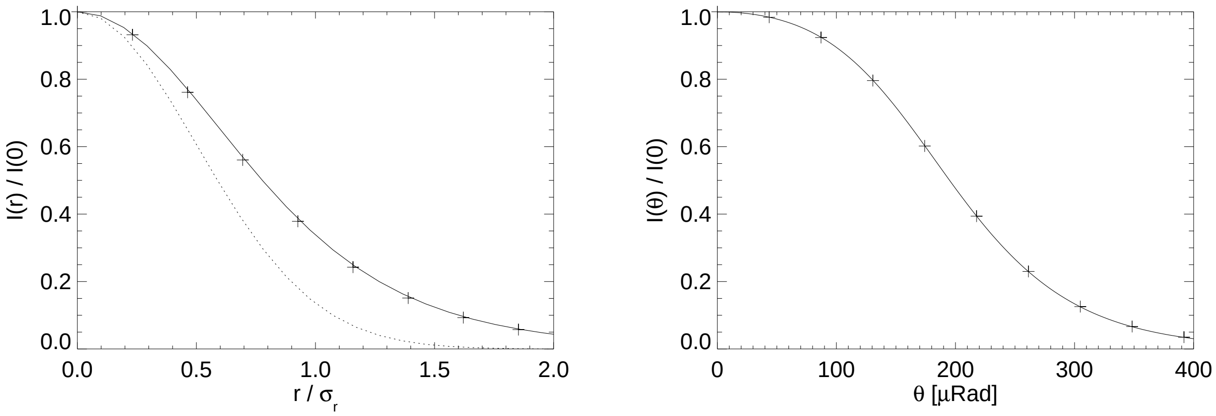


Figure 4.12: *Intensity profile of the radiation field in the Fresnel zone (left) and Fraunhofer zone (right), respectively. The theoretical solution is drawn by a solid line, the results of GENESIS 1.3 are indicated by the markers. In the plot of the Fresnel zone the corresponding electron beam density with arbitrary units is plotted by a dotted line.*

the modes and the fundamental mode is excited regardless of how accurately a higher mode is seeded.

This chapter is concluded with the comparison of GENESIS 1.3 simulations with the experimental data of a high gain SASE FEL experiment. This experiment has been carried out by the collaboration of the University of California in Los Angeles (UCLA), the Los Alamos National Laboratory (LANL), the Russian Research Centre Kurchatov Institute (RRCKI) in Moscow and the Stanford Synchrotron Radiation Laboratory (SSRL). The gain of the amplified spontaneously emitted radiation was measured to be larger than 10^5.

In this experiment the injector was driven by a radio-frequency photo gun and an acceleration structure at a frequency of 1.3 GHz [14]. The electron beam has been boosted to an energy of 18 MeV. The charge of the electron beam was controlled by the power of the rf-gun laser and covered the range between 0.3 and 2.2 nC. The bunch length as well as the transverse emittance depended on the charge, having been measured before the experiment.

The undulator has a length of 2 m with a period length of 2.05 cm and a root-mean-square undulator parameter of 0.74. The electron beam is focused to a size of 115 – 145 μm in both planes by canted pole faces of the undulator magnets. The peak current was in the range of 40 – 170 A.

Using these parameters and adjusting the bunch size and length to the corresponding charge, up to 200 independent runs for four different peak currents have been performed. The averaged energy of the radiation pulse is plotted in Fig. 4.14 for the experimental and simulated results. GENESIS 1.3 underestimates the radiation power in general but agrees very well for large and small values of the electron beam charge. In the mid region at around 1 nC the deviation between simulation and experiment is the largest although it never exceeds one order of magnitude, which is acceptable for an overall gain of $10^4 - 10^5$. In general the dependence

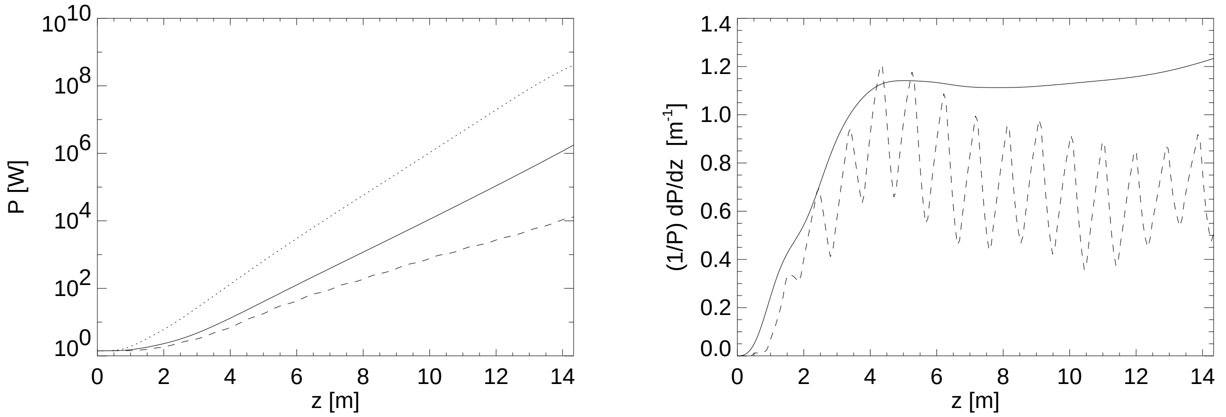


Figure 4.13: *Radiation power and radiation power growth rate (left and right, respectively) for different Hermite-Gaussian modes as the seeding field (dotted line: TEM_{00}, solid line: TEM_{10} with disabled betatron motion, dashed line: TEM_{10} with betatron motion).*

of the radiation pulse energy on the charge, predicted by GENESIS 1.3, is smoother than the observed data. This is in coincidence with the results of FAST [77], another time-dependent FEL simulation code, which has been used to study this experiment in more detail [101].

Typical radiation pulse profiles at different stages of amplification are shown in Fig. 4.15. As is typical for simulations of SASE Free-Electron Lasers the noisy pulse becomes dominated by a few spikes exhibiting the largest growth rate. The initial position of these spikes is random. Averaging 200 independent shots yields a much smoother profile. This profile, where the maximum is shifted in the forward direction relative to the current maximum of the electron beam, is comparable with the radiation profile of an FEL amplifier. For an amplifier the initial fluctuation in the electron density is negligible, because the FEL is started by a seeded radiation field instead.

The number of independent runs for a given charge was not large enough to provide enough statistics. The relative fluctuation of the shot-to-shot radiation energy is approximately 30%. In the experiment this fluctuation is also influenced by the uncertainties of the beam charge, energy spread, beam size and length, which were changing from shot to shot. An estimate of these errors yields a fluctuation of the radiation energy of about 2%. With an expected fluctuation in the radiation power of about 34% this error could hardly be resolved [16].

Ending this chapter it is mentioned that GENESIS 1.3 has been compared with the codes MEDUSA [78], TDA3D, RON [76] and GINGER [98] for the Advance Photon Source (APS) FEL at the Argonne National Lab [102]. The dependence of the beam emittance, energy spread and peak current on the gain length, saturation power and saturation length has been studied in the steady state regime [24]. The results of the various codes agree well with each other. The ongoing work will include time-dependent simulations as well as beam trajectory errors.

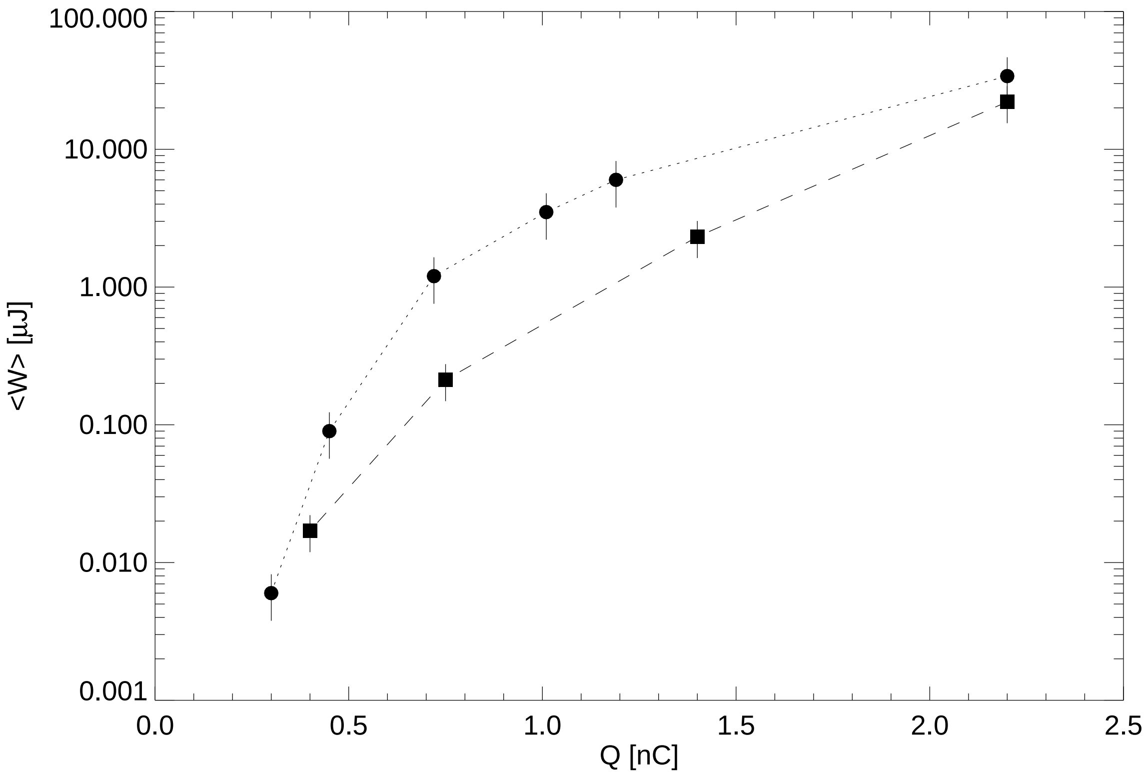


Figure 4.14: *Average radiation pulse energy versus the electron beam charge for the UCLA/LANL/RRCKI/SSRL experiment on SASE-FELs and the numerical results of GENESIS 1.3 (circle and box markers, respectively). The error bars indicate the fluctuation in the radiation power.*

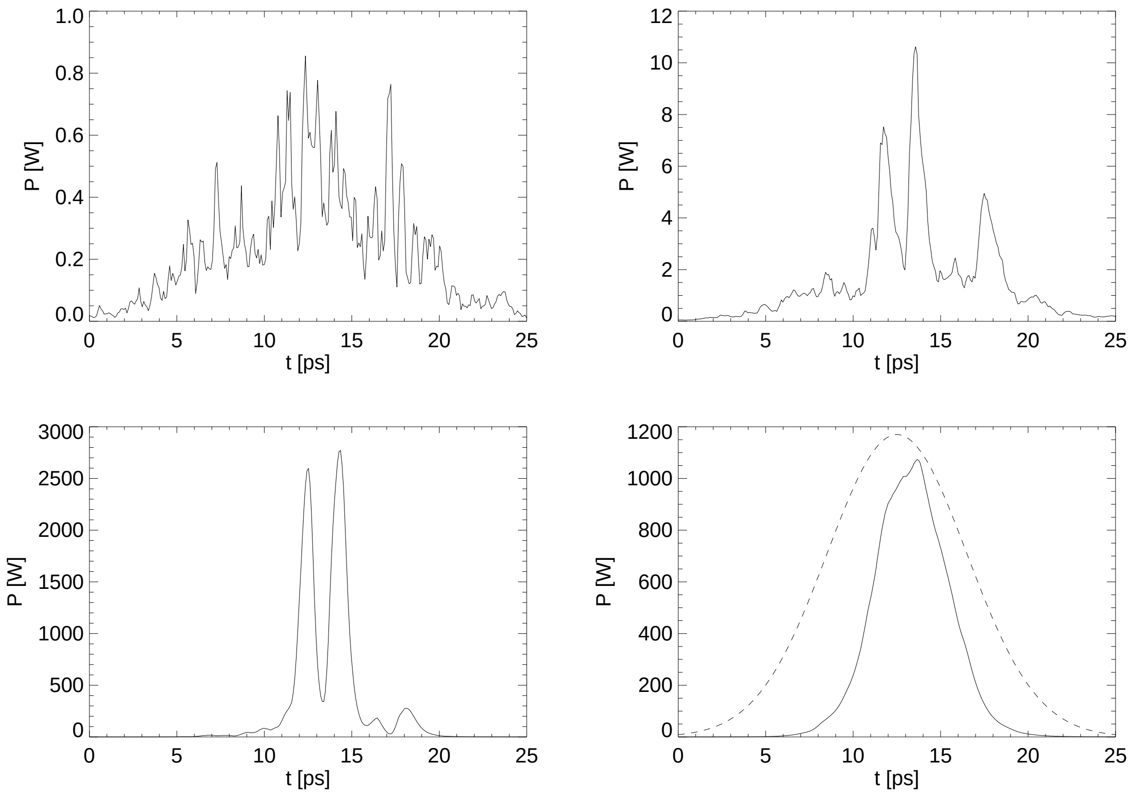


Figure 4.15: *Radiation pulse at 20 cm (upper left), 80 cm (upper right) and at the exit of the undulator (lower left), respectively. The average result at the end of the undulator of 200 independent shots is shown in the lower right plot. The electron beam current profile is drawn by the dashed line.*

Chapter 5

The VUV Free-Electron Laser at the TESLA Test Facility

One of the on-going projects at DESY is the construction and operation of the TESLA Test Facility (TTF) [99], a linear accelerator based on superconducting cavities. A multinational collaboration has been established for developing and testing new technologies in accelerator physics towards the design and realization of the future linear collider TESLA (TeV Superconducting Linear Accelerator) [20] with a center of mass energy of 500 GeV.
Both projects integrate a Free-Electron Laser. The beam parameters for driving an FEL are similar to those of the linear accelerator. Only some minor modifications have to be applied such as a reduced bunch charge. After a brief introduction into the TESLA Test Facility and TTF-FEL, two sections are dedicated to special problems of the FEL operation: the coherent transverse motion of the electron beam within the undulator in Section 5.1 and the longitudinal beam energy modulation due to wake fields in Section 5.2. The discussion of the TESLA X-Ray FEL at a beam energy of 10 – 50 GeV is postponed to Chapter 6.
The TTF linac operation is planned in two phases. They differ in the number of accelerating modules and thus in the possible electron beam energy for driving the Free-Electron Laser. An accelerating module consists of eight 9-cell superconducting cavities, made out of Niobium. These modules are a fundamental component of the TESLA Test Facility as well, as of the TESLA linear collider
With an accelerating gradient of 15 – 25 MV/m, beam energies from 230 MeV to 390 Mev can be achieved with the three modules in phase I of the TTF linac. An rf-photoelectron gun, operating at a frequency of 1.3 GHz, provides a train of short electron bunches. A bunch compressor shortens the root-mean-square bunch length down to 250 μm.
In phase II five more accelerating modules are added to the extended beam line and the bunch length is further compressed by a factor of five. The design beam energy is 1 GeV. A schematic layout of the linac in the two phases are shown in Fig. 5.1.
The undulator has a common design for both phases. Diagnostic blocks are placed between individual undulator modules, as shown in Fig. 5.2. Each module has a length of 4.5 m and consists of 660 permanent magnets and iron poles, which are assembled as hybrid magnets. They produce a peak field of 0.497 T on the undulator axis. Magnetic flux is added in the

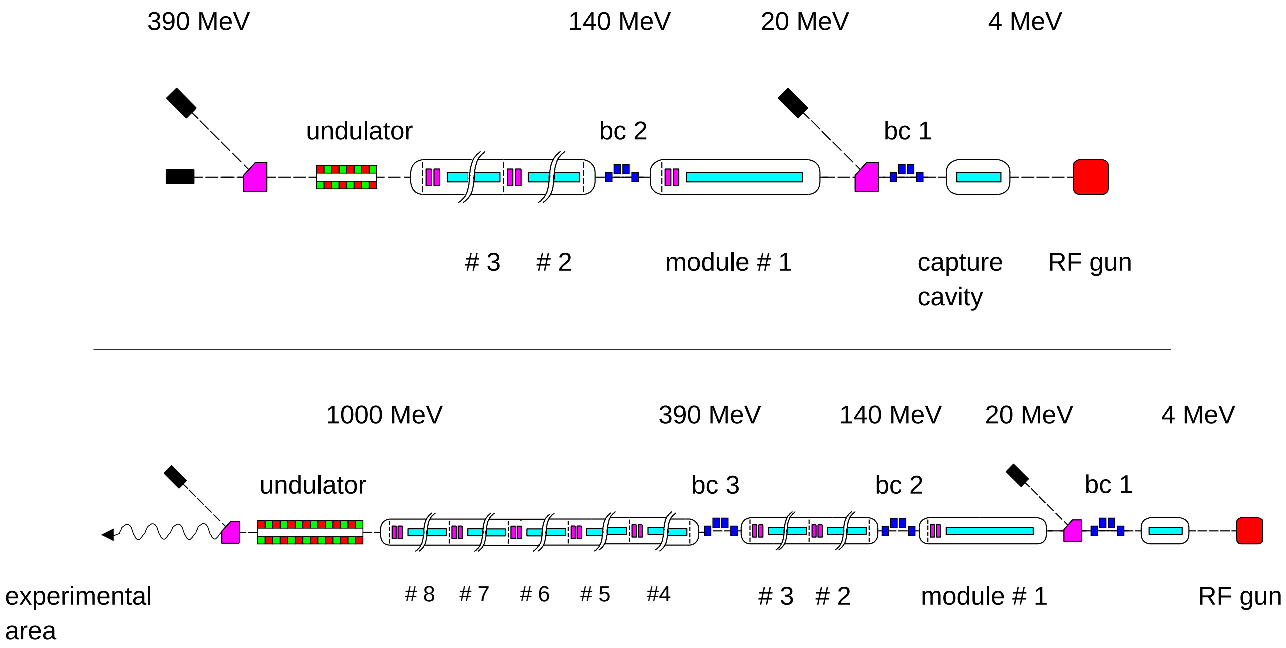


Figure 5.1: *Schematic drawing of the TESLA Test Facility beam line in phase I and II (upper and lower drawing, respectively).*

gap using additional small magnets to form a quadrupole field for focusing. These quadrupoles provide the dominant focusing of the electron beam in addition to the natural focusing of the undulator dipoles. Beam position monitors and steering magnets are located in the diagnostic blocks and are incorporated into the vacuum chamber for use in aligning the electron beam to the undulator axis [103].

For the transition to phase II three more modules are added. The field strength of the quadrupoles is increased in order to optimize the FEL performance. The important parameters are summarized in Tab. 5.1.

An extension of the beam line is planned for both phases. A fraction of the FEL radiation is filtered by a monochromator to improve the longitudinal coherence of the radiation pulse. The radiation is either reflected back to the entrance of the same undulator [104] or guided to an additional undulator section [105]. In both cases the seeding dominates over the shot noise power level of the electron bunch and the radiation characteristics is similar to these of an FEL amplifier.

The general performance of the TTF-Free-Electron Laser, to be operated as a Self-Amplified Spontaneous Emission FEL, is the subject of detailed analytical and numerical studies [99, 100, 106]. Fig. 5.3 shows the typical exponential growth of the radiation pulse energy for a single sample of the initial random distribution of the longitudinal electron position.

Due to the nature of SASE radiation the radiation pulse exhibits many spikes and the fluctuations of the radiation energy follow a Gamma-Distribution (see Section 2.7). The single free parameter of this distribution is related to the average number of spikes per pulse. The

Table 5.1: *Design parameters for the Free-Electron Laser at the TESLA Test Facility.*

Parameter	Phase I	Phase II
Electron Beam		
Energy	230 – 390 MeV	1 GeV
Energy Spread (rms)	500 keV	1 MeV
Charge	1 nC	
Peak Current	500 A	2500 A
Bunch Length	250 μm	50 μm
Normalized Emittance (rms)	2 π mm·mrad	
Beam Size (rms)	66 μm	57 μm
Bunches per Pulse Train	7200	
Repetition Rate	10 Hz	
Undulator		
Type	planar	
Period Length	2.73 cm	
Peak Field	0.497 T	
Undulator Parameter (rms)	0.894	
Total Length	15 m	30 m
Module Length	4.5 m	
Modules in Undulator	3	6
Quadrupole Gradient	12.5 T/m	18.3 T/m
Quadrupole Length	13.65 cm	
FODO Cell Length	95.55 cm	
Radiation		
Wavelength	120 – 40 nm	6.4 nm
Slippage Length	63 – 21 μm	6.8 μm
FEL Parameter (1D)	4.6×10^{-3} – 2.8×10^{-3}	2.1×10^{-3}
Gain Length (1D)	47 – 78 cm	103 cm
Diffraction Parameter	1.1 – 2.3	14.0

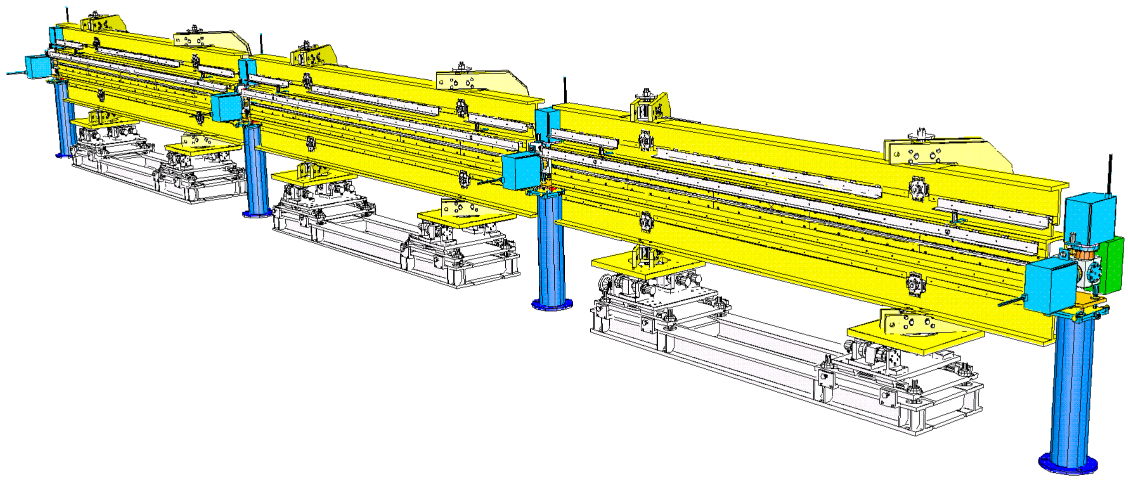


Figure 5.2: *Drawing of the undulator of the TTF-FEL phase I including the undulator modules (light grey) and the diagnostic blocks (dark grey).*

evolution of these spikes can be seen in Fig. 5.4, where the radiation power is plotted as a function of the position within the bunch and in the undulator. The exponential growth of the power makes it difficult to compare the longitudinal radiation power profile at saturation with that in the start-up regime. Therefore the radiation power profile has been normalized to the maximum power at each position along the undulator. Although the TTF-FEL of phase I consists of only three modules, an additional module is included in the simulation to cover the deep saturation regime.

In the start-up regime ($z < 4$ m) the difference in power level of the spikes and the pedestal radiation level is small as indicated by the low contrast in the bottom part of Fig. 5.4. This is caused by the incoherence in the transverse plane, where several independent 'hot spots' – points of high emission due to localized bunching of the electron beam – exist.

These localized emissions add up to give the total power and, unless the radiation field is transversely coherent, it is likely that at least one 'hot spot' is in any beam slice arbitrarily chosen and thus the total power has a significant pedestal value. These fluctuations in the power are described again by a Gamma distribution. The parameter of the Gamma distribution reflects the average number of 'hot spots' per slice. With increasing transverse coherence the contrast in Fig. 5.4 becomes stronger and individual spikes are clearly visible. Full transverse coherence is achieved at $z > 8$ m.

The slopes of the spikes in Fig. 5.4 corresponds to the longitudinal velocity of the spikes which is slower than the speed of light. This effect arises because the electron beam acts as a dispersive medium in the linear regime of the FEL amplification. The slippage of these spikes is identical to the group velocity in this dispersive medium. At saturation the group velocity of the spikes changes to the speed of light as seen in the bend in the slope at approximately $z = 13$ m. In this region new spikes grow and the pulse length increases.

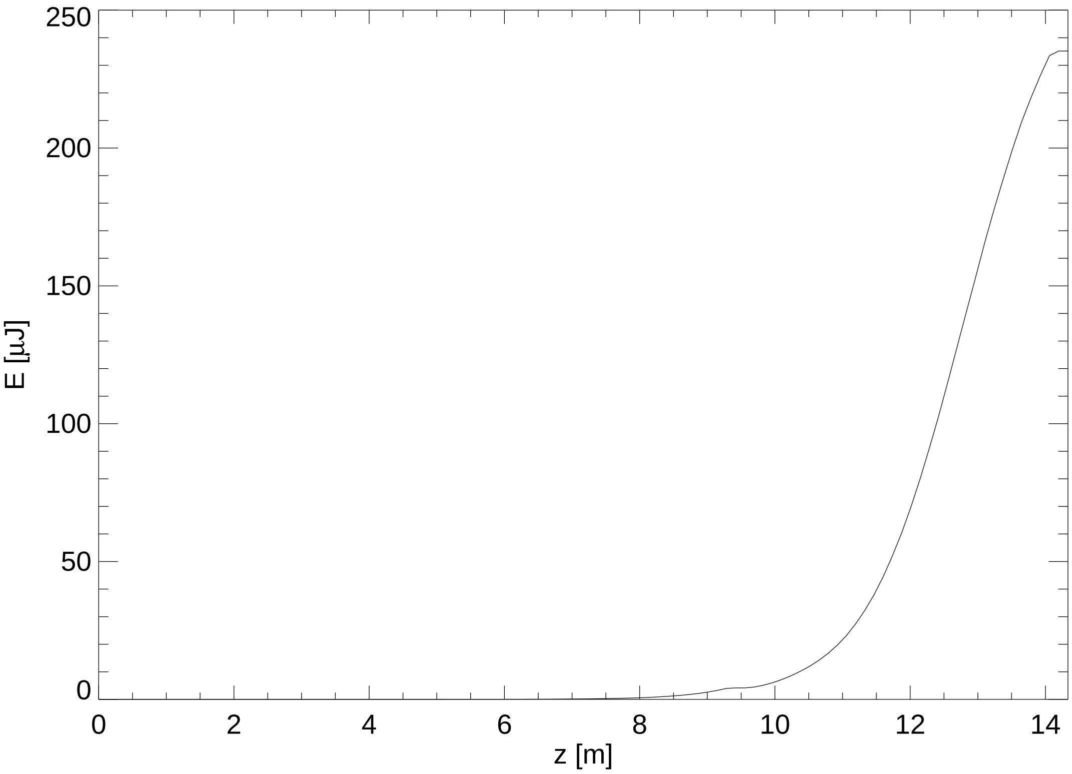


Figure 5.3: *Energy of the radiation pulse along the undulator axis for the TTF-FEL in phase I at 230 MeV.*

5.1 Transverse Motion of the Electron Beam Centroid

The theory of FELs to be applied here is based on some fundamental assumptions to keep the analytical complexity within limits. Even so, the dispersion relation, the solution of which determines the growth rate of the amplification, is difficult to solve. It consists of Bessel functions with complex arguments (Section 2.6). The impact of transverse motion can only be studied using numerical codes, which are capable to cover these aspects, such as GENESIS 1.3. In this section the transverse motion of the electron beam centroid is studied for the parameters of the TESLA Test Facility Free-Electron Laser at phase I.
It can be expected that coherent transverse motion of the electrons is not beneficial to the FEL amplification process. The radiation beam is almost stiff and cannot follow an arbitrary transverse motion of the electron beam, because the gain length is large compared to the characteristic length of transverse motion. The resulting reduction in the overlap of the radiation field and electron beam has a degrading influence on the growth rate of the FEL amplification [107, 108, 109, 110]. The most extreme case occurs if the electron beam and the radiation field are completely separated. While the existing radiation field diffracts the electron beam must build-up a new radiation field at its current position before the amplification can be started

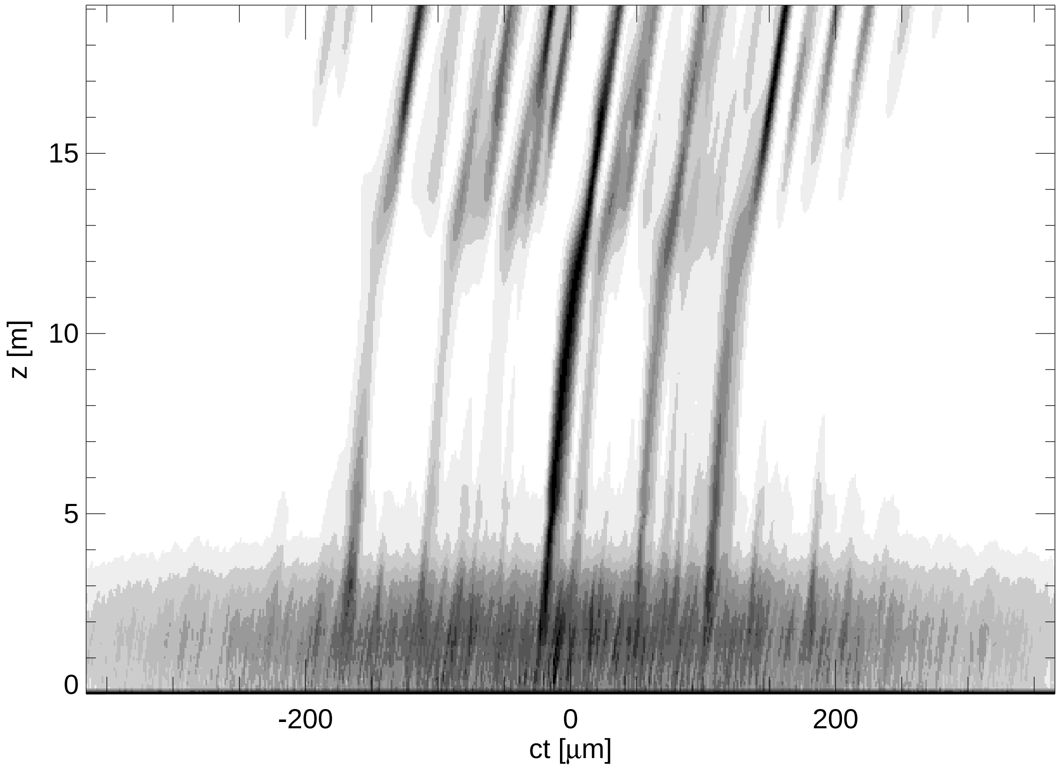


Figure 5.4: *Radiation power as a function of the position (ct, z) within the bunch and in the undulator, horizontal and vertical axis, respectively. For each z the radiation power profile has been normalized to the maximum value in ct.*

again. One consequence is an increase of the saturation length.
One of the major sources of transverse motion is a variation in the amplitude of the undulator field caused by fluctuations in the field strength of the individual undulator magnet poles, referred to as 'undulator field errors'. Because GENESIS 1.3 explicitly excludes the fast transverse oscillation of the electrons from the electron trajectory and only follows the betatron motion, the change in the trajectory caused by undulator field errors is approximated by a random transverse kick. To estimate the magnitude of the kick a pair of undulator poles is arbitrarily chosen. The root-mean-square field amplitudes are K_j and K_{j+1}. In order to provide a continuous and smooth trajectory between these poles the transverse momentum $p_\perp$ of the betatron motion changes for a planar undulator according to

$$\frac{\Delta p_\perp}{mc} = (-1)^j \sqrt{2} \frac{K_j - K_{j+1}}{\gamma} \quad . \tag{5.1}$$

The case of a helical undulator is rather complicated and not discussed here. The random kicks act coherently on the beam and the resulting trajectory is similar to a 'random walk' problem. GENESIS 1.3 generates a distribution of the undulator field errors such that all random kicks add up to zero. It can easily be shown that the sum of all kicks is proportional to the integration of the undulator field $\int_0^z B_y(z')dz'$. This is referred to as the 'first field integral' and is a measure of the transverse momentum of the electron beam along the undulator due to field errors. The beam-wander of the electron beam is given by the 'second field integral' $\int_0^z \int_0^{z'} B_y(z'')dz''dz'$. As mentioned above GENESIS 1.3 forces a zero value for the first field integral. The second integral is not optimized and fluctuates around a mean value of zero for independent samples of field errors. This can be excluded by suppling an additional input file for GENESIS 1.3, which contains the description of an undulator field with zero first and second field integrals.

The undulator of the TTF-FEL has a superimposed quadrupole lattice used to focus the electron beam and thus to reduce the gain length. If the electron beam has a transverse offset the quadrupole field deflects the beam like a dipole in addition to its focusing effect. The centroid position follows the trajectory of a betatron oscillation for a single electron. Therefore the main effect of undulator field errors is more the random excitation of betatron oscillations rather than the behavior of a random transverse walk. The situation is changed if steering magnets are used to align the orbit to the undulator axis in the beam-based alignment procedure. The excited betatron oscillation is disturbed by the kicks.

Other sources of transverse motion are off-axis injection and injection with an angle into the undulator, misalignment of the quadrupoles in the lattice and incorrect settings of the steering magnets, used for the beam-based alignment procedure.

The FEL process is not only affected by transverse motion because of the reduced overlap, but also because any random kick in the transverse direction changes the longitudinal velocity. As a result the longitudinal synchronization of the bunching phase and radiation phase is disturbed. The strength of the electron phase drift θ' depends quadratically on the transverse momentum as can easily be derived from Eq. 2.59. The accumulated phase shift is of less importance because it is compensated by a modified resonance condition. Because transverse motion slows down the electrons, the resonant wavelength is shifted towards longer wavelengths. More severe is the remaining fluctuation in θ', often referred to as 'phase shake'.

In Fig. 5.5 is shown the fluctuation in the transverse position and longitudinal phase drift for 1000 independent samples of undulator field errors. The individual field strengths of the undulator magnets have been taken randomly from a uniform distribution. A clear correlation between the fluctuations is visible which is caused by the excitation of betatron oscillations.

According to Eq. 5.1 large kicks are produced if the field strength of two adjacent undulator poles differs significantly. If the field strength K_j follows any typical distribution such as a uniform, parabolic or Gaussian distribution, the probability of a large kick is smaller than that for a small kick. Summing up over many undulator poles the small kicks, which occur more often, cancel each other. This is not the case for large kicks, which excite a betatron oscillation with rather large amplitude. The probability is low for an exact cancelation by an additional large kick, which has to occur immediately after the first kick. Otherwise the resulting off-axis position of the electron beam does not allow to cancel the betatron oscillation at all.

The transverse position and velocity in the betatron oscillation are deterministic, resulting in the quadratic relation between the two parameters seen in Fig. 5.5. Therefore phase shake and transverse overlap are much more descriptive parameters for use in a discussion of undulator field errors than the root-mean-square value of the magnetic field fluctuation itself [111]. The

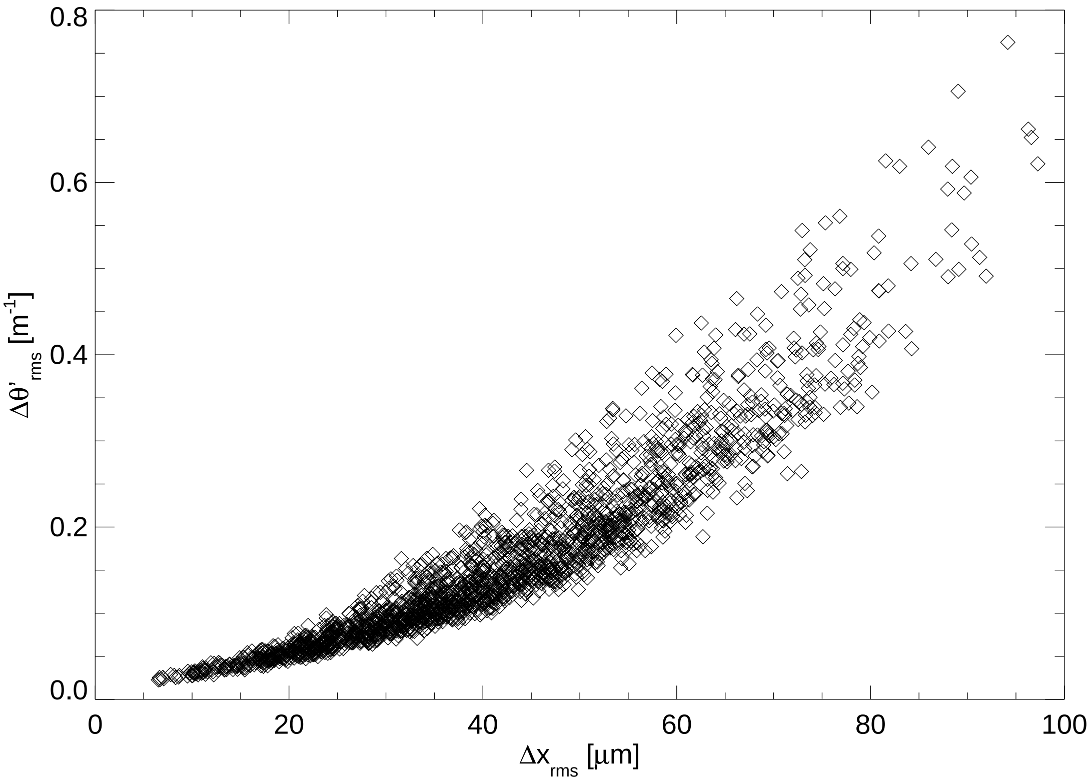


Figure 5.5: *Root-mean-square beam offset and phase shake for 1000 independent samples of undulator field errors.*

same root-mean-square value of the field fluctuations was used in all samples shown in Fig. 5.5. To distinguish between reduced transverse overlap and disturbed synchronization of electron and radiation beam, GENESIS 1.3 has been modified so that one of the two effects can disabled. The results of applying a noise term to the differential equation of the electron phase θ, which acts coherently on the electron beam, are shown in Fig. 5.6. For the simulation the radiation power has been taken at a fixed position within the undulator. In the case of the undisturbed FEL process this position is one gain length before saturation. This method allows one to estimate the change in the average gain length for each set of errors.

As expected the FEL output power drops with increasing amplitude of the phase shake. An improvement can be obtained if the radiation wavelength is optimized for the best performance. The resonance wavelength has been adjusted to the net phase change accumulated over the entire undulator length for each sample of errors.

For some small values of the root-mean-square phase shake a slight enhancement of the output power of up to 3% has been observed. The underlying reason is that for some settings the phase shake shifts the bunching phase relative to the radiation phase so that the emission of radiation is stronger than the absorption. Before the radiation phase adjusts itself to the new bunching

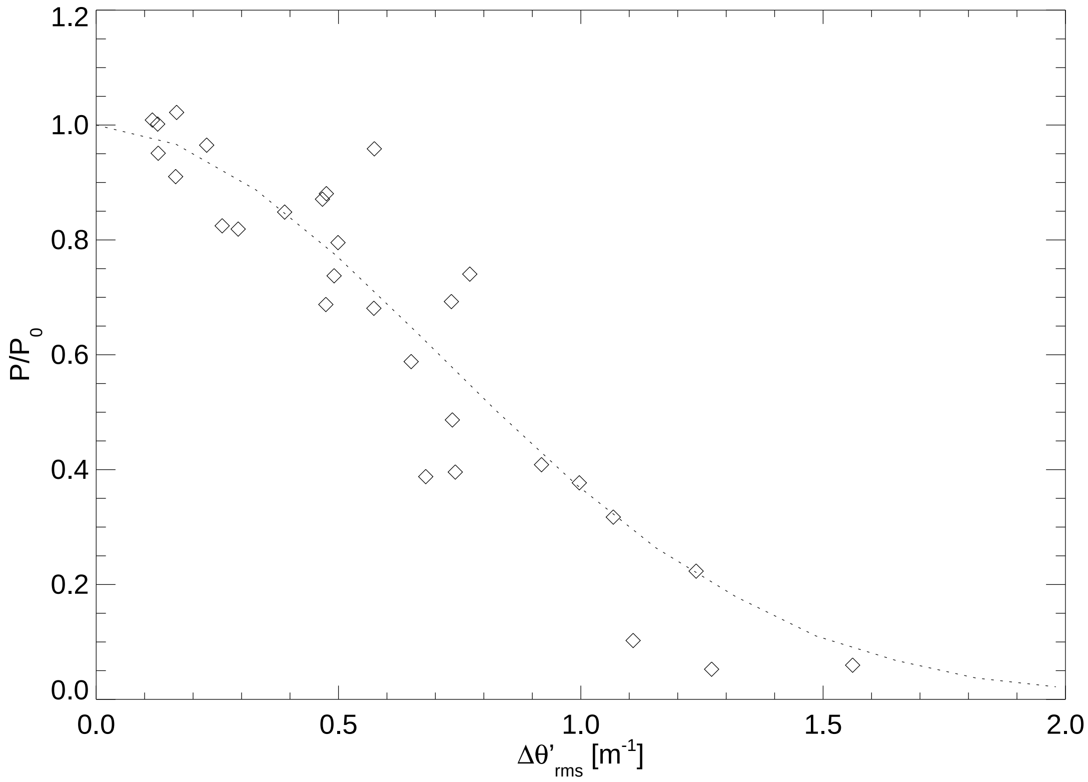


Figure 5.6: *FEL performance for different independent sets of electron phase drift errors, while the transverse motion is disabled. The radiation power has been taken at a fixed position within the undulator, in the case of the undisturbed FEL process (P_0) one gain length before saturation. The dotted line indicates the results for a fixed sample distribution of phase drift errors but scaled in magnitude.*

phase on a typical scale of one gain length the emitted power has become larger because the work of the electron beam against the radiation field has increased. To see a significant effect this phase shift has to occur when the separatrix is already reasonably large as it is the in case close to saturation.

The dashed line in Fig. 5.6 shows the result for a fixed set of phase shakes, being scaled with a varying factor to cover the desired range of the root-mean-square value of the phase shake. Its shape, as well as the distribution of the independent runs itself, is described by a Gaussian dependence on the phase shake.

Shown in Fig. 5.7 are the results for random motion in the transverse direction with the change in the longitudinal electron phase having been disabled. The distribution is similar to that of random phase drifts, Fig. 5.6, with the drop in the radiation power described by a Gaussian. The width of this distribution and therefore a measure for the tolerance on the transverse beam motion is $\sigma_{\Delta x} \approx 33$ μm compared to the root-mean-square size of the electron beam

$\sigma_x = 66\ \mu$m.
For a given set of undulator field errors (Fig. 5.5) the disturbance of the electron phase is negligible as compared to the beam-wander. The degradation of the FEL performance has its sole origin in the missing overlap between radiation and electron beam as long as the transverse beam offset is not larger than roughly 70 μm. The degradation of the FEL process depends as $\exp[-a\Delta x_{rms}^2 - b\Delta x_{rms}^4]$ on the transverse beam centroid motion. The Δx_{rms}^4 term arises from the phase shake, which is dominant for large values of Δx_{rms} [110].
Transverse motion is also generated if the beam is injected off-axis or under an angle. The results for variation in the injection angle and thus the amplitude of the excited betatron oscillation without any undulator field errors are also shown in Fig. 5.7. They are very similar to the dependence on field errors. This indicates that the root-mean-square offset Δx is a universal parameter such that the performance can be predicted by this parameter without knowing the underlying source for the transverse motion. For the TTF FEL at 230 MeV the tolerance lies at roughly 10 μm in order not to suffer from off-axis beam-wander. Similar results have been found for simulations at different beam energies [112, 113].
It is quite surprising that the degradation depends on the root-mean-square offset because it does not include any information about the time scale of the transverse motion. For the TTF-FEL it is implicitly given by the offset because the strong focusing quadrupole lattice allows one to derive the velocity from the amplitude of the excited betatron oscillation with the simple relation $v_0 \approx (c/\beta)x_0$, where v_0 is the amplitude in the oscillation of the transverse velocity, x_0 is the amplitude of the betatron oscillation and β the lattice-dependent beta function. For the TTF-FEL parameters at 230 MeV the beta function has a value of about 1 m.
The simulations exclude any correction of the beam orbit by steering magnets use, for example, in a beam-based alignment procedure. It can be expected that for the same root-mean-square value of the orbit misplacement the average transverse momentum is larger after correction than for an undisturbed betatron oscillation because the electron beam is deflected more strongly to the undulator axis. As a consequence the phase shake is enhanced and the distribution of Fig. 5.7 would become narrower.
A changing transverse position of the electron beam centroid must always be related to the FEL capability to follow the centroid position. A slow variation can be adapted by the FEL amplification for which the characteristic scale is the gain length. For the fixed quadrupole lattice of the TTF-FEL with a rather small phase advance per lattice cell the dependence on the beam energy of the optical lattice function, the period length of the betatron oscillation and the gain length is approximately linear, while the beam spot size is almost constant. The beam-wander over one gain length depends on the betatron amplitude but not on the beam energy because the transverse velocity is reduced but acts on a longer time scale which is proportional to the gain length. Results shown in Fig. 5.7 are valid for the entire energy range of the TTF-FEL Phase I, operating in the energy range between 230 and 390 MeV.
The possibilities for beam diagnostics within the undulator are limited due to the small undulator gap and the beam pipe design. In particular no information can be obtained about the radiation power along the undulator. The only measurement is at the end of the device. In principle the FEL amplification can be stopped by extracting the beam transversely from the undulator but the design of the vacuum chamber of most Free-Electron Lasers, such as the

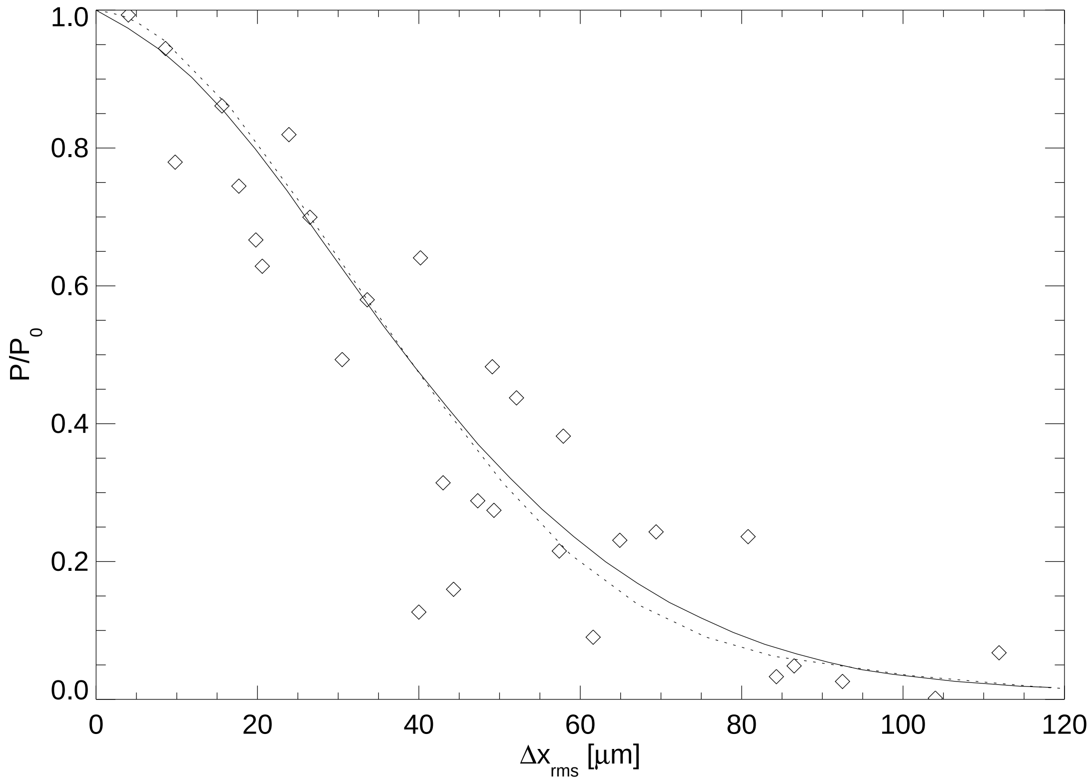


Figure 5.7: *FEL performance versus root-mean-square transverse beam offsets. The radiation power has been taken at a fixed position within the undulator, in the case of the undisturbed FEL process (P_0) one gain length before saturation. The dotted line indicates the results for the same set of random kick but scaled in magnitude while the solid represents the results for an undisturbed betatron oscillation with the same root-mean-square betatron amplitude.*

TTF-FEL, denies this. The beam has to be transported to the end of the undulator through the beam pipe. Over the transport distance the already modulated electron beam radiates at the resonant wavelength although further amplification is reduced.

A possible solution is to steer the electron beam in order to excite a large betatron oscillation. The missing overlap inhibits the FEL amplification by shifting the resonance wavelength. A monochromator filters out the radiation at the undisturbed resonance wavelength to be detected by the diagnostics.

In Fig. 5.8 the radiation power along the undulator is shown for different strengths of the kick up to an excited betatron amplitude of 1 mm. Beyond 1 mm the output power of the FEL grows again as discussed below. The steering magnets of the TTF-FEL are not strong enough to provide an optimum steering by a single magnet. The amplitude is limited to 0.5 mm [114] and two magnets have to be combined to reach the required betatron amplitude of 1 mm.

Using the optimum excitation of the betatron oscillation the FEL amplification can almost be

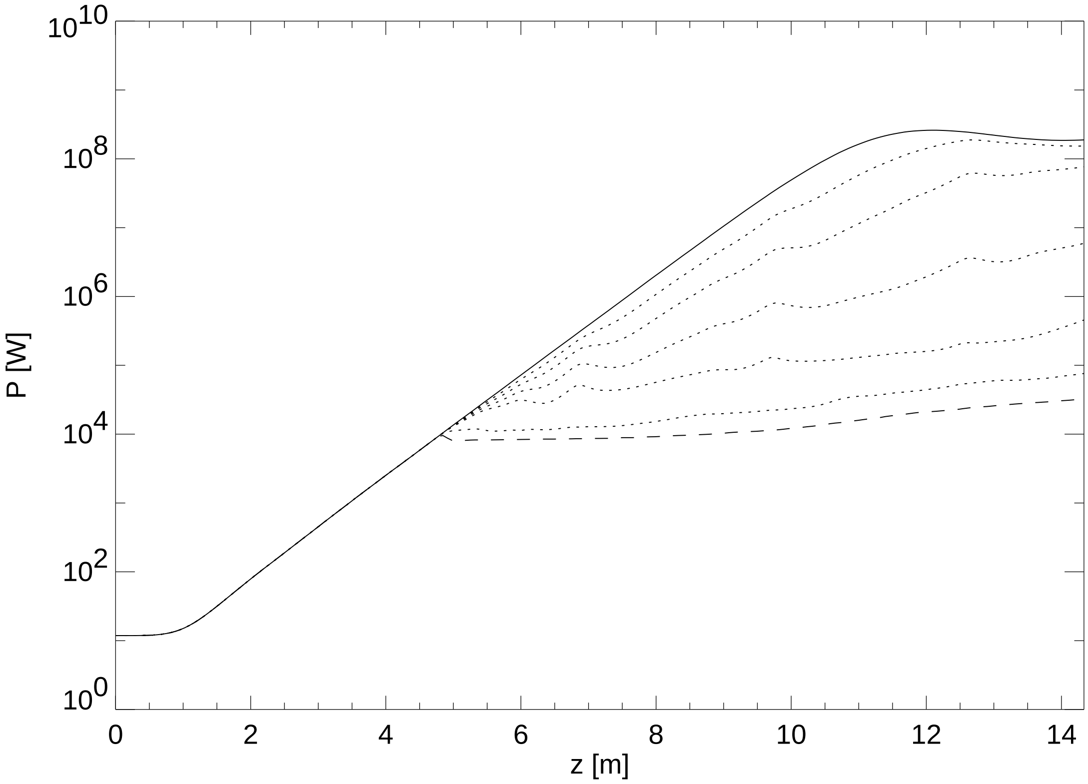


Figure 5.8: *Growth of the radiation power for different strengths of the transverse kick at $z = 5$ m. The solid line shows the undisturbed FEL amplification, the dashed line the best performance of inhibiting the amplification, while the dotted lines indicate the results for insufficient steering.*

turned off and the radiation power is conserved to the end of the undulator as presented in Fig. 5.9 for different positions of the steering magnet. For the case of steering the electron beam at an early stage of the amplification a further growth of the radiation power is unavoidable. The electron beam energy is slightly modulated and the resulting difference in the longitudinal velocity yields a further growth of the bunching factor and thus an enhanced emission of coherent radiation at the resonant wavelength.

The efficiency is not improved for betatron amplitudes exceeding 1 mm as shown in Fig. 5.10 for a steering position at $z = 5$ m. The smallest value of the radiation power occurs for a 1.2 mm offset. The power at the undulator exit is larger by a factor of 3.5 than the power at the steering position. At earlier positions the factor is even larger. Compared to the undisturbed growth of the amplification over the remaining distance to the end of the undulator of 2×10^4 this slight increase of the measured radiation power is acceptable.

The drop in the radiation power for betatron amplitudes $x_0 \approx \sqrt{2}\Delta x_{rms}$ smaller than 0.5 mm exhibits an $\exp[-bx_0^4]$ dependence. This agrees with the results presented above when the phase

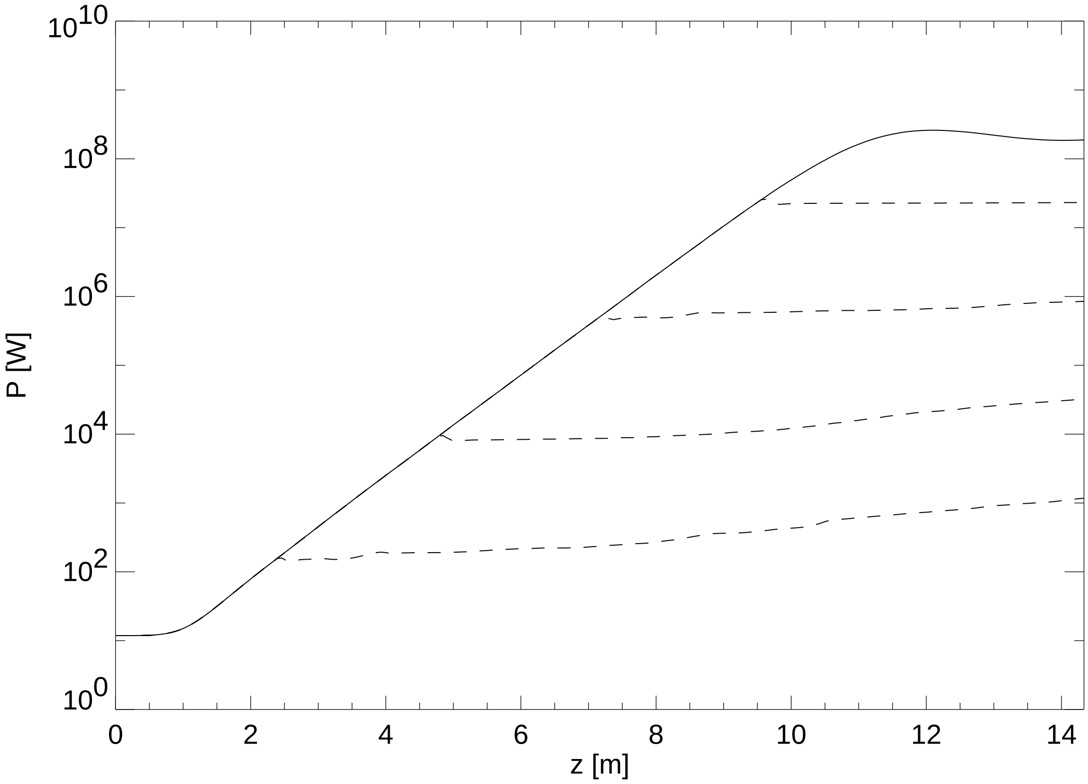


Figure 5.9: *Radiation power along the undulator for different positions of the steering magnet used. The evolution of the radiation power after the excitation of the betatron oscillation is indicated by the dashed line.*

shake becomes dominant. The value of b is larger in this case because the wavelength has not been adjusted for optimum amplification.
The power growth for large betatron amplitudes, limited by the beam pipe diameter, has its sources in the dependency of the excited betatron oscillation amplitude on the energy. Small variations in the beam energy change the path length of the trajectory. The strength of this correlation is often referred to as "momentum compaction" [28]. Therefore the micro bunching is not only enhanced by the longitudinal velocity modulation, as mentioned above, but also by time of flight differences. It is comparable to the functionality of a bunch compressor to compress the electron beam on the scale of the resonant wavelength. The compression increases with increasing dispersion along the excited trajectory. This is the case when the beam deflection is large due to the large beam offsets in the quadrupoles. Fig. 5.11 shows this enhanced emission as a result of compression for an offset up to 3 mm. Two spikes in the growth rate are visible which are clearly correlated with the turning point of the betatron oscillation. For larger amplitudes the growth is not as localized as in Fig. 5.11 but more distributed over the remaining length of the undulator.

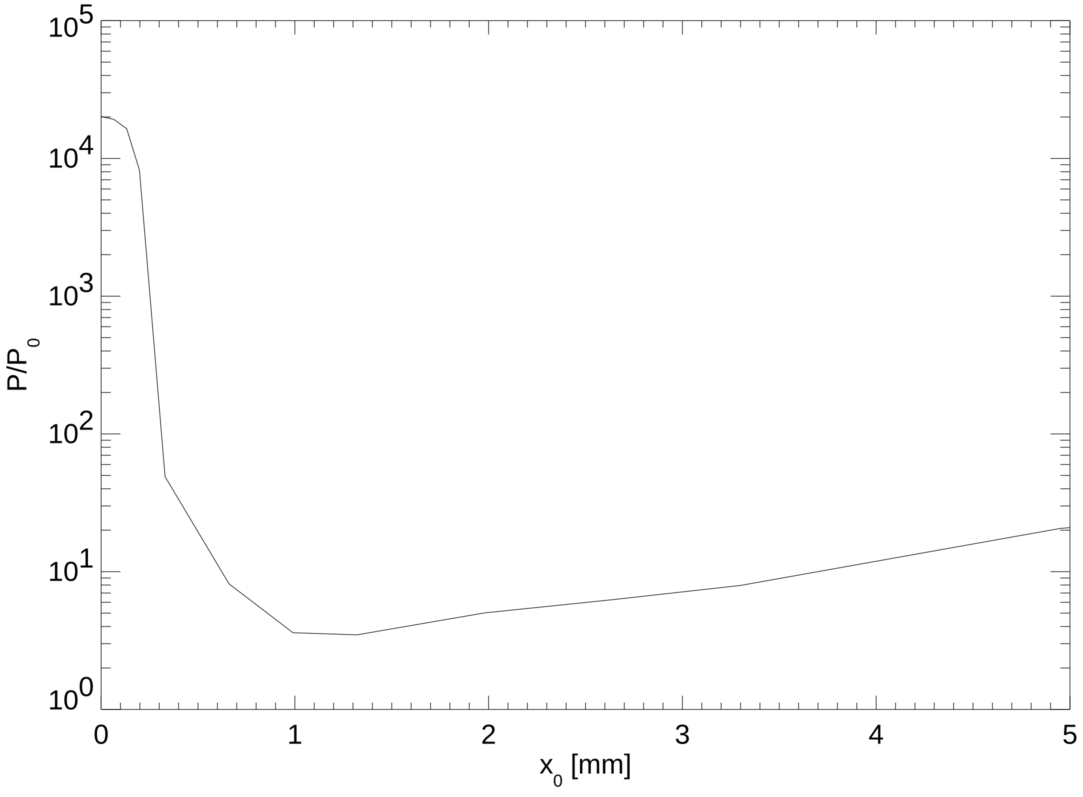


Figure 5.10: *Radiation power at the end of the undulator for different amplitudes of the excited betatron oscillation. The radiation power has been normalized to the power at the steering position $z = 5$ m.*

Even with this enhancement of the radiation power, simulations indicate that a kick of 1 mm is sufficient to nearly disable the FEL process. This allows one to probe the state of amplification at various positions along the undulator, which is an important task to compare measurements with theory and simulations.

5.2 Energy Modulation Induced by Wake Fields

In the design of Free-Electron Lasers for wavelengths in the VUV and X-ray region one has to deal with the increasing saturation length, which scales inversely to the FEL parameter [6]

$$\rho = \left[\frac{K f_c \Omega_p}{4 c \gamma k_U}\right]^{\frac{2}{3}} \quad , \tag{5.2}$$

where K is the dimensionless undulator parameter, f_c is the coupling factor, γ is the beam energy in units of the electron rest mass, k_U is the undulator wavenumber, Ω_p is the plasma

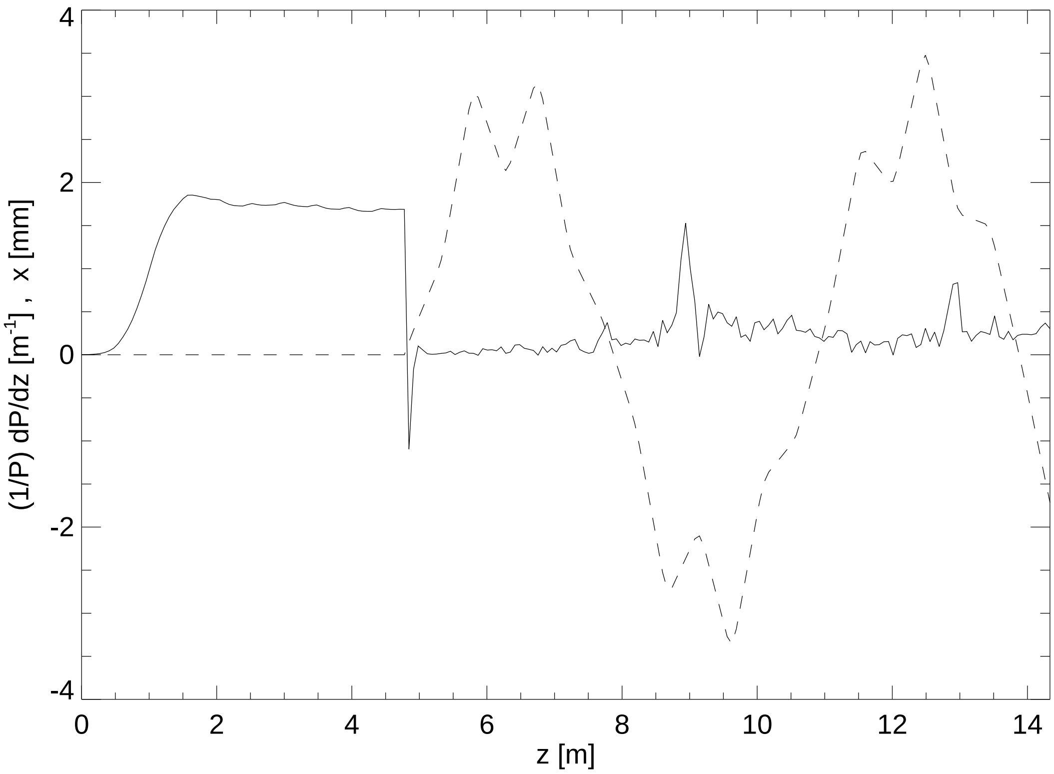


Figure 5.11: *Electron beam trajectory and exponential increment of the radiation field along the undulator (dashed and solid line, respectively).*

frequency and c is the speed of light. The growth of the saturation length is partially compensated by an increase in the undulator field or the peak current of the electron beam. These quantities are expressed by K and Ω_p in Eq. 5.2.

In the case of smaller undulator gaps or shorter bunches to increase the magnetic field or the peak current, respectively, the electromagnetic interaction between the electrons and the surrounding vacuum chamber is no longer negligible. These induced fields are referred to as wake fields [115] and are briefly described in this section. The discussion is focused on the interaction between the electrons and the wake fields within the undulator.

In free space a highly relativistic electron propagates rather uninfluenced by other electrons. The electric field is longitudinally suppressed by a factor γ^{-2} [43]. The electric force on a probe electron, placed with a transverse offset with respect to an electron, is largely compensated by the induced magnetic field of the relativistic electron. The net force on the probe electron, propagating with the same longitudinal velocity, scales as $1 - \beta^2$ and vanishes like the longitudinal force with γ^{-2}.

The situation changes significantly if boundary conditions are applied, in particular the limited aperture imposed by a beam pipe. The electric and magnetic fields are distorted so that the

fields might affect the motion of a trailing electron. In addition to the steady state solution [116, 117] for a constant beam pipe cross section any change in the geometric size of the beam pipe generates radiation [118, 119], which is either trapped in cavity-like structures, propagates along the beam pipe or is damped if the corresponding frequency is below the cut-off frequency of the beam pipe.
These different aspects can be classified into two different types of wake fields:

- wake fields determined by the shape and material of the vacuum chamber,
- wake fields induced by the surface of the beam pipe.

The latter includes effects arising from a rough surface or a dielectric layer on the surface of the pipe. Regarding the undulator of the TTF-FEL only three wake field sources are of importance. The model to estimate their magnitudes are described in the following subsections.
To obtain the total wake field provided by an electron beam the single electron wake field is convoluted with the longitudinal bunch profile for each source of wake. The different wakes are added. A round beam pipe with radius R is assumed as well as that the electron is located at the undulator axis.

5.2.1 Resistive Wall

Electric fields penetrate into the metal of a beam pipe and induce charge at the surface. Propagating along with the electric field of the electron, the induced charge loses energy due to the resistance of the beam pipe material. The energy loss is determined by the electric field with an non-vanishing longitudinal field component at the location of the electron.
From the analysis of Maxwell's equations in the frequency domain it can easily be derived that the longitudinal electric field is constant over the entire transverse plane of the beam pipe ($r < R$) and drops as $\exp[i\lambda(r-R)]$ beyond R. The imaginary part of the complex parameter λ defines the characteristic length δ_S of the electric field penetration, called the 'skin depth', with $\delta_S = 1/\Im\mathrm{m}\lambda = c/\sqrt{2\pi\sigma|\omega|}$, where σ is the conductivity of the material and ω is the frequency under consideration. Due to the continuity of the tangential magnetic field at $r = R$ the amplitude of the Fourier component of the electric field is fully determined. The longitudinal electric field and thus the energy loss per unit length of a following electron $W_z = e(dE_z/dz)$ in the time domain can be calculated by a rather lengthly inverse Fourier transformation [116]. The result for the so-called wake function W_z of the resistive wall wake field within the beam pipe ($r < R$) is

$$W_z(t) = -\frac{4ce^2 Z_0}{\pi R^2}\left(\frac{1}{3}e^{\frac{t}{\tau}}\cos[\sqrt{3}\frac{t}{\tau}] - \frac{\sqrt{2}}{\pi}\int_0^\infty dx \frac{x^2 e^{\frac{t}{\tau}x^2}}{x^6+8}\right) \quad , \tag{5.3}$$

where t is the longitudinal position of the second electron in the moving frame of the first electron,

$$\tau = \frac{1}{c}\left[\frac{2R^2}{Z_0\sigma}\right]^{\frac{1}{3}} \quad , \tag{5.4}$$

and $Z_0 = 377\ \Omega$ is the vacuum impedance. The wake function is evaluated only for positions behind the electron ($t < 0$) in order to not violate causality. In front of the generating and highly relativistic electron the wake field drops rapidly and is $W_z(t) = 0$ for $t > 0$ in good approximation. Another approximation is that the conductivity σ is independent of the frequency, which is well justified for an aluminum beam pipe and a bunch length of 50 μm as in the TTF-FEL.
To be consistent with the coordinate system used in FEL theory the longitudinal position t in the moving frame of the electron beam has the dimension of time to distinguish between the position z in the undulator. The characteristic scale of the resistive wake fields is τ. For the TESLA Test Facility undulator with $R = 4.75$ mm and $\sigma = 3.65 \times 10^7\ \Omega^{-1}\mathrm{m}^{-1}$ for the aluminum vacuum chamber the corresponding length is $c\tau = 15\ \mu$m and thus comparable to the TTF-FEL bunch length.
The limit $t \to 0$ yields the field strength seen by the electron of its own wake field. This is identical with the Ohmic losses of the induced charge in the metal beam pipe. The energy loss of the electron is $ce^2Z_0/2\pi R^2 = -W_z(0)/2$. The factor of one half arises because the electron sees in average only one half of its charge. To understand this the electron charge is artificially constructed by adding up small charges Δq with $\Delta q \ll e$. Each charge Δq sees a wake field which scales linearly with the accumulated charge q. Adding up all infinite small charges yields an energy loss of the electron proportional to $\int_0^e q dq = e^2/2$. This factor is typical for the general problem of the interaction of a charged particle with its own field. It is referred to as the fundamental theorem of beam loading [120].
In the limit of a perfectly conducting beam pipe ($\sigma \to \infty$) the characteristic size τ of the wake potential and thus the convolution with any arbitrary charge distribution vanishes due to the identity

$$\int_{-\infty}^{0} W_z(t)dt = 0 \quad . \tag{5.5}$$

In other words a single electron does not experience an energy loss in this limit [115] as expected due to the fact that the resistance is zero.

5.2.2 Geometric Effects

This classification refers to wake fields caused by any change in the geometry of the beam pipe, such as rf-cavities, bellows or vacuum pump ports. The size of such cavity-like perturbations is typically larger than the bunch length and therefore the fundamental resonant frequency of the cavity is much lower than the frequencies of electric field in the pulse, generated by the passage of an electron bunch. Under the assumption of a metal beam pipe with a high conductivity the electric field lines are nearly perpendicular to the surface before the electron enters the cavity. In this case the estimation of the wake fields is based on a model where the electric field is regarded as a plane wave with the frequency ω, diffracting at a sharp edge at the entrance of a pill-box cavity. The width of the diffracted field at the exit of the cavity is $\sqrt{cg/\pi\omega}$, where g is the cavity length. The field extends into the space of the cavity as well as towards the electron.

When the electron exits this structure, a fractional part of the field is trapped in the cavity, while the remaining field propagates along the beam pipe together with the electron beam. Unlike the steady state resistive wall wake fields these wake fields are not constant in time. The diffraction of the electric field has to catch up with the beam. Therefore the explicit electric field depends on z as well as on t. The calculation is significantly simplified if the wake fields are averaged over a certain distance enclosing the geometric change of the aperture. This implies that the wake fields are given by the electric potential rather than by the electric field. Therefore the wake function W_z, which is proportional to the average longitudinal electric field, is sometimes misleadingly referred to as the wake potential.
The calculation of the geometrical wake field [119] yields

$$W_z(t) = -\frac{ce^2 Z_0}{\pi^2 RL}\sqrt{\frac{g}{2}}\frac{1}{\sqrt{-t}} \quad , \tag{5.6}$$

where the wake function is only non-zero for longitudinal positions behind the electron ($t < 0$). For the TTF-FEL design with beam pipe radius R the changes in the aperture repeat with each undulator module. Therefore the wake field is averaged over one single undulator module with $L = 4.5$ m. Within this distance all sources are added by defining the effective gap $g_{eff} = (\sum_i \sqrt{g_i})^2$, where g_i is the length of the individual gaps.
The limitation of this diffraction model can be seen in Eq. 5.6, where the wake function has a singularity at $t \to 0$. The physical meaning of this is that a single electron loses an infinite amount of energy, because all frequencies are taken into account for the derivation of the wake potential. This problem is avoided when the wake function is convoluted with the longitudinal bunch distribution. The finite length cuts the excited frequency at the characteristic frequency $2\pi/\sigma_t$ where σ_t is the electron bunch length. Otherwise it has to be taken into account that the electric field has a transverse opening angle of $1/\gamma$. For a single electron the effective pulse length in the plane wave model would be finite with a length of R/γ.
The only aspect of quasi-periodic changes in the aperture which may be relevant to the TTF-FEL is the surface roughness of the vacuum chamber. This will be treated in the next section.

5.2.3 Surface Roughness

The beam pipe is a waveguide, in which a radiation field can propagate as a waveguide mode, if the frequency of the field is above the cut-off frequency. The phase velocity is, under normal conditions, faster than the speed of light and the wave cannot couple with an electron beam, propagating along with the radiation field. The situation is changed if the surface is not smooth. The perpendicular electric field lines at the beam pipe surface follow any curvature, slowing down the speed of propagation. If the phase velocity is below the speed of light then the electrons may couple to this waveguide mode. The generated wake fields are trail the electron bunch.
To model the surface roughness a thin layer of dielectric material is assumed, which has the same effect of slowing down the phase velocity. The thickness δ of this layer is set to the root-mean-square size of the roughness. For a smooth surface the phase velocity of the nth waveguide mode is given by

$$v_p = \frac{c}{\sqrt{1-\left(\frac{\omega_n}{\omega}\right)^2}} \quad , \tag{5.7}$$

where ω_n is the cut-off frequency of the nth mode. The fundamental mode has the lowest cut-off frequency and thus the slowest phase velocity of all modes. It is therefore most affected to any surface roughness or to a dielectric layer.
Solving the Maxwell's equations for a thin dielectric layer the resonant wave number of the fundamental waveguide mode is

$$k_0 = \sqrt{\frac{2\epsilon}{R\delta(\epsilon-1)}} \quad . \tag{5.8}$$

To match this model with the surface roughness the dielectric constant ϵ has to be set to $\epsilon \simeq 2$ [121].
The wavelength of this waveguide mode is 120 μm for the TTF-FEL beam pipe with $R = 4.75$ mm and an estimated surface roughness of $\delta \approx 300$ nm. For the electron beam energies of the TTF-FEL the wavelength is much larger than the characteristic width $R/\gamma < 10$ μm of the Lorentz-contracted electric field of an electron at the surface of the beam pipe. The waveguide mode can therefore be excited.
The wake function is

$$W_z(t) = -\frac{ce^2 Z_0}{\pi R^2}\cos(k_0 t) \quad . \tag{5.9}$$

Again the expression is only valid for a position behind the electron and vanishes otherwise. This wake field model is rather conservative and predicts larger wake field amplitudes compared to a different model [122] in particular if the typical longitudinal size of the surface variation in the beam pipe is much larger than the root-mean-square amplitude δ.

5.2.4 Transverse and Higher Order Wake Fields

So far the discussion has been restricted to the longitudinal wake field components for an electron beam propagating on-axis. For a complete description the transverse field as well as the case for an off-axis position of the beam have to be included. Due to the properties of the Maxwel's equations all these wake fields are at least linear in the transverse position and vanish at the axis of the vacuum chamber. In the ideal case that the axis of the undulator field and the vacuum chamber coincide the effect on the FEL performance by these wake fields is negligible. Otherwise the transverse components of the wake fields could deflect the beam. The reduced overlap between radiation field and electron beam could result in a strong degradation of the FEL performance (Section 5.1). This class of effects is beyond the scope of this work. The following discussion is restricted to the lowest order wake fields (Eqs. 5.3, 5.6 and 5.9).

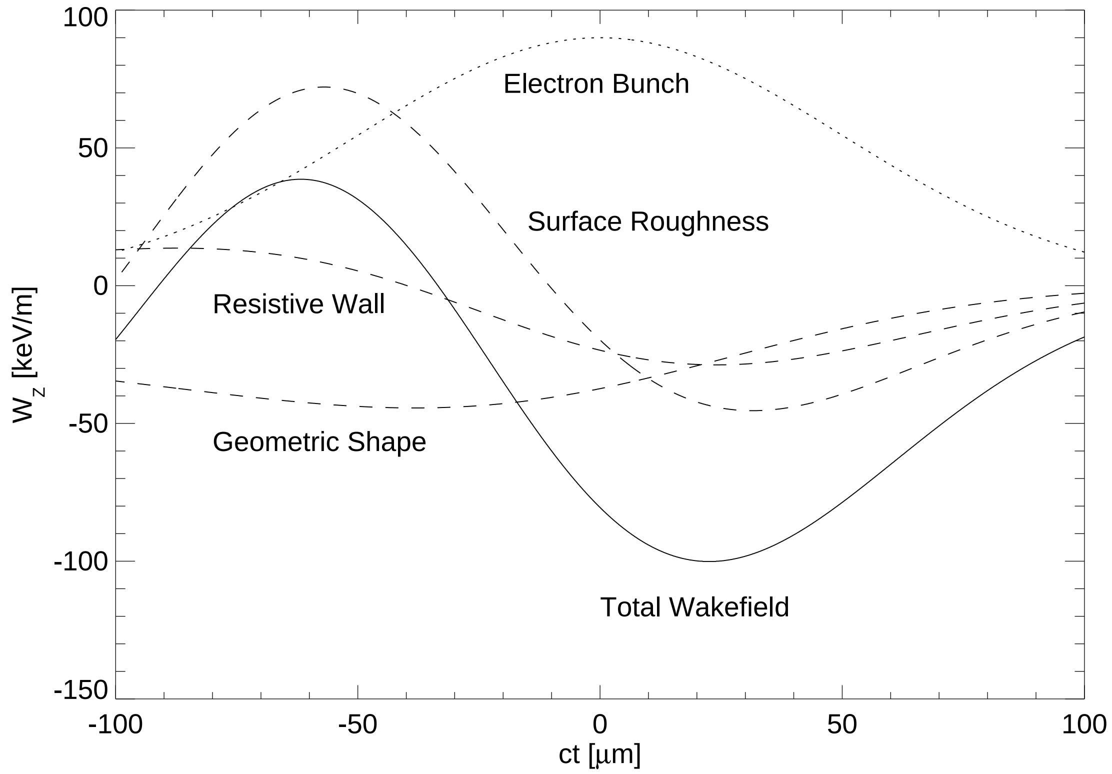


Figure 5.12: *Wake potential for the TESLA Test Facility Free-Electron Laser with a bunch length of* 50 μm*, a beam charge of 1 nC, a beam pipe diameter of 9.5 mm, a root-mean square surface roughness of 350 nm and slits for pumping and beam position monitors in the diagnostic blocks between the undulator modules. The electron profile is drawn for reference. Positive values of* t *correspond to the head of the bunch.*

5.2.5 Energy Modulation by Wake Fields During FEL Amplification

Based on the wake field model described above, the total wake potential as the sum of the wake functions of the resistive wall, the geometric change in the beam pipe and the surface roughness is convoluted with the electron beam profile. It is common to refer to the result as the wake potential [115]. In contrast to the wake function of a single electron it does depend on the electron beam properties. The wake potential is imported into GENESIS 1.3 to simulate a change in the electron beam energy depending on the longitudinal position in the bunch.
The wake potential for the undulator of the TESLA Test Facility is shown in Fig. 5.12. For the calculation the parameter set for phase II of the TTF Free-Electron Laser has been used, namely a bunch length of 50 μm, a beam charge of 1 nC and an energy of 1 GeV. The diameter of the aluminum beam pipe is 9.5 mm. Slits for pumping and beam position monitors in the diagnostic blocks between the undulator modules generate the geometric wake fields and the root-mean-square surface roughness is assumed to be 350 nm.

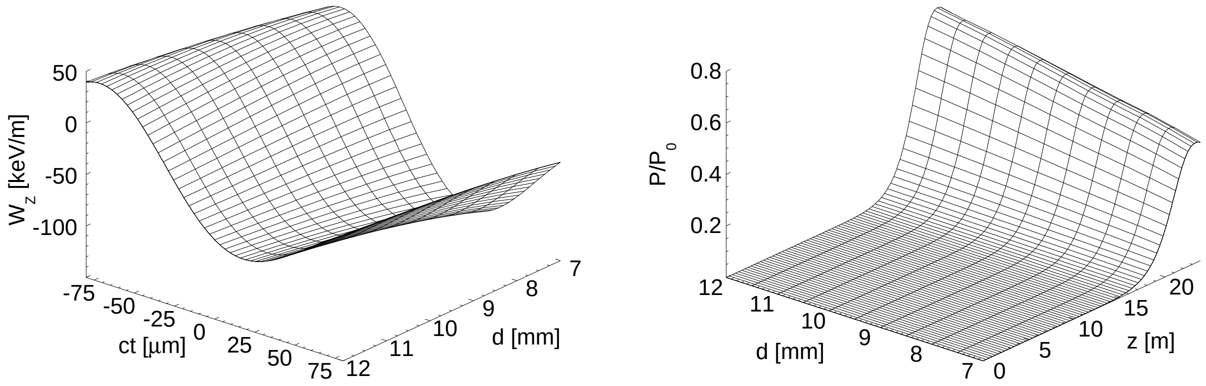


Figure 5.13: *The wake potential and radiation power along the undulator as a function of the beam pipe diameter d for $\delta = 350$ nm and $c\sigma_t = 50$ μm. The radiation power has been normalized to the saturation power P_0 in the case excluding wake fields.*

Most of the beam is decelerated except for the tail where due to the surface roughness wake field the electrons are accelerated. The maximum energy loss occurs 20 μm before the center of the electron beam with a gradient up to 100 keV/m. For the 30 m long undulator the total loss is comparable to the energy bandwidth of the FEL and it can be expected that the part of the bunch with the maximum energy loss does not contribute to the FEL amplification any longer. About 40 μm behind the center the individual wake fields compensate each other and the amplification is not inhibited by energy loss.

Simulations have been performed to study the influence of the beam pipe diameter, surface roughness and bunch length on the TTF-FEL phase II. Simulations for a longer bunch length, as in the case of the TTF-FEL phase I with a length of 250 μm, show that the impact of wake fields is almost negligible.

In the design of undulators for high gain Free-Electron Lasers, the magnetic gap between the undulator poles is a critical parameter. To achieve the highest possible peak field the gap has to be minimized. The limitation is given by practical considerations for the vacuum chamber, which has to fit between the pole faces. In the case of the TTF-FEL the gap is chosen to be 12 mm. With a minimum thickness of the beam pipe wall of 1.25 mm the remaining diameter is 9.5 mm. The induced wake fields for various diameters are shown in Fig. 5.13. The amplitude of the wake potential is rather insensitive to the size of the vacuum chamber d and decreases by only about 50% over the range in d from 7 mm to 12 mm. The shape of the wake potential is unchanged but is shifted towards the head of the beam. The wake fields are then more inductive than resistive and the net energy loss of the beam is reduced by a stronger energy transfer from the electron bunch head to the tail.

The evolution of the amplified radiation field of the Free-Electron Laser is also plotted in Fig. 5.13 for different beam pipe diameters. The only visible effect is in the saturation power which is reduced by roughly 30% from the largest to the smallest diameter. The saturation length remains the same. Compared to the results for undisturbed amplification, in which wake fields were not applied, the degradation of the FEL performance is up to 50%, which is tolerable since the overall gain is larger than 10^7.

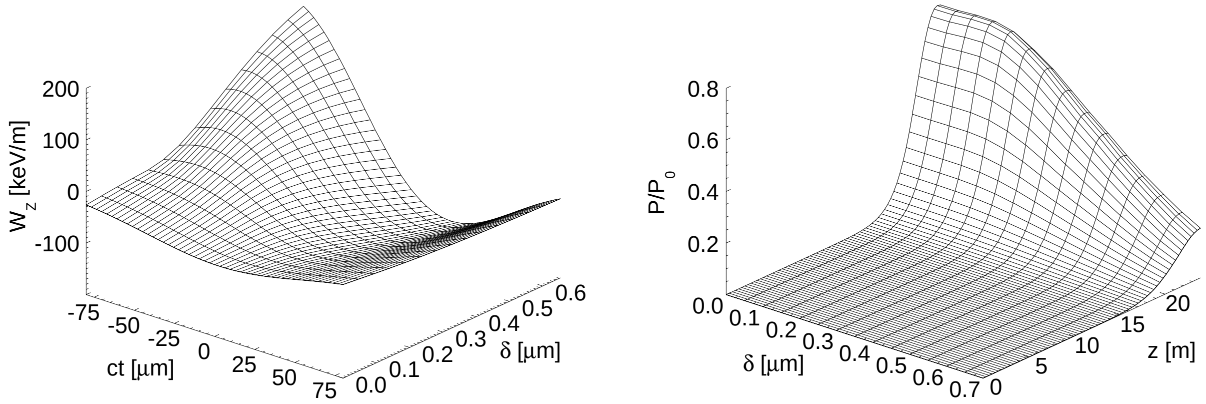


Figure 5.14: *The wake potential and radiation power along the undulator as a function of the surface roughness δ for $d = 9.5$ mm and $c\sigma_t = 50$ μm. The parameter δ defines the root-mean-square amplitude of the roughness. The radiation power has been normalized to the saturation power P_0 in the case excluding wake fields.*

While the gap and therefore the beam pipe radius is specified in the design stage of the undulator, the surface roughness is a problem which arises during the manufacture of the vacuum chambers. The vacuum chambers are made by extrusion which produces structured longitudinal and azimuthal scratches rather than a pure random pattern of surface modulations. Simulations have shown that the explicit structure of the scratches is not important as long as the typical size is much smaller than the bunch length. Three dimensional simulations of repetitive patterns in the surface roughness indicate a reduced wake potential amplitude relative to that predicted with the two dimensional model described above. A correction factor of 0.5 is added to Eq. 5.9 to cover these aspects [123].
The wake potential becomes strongly modulated by the surface roughness wake fields in the range of δ from 0 nm to 700 nm. At $\delta = 700$ nm the surface roughness wake fields dominates with an amplitude of approximately 160 keV/m. The wake potentials used for the simulations are shown in Fig. 5.14. The degradation of the FEL gain over the chosen range is larger than that for the various beam pipe diameters. The minimum reduction of the radiation power is only 25% in the case where only the wake fields of the resistant wall and change in the geometric aperture are incorporated, excluding the surface roughness. The impact of the surface roughness wake is negligible up to small values of δ of 200 nm. For larger variations in the pipe wall surface the saturation power drops to 20% of that of the undisturbed FEL performance. The saturation length remains constant except for very large values of δ, where it increases slightly as seen in Fig. 5.14. On the basis of these simulations the constraint for polishing the beam pipe for the TTF-FEL undulator is a surface roughness below the level of 200 nm. Any further improvement

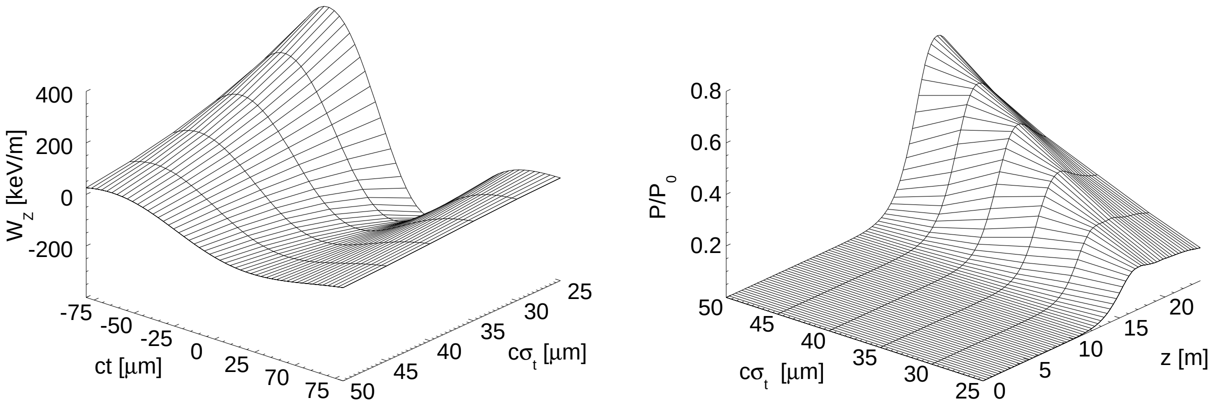


Figure 5.15: *The wake potential and radiation power along the undulator as a function of the bunch length for $d = 9.5$ mm and $\delta = 350$ nm. The total charge is kept constant at 1 nC. The radiation power has been normalized to the saturation power P_0 in the case excluding wake fields.*

in surface quality is not rewarded by an improved FEL performance.

In contrast to the beam pipe diameter and surface roughness the bunch length is not a fixed parameter. Within a limited range the bunch length can be varied and optimized for the highest efficiency of the FEL. A general tendency is that for higher peak currents the output power of the Free-Electron Laser is enhanced, while the saturation length decreases. The bunch length and the beam charge, which together define the peak current, are coupled with other beam parameters. In addition to the emittance and energy spread, which are both enlarged by space charge forces or by the compression scheme if the charge is increased or the bunch length further compressed, the wake field amplitude is also such a coupled parameter. The net change in the FEL efficiency for higher peak currents is not necessarily positive when all effects are involved. If the bunch length is reduced then the single electron wake functions (Eqs. 5.3, 5.6 and 5.9) adds more coherently in the convolution with the bunch profile and the amplitude of the wake potential becomes larger. The wake potential for bunch lengths down to 25 μm is presented in Fig. 5.15. As noted previously, the wake field for phase I of the TTF-FEL with a bunch length of 250 μm is negligible and becomes first noticeable at 50 μm, the design value for phase II at a beam energy of 1 GeV. Although a bunch length shorter than 50 μm is not intended for the FEL, it is of importance to study the dependency on this parameter. It is not intuitively clear that a short bunch implies a worse amplification because the degradation by stronger wake fields is partially compensated by a higher efficiency and a shorter gain length. The energy loss is not accumulated so much if the interaction length for the FEL is shorter.

For the simulations the bunch length has been varied from 50 μm to 25 μm in steps of 5 μm. To simplify the interpretation of the results, which are shown in Fig. 5.15, the energy spread and emittance have been kept constant. The saturation length decreases from 23 m to 17 m and the saturation power drops to a level of 18% compared to the case excluding all wake fields

as the bunch length is reduced to 25 μm..
Besides the change in the saturation power and saturation length, the evolution of the radiation power beyond saturation differs for short bunch lengths below 30 μm. The power remains constant in contrast to the typical behavior of power oscillations with an amplitude of typically 50%. The damping of the oscillation for larger wake fields can be explained by the fact that more and more parts of the electron bunch get out of the bandwidth of the FEL amplification due to the accumulated energy losses. Finally the amplification comes to a complete halt.
The maximum energy loss so that saturation is still reached before the average electron energy slips out of the amplification bandwidth the acceptable energy loss can approximately be estimated by the constraint

$$\frac{d\gamma}{dz} \ll \frac{1}{4} k_U \rho^2 \gamma_0 \quad , \tag{5.10}$$

where γ_0 is the initial beam energy and k_U is the undulator wavenumber. For the parameters of the TTF-FEL phase II the limit lies at 250 keV/m. This constraint is exceeded for bunch lengths smaller than 35 μm in agreement with the simulation in which the radiation power does not oscillate in the deep saturation regime.
A similar problem arises from the fact that the radiation, which propagates faster than the electron beam, slips into a region where the wake fields have altered the beam energy and thus the resonance condition. The tolerance for the beam energy modulation is similar to Eq. 5.10 with

$$\frac{d\gamma}{cdt} \ll \frac{1}{4} k \rho^2 \gamma_0 \quad . \tag{5.11}$$

Due to the dependence of the radiation wave number k on the energy as γ^2 this constraint is relaxed for a high beam energy as is the case for the TTF-FEL. Eq. 5.11 gives an upper limit of 1 MeV/μm for the energy modulation along the electron bunch. Assuming the worst case ($c\sigma_t = 25$ μm) this gradient is accumulated after 600 m. It is most unlikely that this constraint can be violated by the initial energy modulation, caused during the beam transport from the electron beam source to the undulator.
The FEL performance including wake fields can be estimated by a simple approximation. Significant wake fields are only expected for short bunches as required in current and future projects VUV and X-ray Free-Electron Lasers. The total slippage of the radiation field in this wavelength region is small compared to the size of the electron beam and the longitudinal variation of the wake fields. Each part of the electron beam can be regarded independently, where the beam current and wake potential amplitude is kept constant over the slippage length. The gain is a universal function, which depends only on the local current and the energy loss. In Fig. 5.16 is shown the radiation power at $z = 20$ m for different settings of these parameters. Knowing this dependency, either by an interpolated table or by fitting a function with two arguments to the data, the effect of any wake field can quickly be estimated by integrating the power as a function of I and W_z along the electron bunch.
All the simulations, presented so far, are based on an FEL amplifier. This is a valid approach because the difference to the radiation pulse of a SASE FEL, averaged over many independent

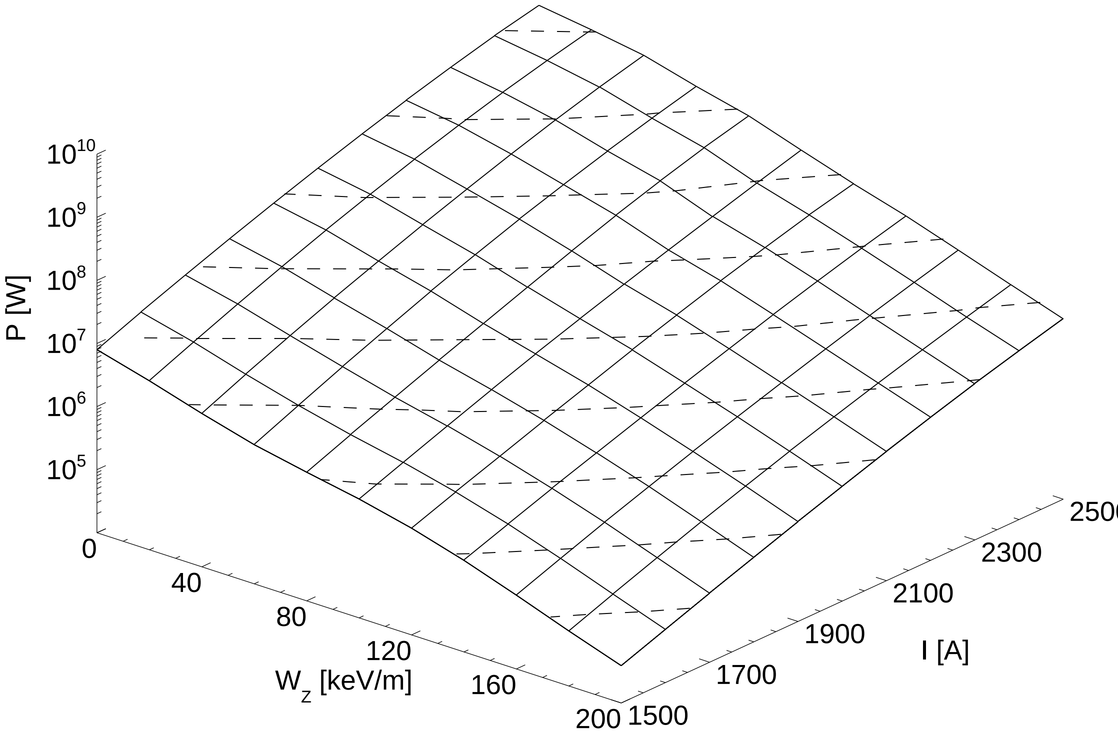


Figure 5.16: *Radiation power in the steady state regime at 20 m for different peak currents and energy losses. Contour levels are drawn by dashed lines.*

shots, is small and can be neglected. For the simulation of a single shot the radiation pulse is dominated by its noisy pulse shape although the amplitude of those spikes are reduced where the electrons experience a large energy loss. The energy of the radiation pulse for a single shot is plotted in Fig. 5.17, together with the result for the same initial conditions but without wake fields.

The reduction of the power is 35% and comparable to the results for the FEL amplifier, presented above. This may to be surprising since a SASE FEL has the general capability to be always on resonance to the wavelength with the maximum growth rate and thus to follow the changing resonance condition. This argument is only partially true because the change in the resonant wavelength implies that the spacing of the micro bunching has also changed, which is not the case. The beginning and the end of a longitudinally coherent part of the pulse – normally a spike – are not in phase anymore and the coherence length would be reduced by the continued FEL interaction. Only the lower frequencies in the FEL amplification bandwidth would benefit from an energy loss of the beam. A discussion of the SASE FEL amplification bandwidth is almost equivalent to that of the energy bandwidth of the FEL amplifier and would yield the same constraints Eqs. 5.10 and 5.11. In Fig. 5.18 a typical radiation pulse and spectrum is

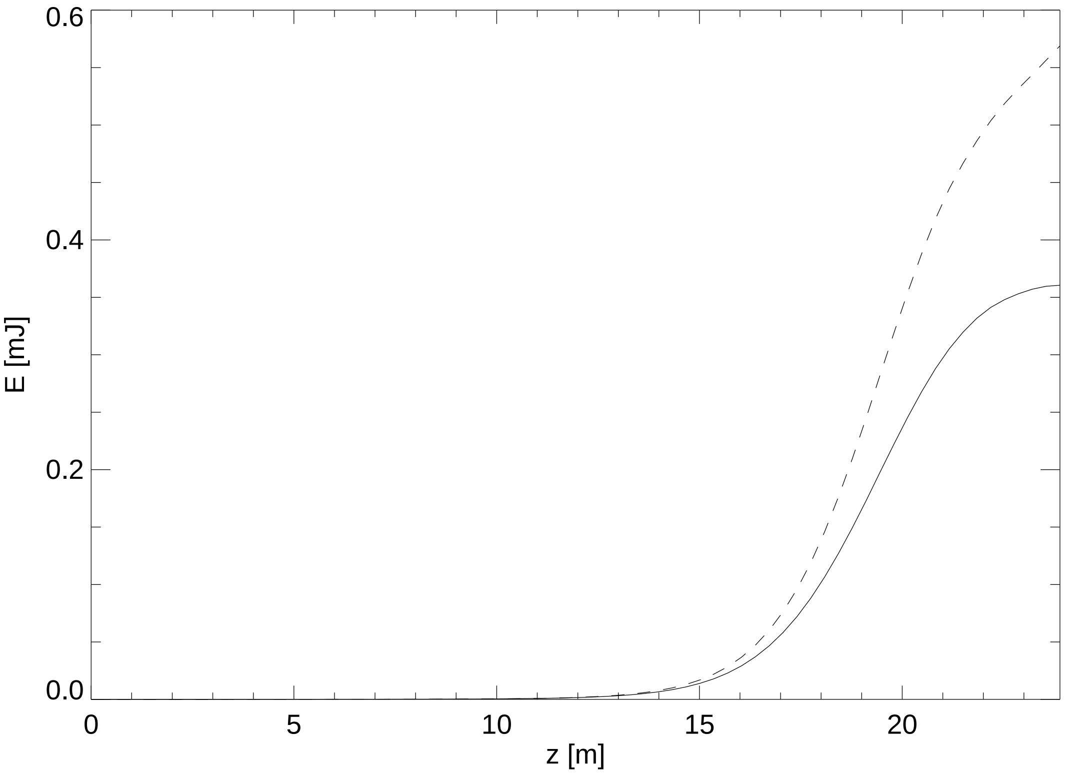


Figure 5.17: *Radiation energy for a SASE FEL simulation. The dashed line shows the results for the same initial settings but excluding wake fields in the simulation.*

shown. The wake fields shift the spectra about 0.7% towards lower frequencies, while most of the spikes in the time and frequency domain coincide with the simulation results excluding wake fields.

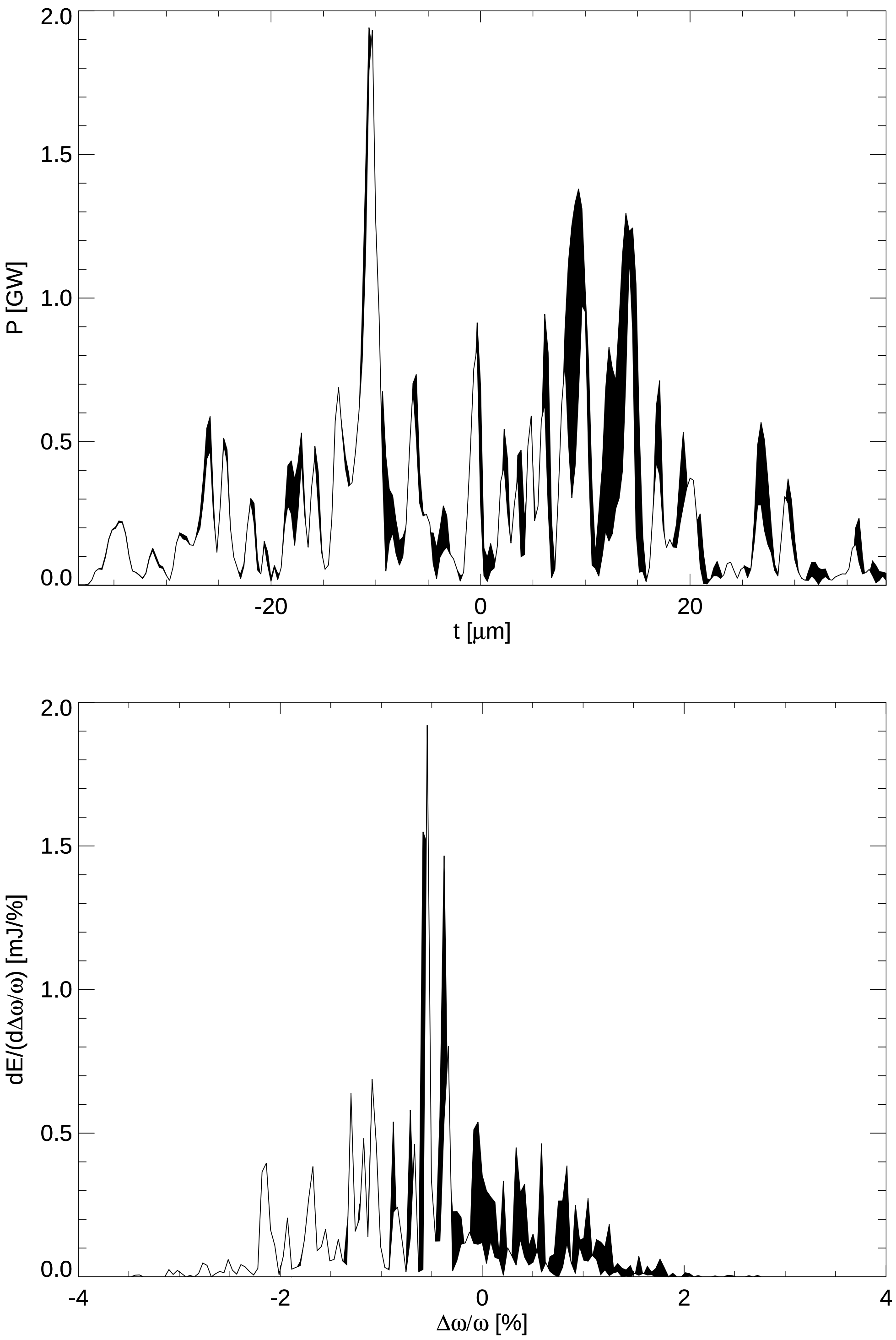


Figure 5.18: *Typical radiation pulse and spectrum of a SASE FEL after 18 m. The black envelope indicates the difference with the results excluding wake fields.*

Chapter 6

The X-Ray Free-Electron Laser at the TESLA Linear Accelerator

One of the major goals for Free-Electron Lasers is to generate intense radiation in the X-ray region, which cannot be reached by conventional lasers using mirrors. This is achieved by accelerating the electrons to high energies. A resonant radiation wavelength λ_r below 1 Å demands a beam energy above 10 GeV according to

$$\lambda_r = \frac{\lambda_U}{2\gamma^2}(1 + K^2) \quad , \tag{6.1}$$

where λ_U is the undulator period length, γ the electron energy and K the dimensionless undulator parameter. Within the TESLA linear collider project [20] with a center of mass energy of 500 GeV, one of the two linacs can easily provide this beam energy to drive an Å FEL. The electron beam is extracted from the accelerating beam line of superconducting cavities at an energy of up to 50 GeV and is transported to the undulator. Besides the general problem of providing a sufficient beam quality for the FEL such as high peak current and small emittance, the designers and users of the FEL are confronted with additional problems in this wavelength region.

The nature of X-rays excludes the operation of the FEL as an oscillator or amplifier, because neither mirrors nor gratings exist to reflect the radiation sufficiently nor can a seeding field be generated by conventional radiation sources in this wavelength region. Only the Self-Amplified Spontaneous Emission FEL remains as a realizable device. To keep the undulator length within a reasonable limit the constraints for the peak current, energy spread and beam emittance are very high. The energy of the driving electron beam is limited to 50 GeV because beyond this energy the quantum fluctuation of the incoherently emitted radiation yields a growth in the energy spread of the electron beam.

The diffraction of the highly brilliant X-ray beam is very small so that several hundreds of meters drift space must follow the undulator to enlarge the beam spot up to a point most suitable for the X-ray experiments. Section 6.1 presents and discusses several methods to enhanced the diffraction of the X-ray beam and thus reduce this length of drift space.

Another limiting factor of X-ray FELs is the transverse emittance of the electron beam. The

spread in the transverse motion reduces the synchronization of the electrons with the radiation field and degrades the rate of amplification. The impact of the emittance can be partially compensated by a correlation between energy and the average transverse momentum as is described in Section 6.2.
The impact of a higher beam energy is given in Section 6.3. It includes the degradation of the FEL performance due to the increase of the energy spread by the quantum fluctuation of the incoherently emitted radiation.
For all simulations the same parameter set for the TESLA FEL undulator has been used. The focusing of the electron beam is done by an alternating sequence of undulator modules and focusing elements, either single quadrupoles or triplets. In the simulations the first method has been chosen throughout with a quadrupole length of 0.4 m and an undulator module length of 2.2 m. The root-mean-square beam size is 20 μm in both transverse directions.
For a beam energy of 25 GeV the resonant radiation wavelength is $\lambda = 2.4$ Å and the estimated shot noise power of the fluctuation in the electron positions is 11.5 kW. Saturation occurs after 90 m, demanding at least 34 undulator modules. If the FEL operates under these conditions the root-mean-square opening angle of the intensity directivity is 1.3 μrad at saturation and 1.9 μrad in the linear regime of the FEL, when the amplification would cease at this point.

6.1 Enhanced Diffraction for X-Ray Free-Electron Lasers

At short radiation wavelengths in the VUV or X-Ray regime the diffraction of the radiation field plays a less dominant role and the FEL amplification operates almost in one dimension. This is not necessarily a positive aspect because the growth of the transverse coherence is inhibited. The information of the radiation field amplitude and phase at a certain transverse position is poorly transported by the reduced diffraction to other regions in the transverse plane. A quantitative estimate of this behavior is the diffraction parameter B (see Section 2.6) which is the ratio between the Rayleigh length and the gain length. For a value larger than 10 the 3D FEL theory predicts comparable growth rates for all eigenmodes of the FEL amplification. This drawback for achieving transverse coherence is relaxed if the transverse betatron oscillation is taken into account. The higher modes are more sensitive to any variation of the beam size and their amplification is reduced.
Another problem, arising from small diffraction, has a practical origin. The X-Ray FEL radiation is an intense beam localized into a narrow solid angle and frequency range, which is the property of a high brilliance radiation source. The power density might exceed the limits of the adjacent experiments. The radiation could harm the probe, mirrors or gratings. Also the radiation cone may not illuminate the full area of the gratings or probe surface. A drift length of 430 m is needed to widen the beam to an incidence area of 1 mm^2 for the radiation properties of the TESLA FEL at a beam energy of 25 GeV.
The problem to decrease the radiation power and/or to widen the opening angle of the radiation beam is difficult because certain radiation parameters cannot be changed without degrading others. In particular 'switching off' the FEL process at an early stage of amplification has the major drawback of a wider bandwidth of the radiation spectrum and a larger fluctuation of the

radiation power.
The solution to put a small aperture in the way of propagation in order to increase diffraction and to reduce intensity is impractical due to the X-ray radiation causing serious damage. The operational life time of such an insertion device would be short and thus require frequent replacement. A similar more promising approach is to place a liquid or gaseous absorber in the radiation beam. The deposited power can be cooled by the gas or liquid. Despite this the method does not enhance the diffraction of the FEL radiation and it is not included in the discussion.
If the undulator is succeeded by a long drift space in the range of hundreds of meters for an X-ray FEL the natural diffraction reduces the intensity to the desired level of the experiment. This is the straight forward solution but demands a large facility site for the FEL and the experiments, where most of the space is 'wasted' by drift spaces.
To reduce the drift length and thus the cost for the construction, four different methods are presented and discussed in this section. All of them have in common that the increased diffraction is produced by influencing the FEL amplification process. They do not rely on passive elements in the beam line such as a small aperture.
The TESLA-FEL is designed to operate in an electron energy region between 10 and 50 GeV. A 10 GeV beam energy corresponds to a resonant wavelength of 15.2 Å and a saturation length of about 40 m. The root-mean-square opening angle of the radiation cone is 6 μrad. With the model of alternating undulator modules and quadrupoles the FEL is designed to achieve the shortest wavelength of 2.4 Å for an electron beam energy of 25 GeV. At this limit the diffraction of the radiation field is the smallest and a control of the radiation intensity is demanded. Therefore the results of the simulations, presented in this section, are done for the shortest wavelength.

6.1.1 Longitudinal Mismatch of Undulator Modules

If an undulator is built up from several modules the separating drift spaces cannot be chosen arbitrarily. Otherwise the synchronization between the electron beam and the radiation field would not be conserved when both enter the next module. Besides a negligible diffraction of the radiation field in the drift space because the gain guiding of the FEL amplification does not apply there, the electrons move with a different longitudinal velocity due to the missing transverse oscillation. To restart the FEL amplification at the entrance of the next module under the optimum conditions the modules must be aligned to each other with a tolerance smaller than the undulator period.
The worst case regarding the FEL amplification is a complete mismatch, where an artificial phase shift of $\Delta\theta = \pi$ in the electron ponderomotive phase has been introduced. The energy transfer of the amplification from the electron beam to the radiation field is inverted and the field is weakened. In addition the forming of the electron beam modulation — the micro bunching — is not driven by the radiation field anymore. The FEL interaction reacts to such a disturbance by adapting the radiation field to the new phase of the electron bunching. The scale of this phase change is the gain length of typically several meters.
The sudden and coherent change in the electron ponderomotive phase after a mismatched

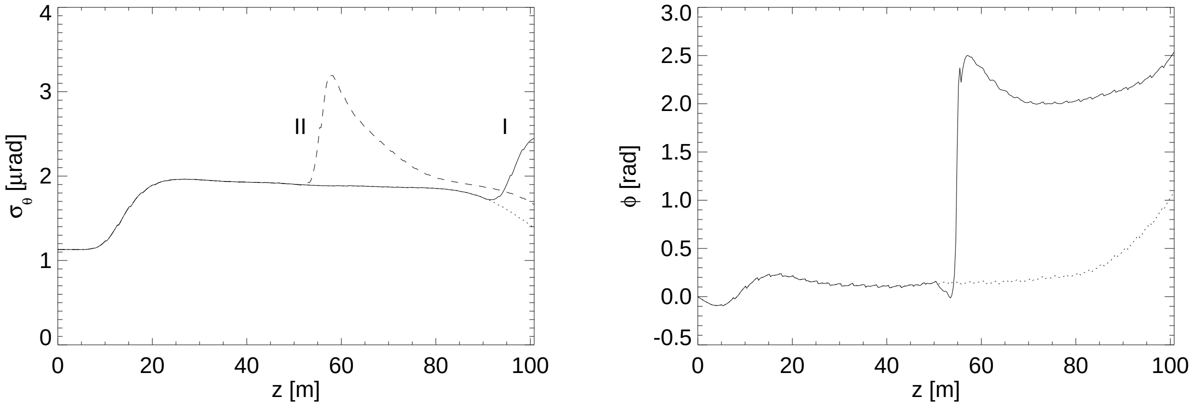


Figure 6.1: *Root-mean-square diffraction angle along the undulator for a mismatched undulator module close to the end or in the linear amplification regime (curve I and II in the left plot, respectively). The undisturbed case is drawn by a dotted line. The change in the radiation phase for a mismatch at $z = 50$ m is shown in the right plot.*

undulator module is connected with an enhancement of the diffraction. The Gaussian-like radiation profile of the undisturbed FEL amplification is always larger than the electron beam, in particular after a drift space, where the field diffracts due to the missing gain guiding. With a phase difference in the ponderomotive wave of $\Delta\theta = \pi$ the electron beam extracts almost all the energy from the radiation field. The radiation amplitude at the location of the electrons is significantly reduced to a level where the coherent emission of the electron beam dominates. Then the radiation field is built up again but at the new ponderomotive phase. The remains of the former radiation field, which has no overlap with the electron beam, forms a 'ring'-like radiation profile. While it diffracts, the FEL amplification starts over again with the newly built up radiation field within the ring. After several gain lengths the FEL process has converged back to the undisturbed case.

Because the radiation field distribution differs from that of the fundamental mode, the diffraction is enhanced. In Fig. 6.1 the root-mean-square diffraction angle is plotted versus the position along the undulator. The value of the diffraction angle is only meaningful if the undulator would end at that position. Two cases are considered where the undulator mismatch is either placed in the saturation or in the linear regime of the FEL process (curve I and II in the plot, respectively).

Eight meters behind the mismatched undulator module the enhanced diffraction has its maximum before it is slowly reduced back to the undisturbed value by the restarted FEL amplification. For the case of a mismatch in the linear regime the power level of the newly amplified field dominates above the diffracted 'ring' after 20 m. The disturbance of the FEL is hardly noticeable anymore except for an increase of the saturation length.

In the right plot of Fig. 6.1 the phase of the radiation field at the undulator axis is shown. The fast change of the radiation phase is clearly visible. If the undulator ends at maximum

diffraction, which occurs 8 m after the beginning of the mismatched undulator module, the intensity in the far field zone will be reduced by a factor of three.
Even for perfectly matched undulator modules the diffraction is reduced in the saturation regime, because the fundamental mode of the FEL amplification approaches that of free space propagation. It takes longer ($\approx$20 m) to reach the diffraction maximum value. The reduction of the intensity is roughly one third and thus the efficiency is comparable with the perturbation in the linear regime of the FEL.
Due to the shift in the electron phases a fraction of the radiation power is reabsorbed by the electron beam. The radiation loses up to 80% of the power. However the overall gain is still larger than six orders of magnitude.

6.1.2 Electron Delay by Transverse Motion

Another method works similarly to the method of a mismatched undulator module. Both have in common that the electrons phases are changed suddenly relative to the radiation field. While a mismatched module determines the electron phase at the entrance of the module by the distance to the previous module, the electrons can also be delayed by applying a short but strong transverse coherent motion using a pair of steering magnets. For the simulations the two steering magnets are separated by 0.4 m. This distance is small compared to a gain length and the change in the electron phase is almost instantaneous ('phase shake', see Section 5.1).
With a deflection angle of 15 μrad the generated transverse offset of the electron beam is not larger than 8 μm and the FEL amplification process is only slightly reduced by the off-axis position and the succeeding betatron oscillation. A set-up of three steerers would even avoid this excitation of the betatron oscillation.
Because the applied phase shift is similar to the first method the results are also similar, as can been seen in Fig. 6.2. The only visible difference is that the radiation field is not axi-symmetric in the near field as well as in the far field region.

6.1.3 Reduction of the Transverse Electron Beam Size

A different approach is to modify those parameters which determine the magnitude of diffraction. A rough estimate is given by the Rayleigh length

$$z_R = \frac{\Sigma k}{2} \quad , \tag{6.2}$$

where k is the radiation wavenumber and Σ is the radiation spot size. The value of Σ can be approximated by the electron beam size $\Sigma_e \approx \pi\sigma_x\sigma_y$, which is for an X-ray FEL slightly smaller than Σ. The root-mean-beam sizes $\sigma_{x,y}$ in the x- and y-direction, respectively, vary for the chosen quadrupole lattice only by 14% and the eccentricity of the beam spot is small. Because the radiation wavenumber k is fixed, the only possible way to shorten the Rayleigh length, and thus to increase the diffraction, is to reduce the spot size.
According to the conceptional design report of the TESLA Free-Electron Laser [20] the quadrupoles can be operated up to a gradient of 100 T/m. To obtain the design values of the beam

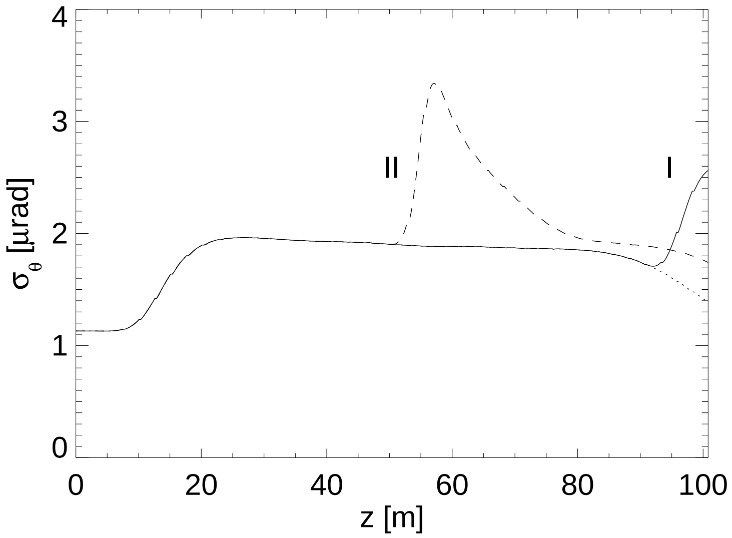


Figure 6.2: *Root-mean-square diffraction angle along the undulator for an electron beam, strongly steered in the transverse direction close to saturation (curve I) and in the linear regime (curve II) of the FEL amplification. The undisturbed case is drawn by a dotted line.*

sizes $\sigma_{x,y} \approx 20$ μm for a normalized emittance of 1 π mm-mrad, a quadrupole length of 0.4 m and a drift length of 2.2 m, the gradient has to be set to 35 T/m.
In principle the best idea seems to be to maximize the focusing strength over the entire undulator length to improve the overall FEL performance, but this is not advised for three reasons. With a smaller beam size the transverse momentum spread is larger and the radiation field growth rate is reduced because the electrons are less synchronized with respect to the radiation field. Also the centroid motion, caused by quadrupole misalignment or undulator field errors, is a more critical issue to conserve the overlap between the electron beam and the radiation field. As a third point the matching of an electron beam to a smaller beta function of the quadrupole lattice is more sensitive to any errors or wrong settings of the focusing magnets.
With a single quadrupole it is impossible to focus the electron beam in both planes. To keep the spot shape almost round the field gradient of the quadrupoles are increased quadratically over the last quarter of the undulator. The initial and final gradients are 35 T/m and 100 T/m, respectively. The resulting beam sizes in the x- and y-plane are drawn in the left plot of Fig. 6.3. With the chosen quadrupole lattice the electron beam spot size, and thus the Rayleigh length, is reduced to one quarter of its initial value. In the case that the electron beam is focusing during the exponential growth in the linear regime of the FEL the simulation is stopped at maximum diffraction. After the minimal spot size has been obtained the diffraction angle would converge back to the undisturbed value if the beam were relaxed to the initial size.
The initial condition for the simulation has been altered by using a ten times larger radiation power. It shortens the integration length by 15 m without changing the results. In this way it

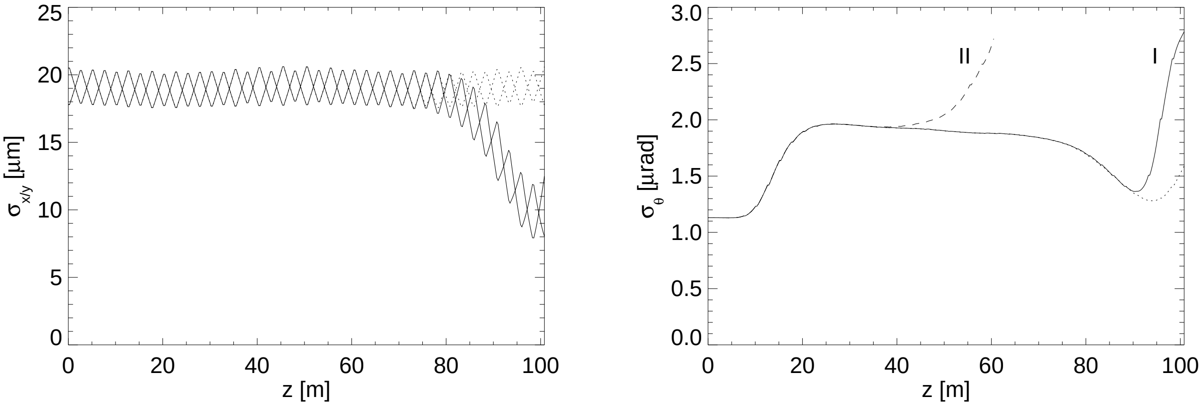


Figure 6.3: *Electron beam size and diffraction angle (left and right plot, respectively) along the undulator for an increasing focusing strength at the end of the undulator. For the diffraction angle a second case (curve II) of a shorter undulator (60 m) has been considered.*

is easier to compare the results with the other methods. Using the same input power 7 more modules with a length of 2.2 m each are needed. The reason is that the diffraction is less sensitive to the beam spot size in the linear regime or in the transition to the saturation regime than to the change of the electron ponderomotive phases. Only in the far saturation regime has this method the same efficiency of reducing the intensity in the far field range by a factor of 3. The benefit of this method is that no radiation power is extracted by the electron beam. The saturation power is conserved.

6.1.4 Enhancement of the Electron Beam Divergence

The last method, discussed here, is motivated by a more subtle argument. If two radiation distributions have the same size in the intensity profile, they do not necessarily have the same diffraction. For a complete determination of the far field the dependence of the radiation field amplitude and phase on the transverse position is of importance. The radial dependence of the phase must be stronger than that of the amplitude to significantly enhance the diffraction. In the opposite case where the phase is constant over the entire transverse plane the opening angle of the far field has its smallest value. Focusing the electron spot size, as described above, has some influence on the radiation size but almost none on the phase variation in the radial direction. Indeed it is very difficult to generate a correlation between the transverse position of the electrons and the ponderomotive phase. The only possible way to think of is to let the electron beam diverge faster than the diffraction of the radiation field, assuming that the radiation field would somehow 'follow' the electron beam.
Despite the problem how to realize a fast diverging electron beam, the overall efficiency of increasing the diffraction is poor. If the divergence of the electron beam is comparable to or slightly stronger than the radiation field the resulting diffraction is even reduced. It is

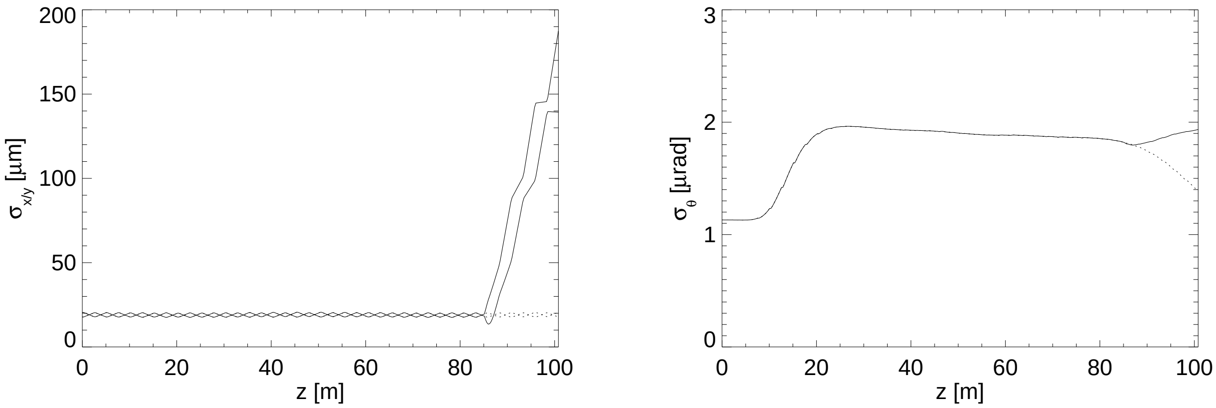


Figure 6.4: *Electron beam size and diffraction angle (left and right plot, respectively) along the undulator for a fast diverging electron beam.*

the inverse process of the previously described method of focusing the electron beam, where the diffraction is mainly defined by the beam size instead of the transverse variation in the ponderomotive phase.

An extremely divergent electron beam does not give any better results as can be seen in Fig. 6.4. The only visible effect is the instantaneous 'turn off' of the the FEL amplification process.

6.1.5 Comparison of the Methods

Regarding the different methods to enhance the diffraction of the FEL radiation each of them has some positive as well as negative aspects. The most inefficient way is to force a strong divergence of the electron beam, which is hardly better than just to simply shorten the undulator by the length where the electron beam is diverging. The two methods of a mismatched undulator and the forced delay of the electron are almost identical concerning the efficiency. The mismatching of a given undulator module can be done with no extra cost. The consequence is that this fixed solution can hardly be altered for changing beam properties. The delay of the electron beam relative to the radiation wave is done by steering magnets. Even if such a phase delay is not considered for the operation of the FEL steering magnets are placed all along the undulator. They are needed for the beam-based alignment to compensate the beam wander, caused by undulator field errors or quadrupole misalignments. Focusing the electron beam to the end of the undulator yields almost the same results as the other methods but has the benefit that no power is extracted from the radiation field.

Another issue is the intensity distribution in the far field region itself. Fig. 6.5 shows the intensity profiles for all methods presented. A smooth distribution is generated by the mismatch of an undulator module. The radiation profile resulting from transverse coherent motion of the electron beam is similar but slightly asymmetric.

A richer pattern is visible for the case of excessively focused beam. The far field profile consists

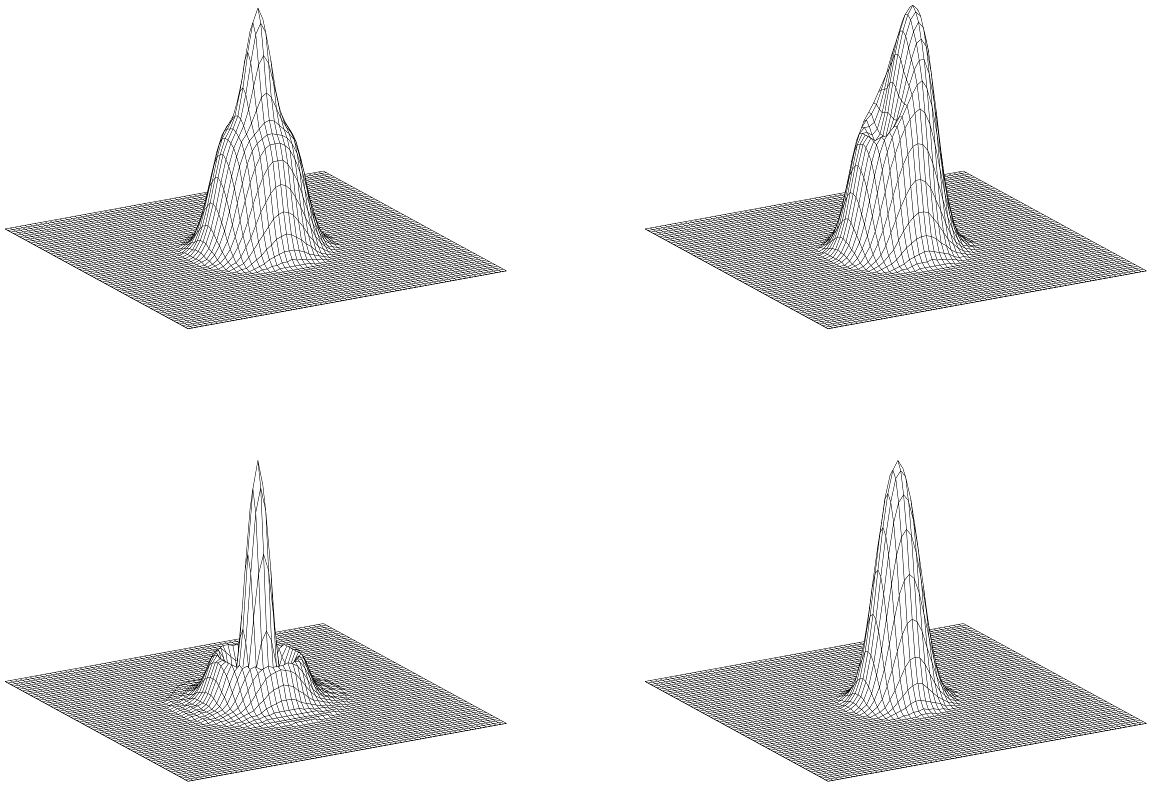


Figure 6.5: *Far field intensity profiles [a.u.] for different methods to increase the diffraction by undulator mismatch (upper left), by delayed electrons due to applied transverse motion (upper right), by a focused or divergent electron beam (lower left and right, respectively).*

mainly of a narrow cone and a wider ring around. Suitable for an experiment which needs to illuminate a certain area homogeneously, such an incident field is not well. This is the major disadvantage of this method.
For completeness the far field profile for the case of the diverging beam is plotted. The distribution is in a good approximation identical to the profile if the undulator ends in the linear regime of the FEL amplification. It is narrower than all other profiles unless the inner cone of the focusing method is regarded alone.
In conclusion steering the electron beam in order to change the ponderomotive phase relatively to the radiation phase shows the best performance because it provides the largest reduction of the intensity with a smooth profile in the far field zone while being fully controllable during runtime of the Free-Electron Laser.

6.2 Compensation of the Axial Electron Velocity by a Correlated Energy Spread

One of the major problems for a X-ray Free-Electron Laser is to provide high quality electron beam parameters, in particular for the energy spread σ_γ and the transverse normalized emittance ϵ_N. The main working principle of any FEL is the reorganization of the longitudinal phase space distribution to obtain a bunching in the longitudinal position and thus an enhancement of coherent radiation at the resonant wavelength λ_r. Because the phase space density is conserved during the FEL interaction, the bunching is connected with a simultaneously increased energy spread. The amplification stops at saturation when the energy spread covers the complete bandwidth $\pm\rho\gamma_0$ of the FEL amplification (Eq. 2.95), where γ_0 is the beam energy and ρ is the FEL parameter (Eq. 2.69).

If the initial energy spread of the electron beam is already larger than ρ the FEL amplification cannot be started. This yields the constraint

$$\frac{\sigma_\gamma}{\gamma_0} \ll \rho \tag{6.3}$$

for the energy spread. Unless the energy spread is not significantly increased during the acceleration of the electron beam this constraint does not become more stringent for higher energies because the FEL parameter ρ drops as γ_0^{-1} or slower if the transverse beam size is kept constant. The situation changes if the spread in the axial velocity of the electron betatron oscillation is analyzed. The transverse motion reduces the longitudinal velocity

$$v_z = c\left[1 - \frac{1+K^2}{2\gamma^2} - \frac{p_x^2+p_y^2}{2\gamma^2}\right] \quad , \tag{6.4}$$

where K is the undulator parameter and $p_{x,y}$ are the canonical momenta in the transverse directions normalized by mc. Regarding all electrons the longitudinal velocity v_z is smeared like for a non vanishing energy spread. In fact the transverse betatron motion can be approximated in this treatment by an artificial energy spread. As long as the transverse motion and the energy spread are uncorrelated the effective energy spread is quadrature sum (see Eq. 2.96). The part of the energy spread belonging to the transverse motion has to satisfy the condition 6.3 as well. This yields the constraint for the normalized emittance with

$$\epsilon_N \ll \frac{4\rho\beta\gamma_0\lambda_r}{\lambda_U} \quad . \tag{6.5}$$

While the normalized emittance ϵ_N is constant in the linear optics of the electron beam transport, the beta function β and the resonant wave length λ_r depend on γ_0 so that the product $\beta\gamma_0\lambda_r$ is almost independent of the energy. Therefore the electron beam cannot be accelerated to any arbitrary energy without violating the inequality Eq. 6.5 above a certain energy threshold.

Energy spread and the spread in the axial velocity are, under normal circumstances, uncorrelated and will be added up unless the beam is conditioned in order to generate a correlation between energy and transverse position.

The simple model behind the correlation is that the slowing down of the longitudinal motion due to the transverse motion is compensated by a higher energy. Nevertheless the compensation cannot be perfect because p_x and p_y are changing with the betatron oscillation. If no external focusing is added to the natural focusing of the undulator the momentum oscillates as $p_x(z) = P_x \sin(k_\beta z + \phi_0)$, where P_x is the amplitude of the transverse momentum, k_β is the wave number of the betatron oscillation and ϕ_0 is the initial phase. For this case the variation in the transverse momentum is completely compensated by the dependence of the undulator field K on the transverse position (Eqs. 2.12 and 2.20) and the longitudinal velocity is constant over the entire undulator length [26]. This is the basis for increasing the energy proportional to $P_x^2 + P_y^2$ and thus removing any spread in the longitudinal velocity caused by a large emittance [124, 125].

For X-ray Free-Electron Lasers the natural focusing is not strong enough to provide sufficiently small beam sizes. Therefore the undulator is superimposed with a quadrupole lattice of alternating polarity. One disadvantage of such strong focusing is that the variation in the transverse momenta is no longer compensated by the undulator field. To what extent the compensation scheme is applicable for a strong focusing lattice is discussed in this section.

For the TESLA-FEL with a quadrupole length of 0.4 m and a drift space of 2.2 m the calculation of the betatron motion is preferably described by the thin lens approach. In this model the quadrupoles act solely on the momentum of the electron while the transverse position is kept constant.

The treatment of the betatron motion is based on the optical functions β, α and γ [28] with the identity $\beta\gamma - \alpha^2 = 1$. Unless mentioned otherwise the following discussion refers to γ as an optical function and not as the beam energy. The quadrupoles are separated by the distance L and the lattice starts at $z = 0$ with focusing in the x-direction. For the drift section to the defocusing quadrupole ($z < L$) the formal solution

$$x(z) = \sqrt{I_x \beta(z)} \sin(\Psi(z) + \phi_0) \tag{6.6}$$

is given by

$$\beta(z) = \beta_0 - 2\alpha_0 z + \gamma_0 z^2 \tag{6.7}$$

and

$$\Psi(z) = \arctan(\gamma_0 z - \alpha_0) + \arctan(\alpha_0) \quad , \tag{6.8}$$

where the index '0' refers to the initial values of the optical functions right behind the focusing quadrupole and I_x is the Courant-Snyder invariant of the betatron motion [126]. For this special lattice, where the drift spaces fill out the entire longitudinal space, γ_0 is also a constant of motion. The orbit angle x', the ratio between transverse and longitudinal momentum, is given by

$$x'(z) = \sqrt{\frac{I_x}{\beta_0}} \left[\cos(\phi_0) - \alpha_0 \sin(\phi_0)\right] \quad , \tag{6.9}$$

which is independent of z as is always the case for a drift. In the second half of the quadrupole lattice cell with $L < z < 2L$ the solution for x' is

$$\begin{aligned} x'(z) &= \sqrt{\frac{I_x}{\beta_1}}\left[\cos(\phi_1) + \alpha_1 \sin(\phi_1)\right] \\ &= \sqrt{\frac{I_x}{\beta_0}}\left[\left(1 + \frac{2\alpha_1 L}{\beta_1}\right)\cos(\phi_0) + \left(1 - \frac{2L}{\alpha_1\beta_1}\right)\alpha_0 \sin(\phi_0)\right] \end{aligned} \tag{6.10}$$

with $\beta_1 = \beta(L)$, $\alpha_1 = \alpha_0 - \gamma_0 L$ and $\phi_1 = \phi_0 + \Psi(L)$. The orbit angle of the betatron motion consists of two terms. The cosine term has the same sign for both drift sections of the lattice cell and adds up to the average betatron motion with its given periodicity and amplitude comparable to the natural focusing of the undulator. The sine term has different signs for the two drift sections after a focusing and defocusing quadrupole, respectively. The net change in the electron position is less dominant than for the cosine term as can be seen in the left plot of Fig. 6.6 for the electron trajectory in the undulator of the TESLA FEL.

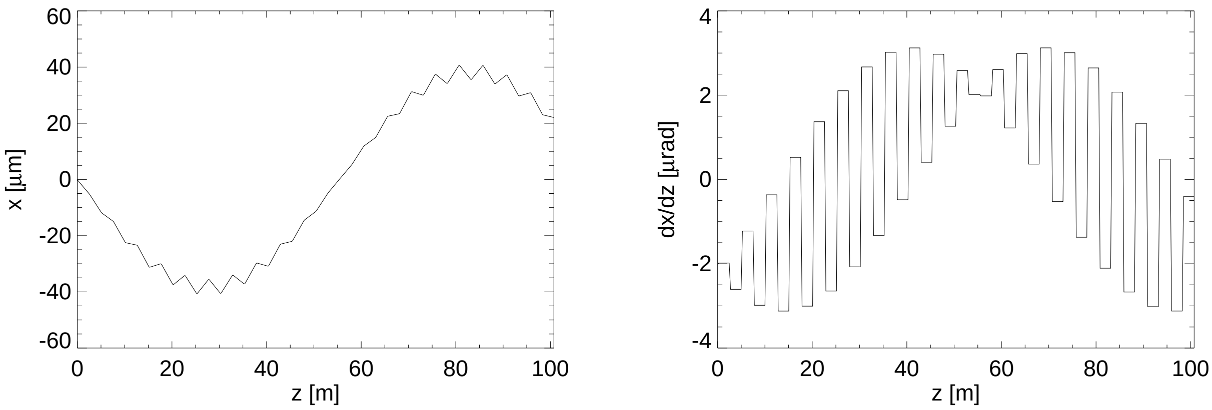


Figure 6.6: *Transverse orbit (left) and orbit angle (right) of an electron, moving through the strong focusing quadrupole lattice of the FEL undulator at TESLA.*

Although this 'sawtooth'-like oscillation has a small impact on a significant change in the transverse position it stabilizes the synchronization of the longitudinal velocity (Eq. 6.4). For a focusing lattice with a rather small phase advance per cell the values of the optical function α_0 and α_1 are of the order of unity with

$$\alpha_{0,1} = \sqrt{\frac{1 \pm \frac{L}{\bar\beta}}{1 \mp \frac{L}{\bar\beta}}} \approx 1 \pm \frac{L^2}{2\bar\beta^2} \tag{6.11}$$

where $\bar\beta = (\beta_0 + \beta_1)/2$ is the average beta function identical with the focal length of the quadrupoles. The upper sign denotes the value of α_0, the lower of α_1, respectively. As a result

the variance of the transverse momenta $< p_x^2 + p_y^2 >$, averaged over one betatron oscillation, is enhanced by a factor of 2 compared to the natural focusing of the undulator. In addition the square of the transverse momenta fluctuates rapidly with a characteristic length of the lattice cell, while the long term variation of the betatron oscillation is strongly suppressed because the cosine and sine cross terms in the square of Eqs. 6.9 and 6.10 cancel each other. The remaining oscillation is of second order $\mathcal{O}([2L/\bar{\beta}]^2)$, which is about 6% for the TESLA undulator. Because the length of a focusing lattice cell is about three times shorter than the gain length (approximately 14 m) the axial velocity term $p_x^2 + p_y^2$ can be regarded as almost constant in Eq. 6.4.
The desynchronization in the longitudinal velocity can be compensated if each electron has a shift in its energy γ_0 proportional to the Courant-Snyder invariant, namely

$$\Delta\gamma = \frac{\lambda_U \gamma_0}{4\lambda\bar{\beta}}(I_x + I_y) \quad . \tag{6.12}$$

In Eq. 6.12 the scaling factor is the same for the compensation of the motion in the x- and y-plane because the average betatron function $\bar{\beta}$ is dominated by the strong focusing and thus almost identical for each plane. The influence of the natural undulator focusing can be ignored.

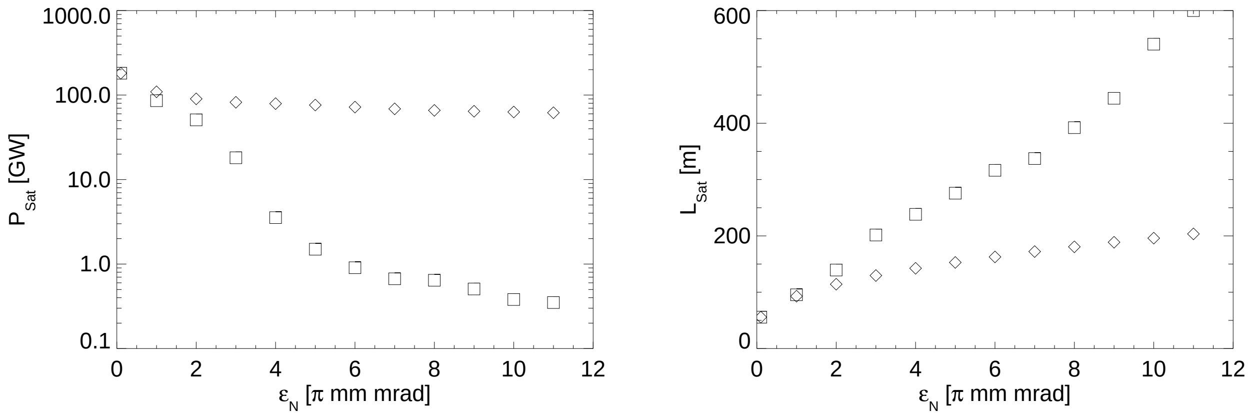


Figure 6.7: *Saturation power and length (left and right, respectively) versus normalized emittance for a conditioned (◇) and unconditioned (□) electron beam.*

Assuming that the electron beam has been conditioned in order to fulfill Eq. 6.12 simulations are done to study the applicable range of the emittance compensation scheme. Fig. 6.7 shows the saturation power and length of the TESLA X-ray Free-Electron Laser for an electron beam energy of 25 GeV and various transverse emittances. The design emittance is $\epsilon_n = 1$ π mm mrad for both transverse planes. Without conditioning the FEL saturates in less than 100 m. At these design values the FEL performance is hardly degraded by the emittance effect and a conditioning of the beam is not necessarily needed.
Above 1 π mm mrad the efficiency of the FEL drops rapidly. For an emittance $\epsilon_N = 11$ π mm mrad the saturation power is two orders of magnitude smaller and the FEL saturates

beyond 600 m. If the correlation between energy and betatron amplitude is applied the FEL performance is improved significantly, although the saturation power and length remain worse than for the $\epsilon_n = 1\ \pi$ mm mrad case.
The reason lies in a larger beam size due to the higher emittance and thus in a reduction of the electron density. The FEL parameter ρ depends on the emittance as $\epsilon_N^{-1/3}$ and thus affects the saturation power and length approximately as $\epsilon_N^{-1/3}$ or $\epsilon_N^{1/3}$, respectively [56]. This agrees with the results for the conditioned beam in Fig. 6.7.

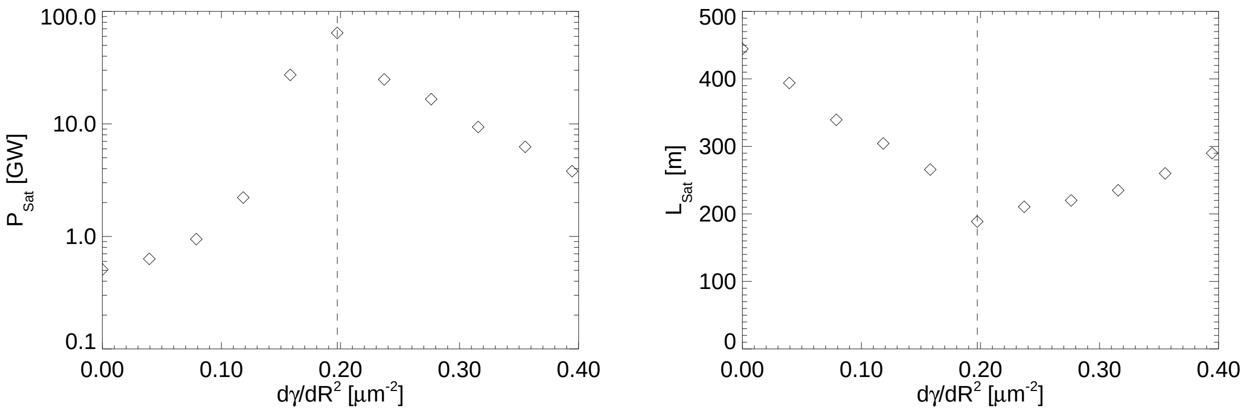


Figure 6.8: *Saturation power and length (left and right, respectively) for different correlation strengths between energy and the square of the maximum betatron amplitude R^2 of the electrons. The optimum correlation yields $d\gamma/dR^2 = 0.197\ \mu m^{-2}$ and is indicated by the dashed line. The root-mean-square radius of the beam is 81 μm corresponding to a normalized emittance of $\epsilon_N = 9\pi$ mm mrad.*

To estimate the efficiency of the compensation scheme the results for the conditioned electron beam are compared with simulations, where the axial velocity term has been disabled in the differential equation for the electron longitudinal position. The difference in the saturation power is less than 10% and in the saturation length even less than 5%. This indicates that the correlation factor $\lambda_U\gamma_0/4\lambda_r\bar{\beta}$ provides the optimum performance. Any stronger or weaker correlation would reduce the efficiency as can be seen in Fig. 6.8.
It is quite surprising that the results for saturation power and length are asymmetric concerning the degree of correlation and that exceeding the optimum correlation is better than a less than optimal conditioning. This can be understood if the motion in the longitudinal phase space is analyzed. If the electron beam is less than optimally conditioned the tail in the phase space distribution is slower than the core. The equivalent energy distribution is asymmetric with a tail towards lower energy. In the case of overconditioning this energy distribution is inverted.
For comparison a simplified effective energy distribution

$$f_0(\eta) = \frac{1}{2\Delta}\left[1 - \frac{\eta}{\Delta}\right] \quad , \tag{6.13}$$

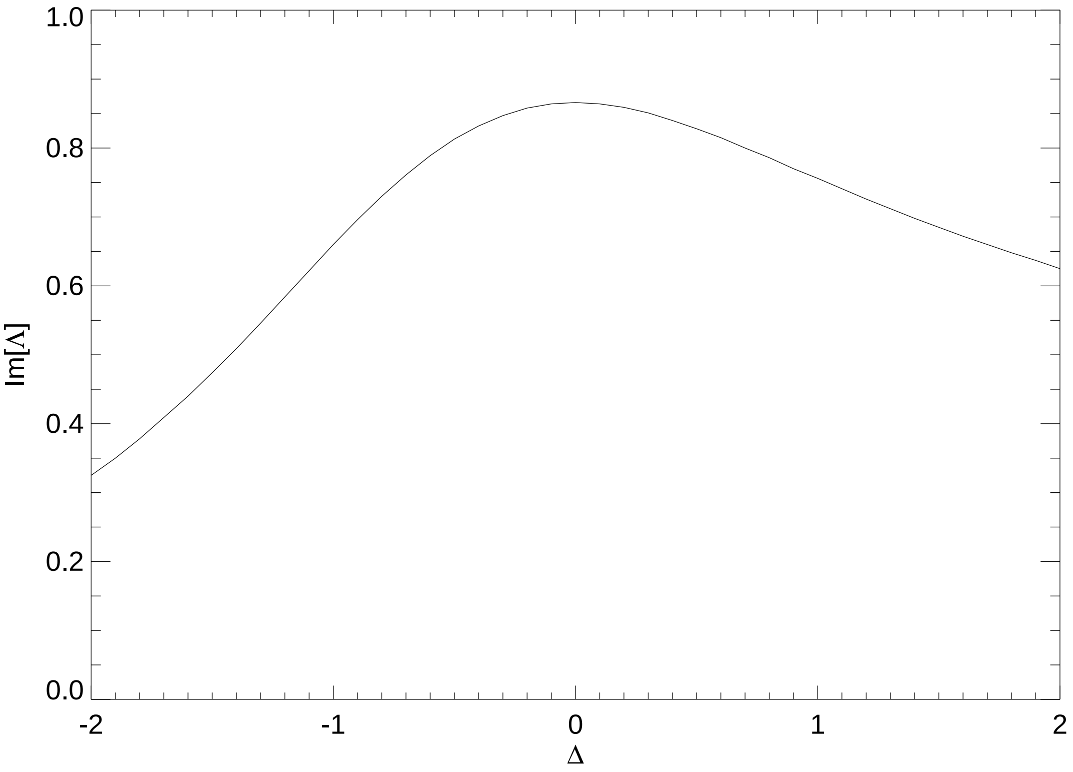


Figure 6.9: *Dependence of the maximum growth rate on the asymmetric energy distribution* $f(\eta) = 1/2\Delta(1 - \eta/\Delta)$ *with* $\eta = (\gamma - \gamma_0)/\rho\gamma_0$ *and* $-|\Delta| \leq \eta \leq |\Delta|$ *in the 1D FEL model.*

is regarded in the 1D FEL model with $\eta = (\gamma - \gamma_0)/\rho\gamma_0$. A negative value of Δ indicates a tail at lower energy as is the case for an underconditioned beam. In the limit $\Delta \to 0$ the effective energy distribution $f_0(\eta)$ becomes a Dirac function, corresponding to a perfect correlation between electron energy and transverse position. If the correlation is stronger than perfect the distribution gets wider again with $\Delta > 0$. The exponential increments of the radiation field are defined by the dispersion equation (Eq. 2.83)

$$(\Lambda - \delta)(\Lambda^2 - \Delta^2) + 1 + \frac{\Lambda}{\Delta} - \frac{\Lambda^2 - \Delta^2}{2\Delta^2} \ln\left[\frac{\Lambda + \Delta}{\Lambda - \Delta}\right] = 0 \quad , \tag{6.14}$$

where δ is the detuning of the electron beam energy relatively to a fixed resonant wavelength. The dependence of the maximum growth rate $\Im\mathrm{m}(\Lambda)$ is shown in Fig. 6.9. The general shape agrees quite well with the saturation power in Fig. 6.8 even for this rough estimate of the 1D FEL model.

A qualitative explanation for the asymmetry comes from the motion in the longitudinal phase space. The bucket of the radiation phase traps a certain part of the electron beam and forces it to rotate within the bucket to form the micro bunching. For an overconditioned beam the core of the phase space distribution lies close to the resonant energy while the tail populates the phase space in the bucket towards higher energies. During FEL interaction, which is mainly driven by that part of the distribution with the highest density, the tail is rotated towards lower energies. This corresponds to an energy transfer from the electrons in this tail to the radiation field and thus supports the FEL amplification by the core electrons. In the other case, where the tail of the distribution lies in the lower part of the bucket, the electrons in the tail weaken the radiation field and reduce the efficiency of the amplification.

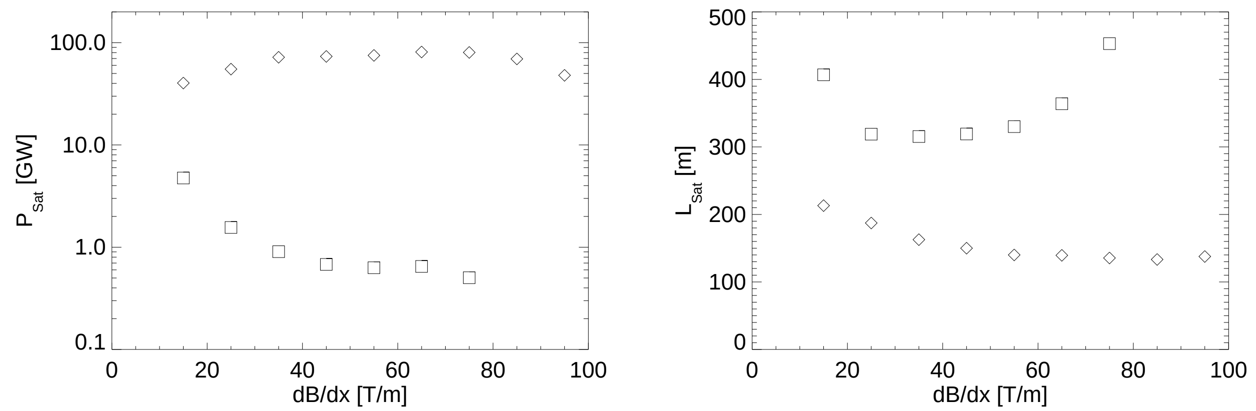


Figure 6.10: *Saturation power and length (left and right, respectively) versus field strength of the quadrupole lattice for a conditioned (◇) and unconditioned (□) electron beam with a normalized emittance of* $\epsilon_N = 6\pi$ *mm mrad.*

The optimum strength of the correlation between the energy and the betatron amplitude depends on the period length of the undulator, the radiation as well as on the betatron oscillation. Only the last parameter can be varied in a limited range by changing the quadrupole field strength. The nominal value of the quadrupole field gradient is 35 T/m to obtain the optimum beta function. Any lower or higher gradient reduces the efficiency of the Free-Electron Laser because either the electron density is lower for a weaker focusing or the degradation by the axial velocity spread dominates over the benefit of a higher density for a stronger focusing. If the beam is conditioned the latter argument becomes obsolete. Pushing the focusing to its theoretical limit to reduce the spot size as much as possible would give the best performance. The proposed quadrupoles of the TESLA FEL can operate at gradients up to 100 T/m.
Fig. 6.10 presents the results of the emittance compensation for various field gradients. The efficiency of conditioning the electron beam is improved at higher gradients, both for saturation length and power, although the saturation power tends to drop again for the largest gradients and a conditioned beam. For these field gradients the remaining fluctuation in the longitudinal velocity affects the amplification. One source is the large variation in the beta function. For gradients above 80 T/m the ratio between maximum and minimum value of $\beta(z)$ exceeds 2 and the correlation (Eq. 6.12) is not optimal anymore because the compensation scheme is based on the assumption that the deviation of the beta function from its average value is negligible.

In addition the short term averaging of $p_x^2 + p_y^2$ in Eq. 6.4 over a single cell of the quadrupole lattice is no longer constant. The assumption that the betatron phase advance per cell is small is violated. As a consequence α_0 and α_1 are not close to unity and the sine and cosine cross terms have a non negligible amplitude of $(2L/\bar{\beta})^2$. This slow modulation will be added up over several gain lengths similar to an uncorrelated energy spread which cannot be compensated.
For the optimum quadrupole gradient range between 35 and 75 T/m the change in the saturation power and length is weak. The compensation scheme is efficient if the emittance is large but not very useful to focus the beam to a smaller spot size, because the benefit for that is small.
The section is concluded with a possible method to correlate the energy with the Courant-Snyder invariants I_x and I_y. The easiest way to measure I_x and I_y and thus control the correlation is given rather by the average position of the electron in a periodic focusing lattice than by its transverse momentum. With the identity $I_x = 2 < x^2 > /\beta_x$ the correlation is introduced by accelerating the electron beam with a longitudinal electric field which depends quadratically on x and y. For a cylindrical cavity only the TM_{210}-mode provides this feature [43] with

$$E_z(x,y) = \bar{\mathcal{E}}(x^2 - y^2) \quad , \tag{6.15}$$

where $\bar{\mathcal{E}}$ is the effective accelerating gradient.
The schematic layout of the conditioning beam line is a quadrupole lattice, where TM_{210} cavities are placed between two quadrupoles with the same polarity. The phase difference of the electric field, as seen by the electrons, between two successive cavities is $\Delta\phi = \pi$. Assuming the same periodic betatron function in both planes but shifted half a period relative to each other the change in energy is

$$\Delta\gamma = \frac{eNL_{\mathcal{E}}\bar{\mathcal{E}}}{mc^2}\frac{\beta_{max} - \beta_{min}}{2}(I_x + I_y) \quad , \tag{6.16}$$

where $L_{\mathcal{E}}$ and N are the length and the number of the cavities respectively, β_{max} and β_{min} are the maximum and minimum values of the betatron function. The number of cavities has to be sufficiently large so that the electrons enter the accelerating structures under various betatron phases with

$$\frac{1}{N}\sum_{n=1}^{N}\sin(\psi_0 + 2n\Delta\psi) \ll 1 \tag{6.17}$$

for any arbitrary initial betatron phase ψ_0 and the phase advance $\Delta\psi$ over one cell. Otherwise a correction term has to be added to Eq. 6.16 because the averaged transverse position of the electron might deviate from $\sqrt{\beta_x I_x/2}$ due to the small number of 'samples'.
The total length of the conditioning beam line and the number of cavities depends on the degree of the correlation. For a normalized emittance $\epsilon_N = 9\pi$ mm mrad the average shift in energy is $< \Delta\gamma/\gamma_0 >= 2.7$ %. The explicit design of the conditioning section is driven by the achievable electric field in the TM_{210} cavities. The electron beam must be conditioned in an early stage of acceleration where the larger spot size of the electron beam is advantageous.

6.3 Wavelength Limitation by Incoherent Emission of Spontaneous Undulator Radiation

The advantage of a Free-Electron Laser is that the radiation wavelength is tunable by varying the electron beam energy. The radiation is not bounded to any transition of electrons between quantum states. Using beam energies in the GeV region the wavelength becomes shorter than 10 nm and the FEL radiates in the X-ray regime, which opens new perspectives of natural science using high brilliance X-ray radiation.

Besides the technical problem to accelerate the electron beam to the desired energy the FEL performance is affected or limited by phenomena which only occur at that wavelength region. One of them is the problem of the axial velocity spread for large emittance values. This can be partly compensated by conditioning the beam as is described in the previous Section 6.2.

Another degradation of the FEL performance is caused by wake fields. In order to keep the total length of the undulator within acceptable limits the pulse length is shortened to achieve a high peak current. As a negative effect the amplitudes of the wake fields are enlarged introducing a coherent energy spread along the bunch. The resulting reduction of the amplification is discussed in Section 5.2 for the VUV FEL at the TESLA Test Facility.

In addition the incoherent part of the spontaneous undulator radiation plays a significant role. A higher beam energy widens the complete emission spectrum and thus increases the total power of incoherent radiation. Using the classical formula for an accelerated electron [43] the average energy loss is

$$\frac{d<\gamma>}{dz} = -\frac{2}{3} r_e \gamma^2 k_U^2 K^2 \quad , \tag{6.18}$$

where $r_e = e^2/4\pi\epsilon_0 mc^2 = 2.818 \cdot 10^{-15}$ m is the classical electron radius, $k_U = 2\pi/\lambda_U$ is the undulator wave number and K is the undulator parameter. The relative change in $(\Delta\gamma/\gamma)$ is linear in γ and might be larger than the energy bandwidth of the FEL amplification. The constraint for the tolerable energy loss $L_u d<\gamma>/dz$ and thus for the maximum undulator length L_U is give, analogous to the emittance (Eq. 6.5), by

$$L_U \ll \frac{3\rho}{2 r_e \gamma k_u^2 K^2} \quad . \tag{6.19}$$

For the design parameters of the X-ray FEL at TESLA at an electron energy of 25 GeV the undulator length would be limited to roughly 50 m. Compared to the design value of $L_U = 100$ m in order to reach saturation the constraint Eq. 6.19 is well exceeded.

The change in the resonance condition can be avoided if the undulator field is matched to the energy loss. The variation of the undulator field is call 'tapering', which is often used to push the saturation power to a higher level. To keep the resonant wavelength constant the undulator field compensates the energy loss with

$$\frac{dK}{dz} = \frac{1+K^2}{\gamma K} \frac{d<\gamma>}{dz} \quad . \tag{6.20}$$

For the TESLA FEL at a beam energy of 25 GeV the relative change in K over the entire undulator is 0.18%.
The taper of the undulator field introduces the problem that the resonant condition is disturbed for lower energies, where the energy loss is overcompensated. Fortunately this problem is less severe because the saturation length is shorter and the energy bandwidth is larger, which makes the FEL amplification overall less sensitive to the variation in the undulator parameter K.

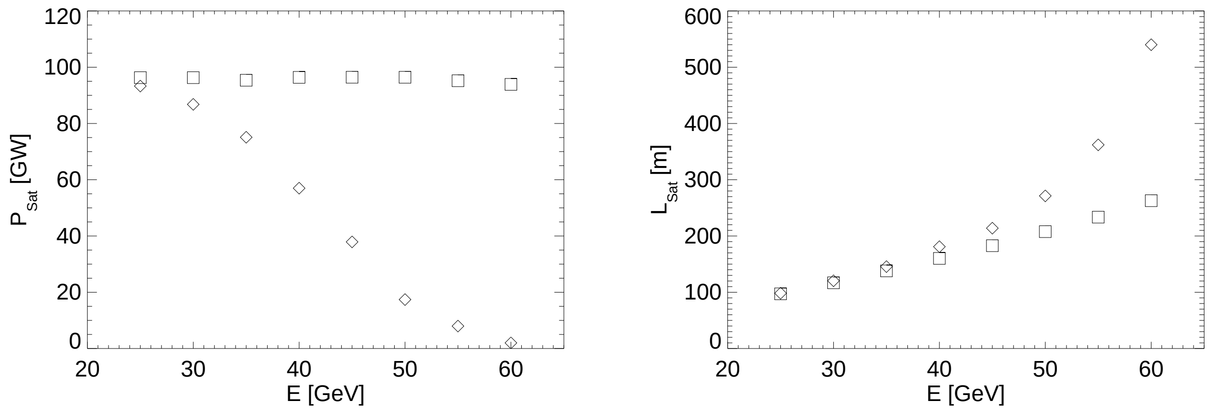


Figure 6.11: *Saturation power and length (left and right, respectively) for different electron beam energies including the increase of the energy spread by incoherent emission of spontaneous undulator radiation ($\diamond$). The case of excluding the emission is indicated by the $\square$-markers.*

While the coherent energy loss can by compensated by a tapered undulator the incoherent emission affects the energy spread of the beam. For higher beam energies the spectrum of the spontaneous undulator radiation covers frequencies where the number of emitted X-ray photons is small. The emission does not follow the classical description anymore and only quantum mechanics can estimate the probability of the emission correctly. This random process yields an increase of the energy spread due to the fluctuation in the emitted power. The growth rate is given by the formula [127]

$$\frac{d < \Delta\gamma^2 >}{dz} = \frac{14}{15} \lambda\!\!\!^{-}_c r_e \gamma^4 k_U^3 K^2 F(K) \quad , \tag{6.21}$$

where $\lambda\!\!\!^{-}_c = \hbar/mc$ is the normalized Compton wavelength, $\hbar$ is the Planck's constant and $F(K)$ is a fit to simplify the dependence on various Bessel functions [50] with

$$F(K) = \begin{cases} 1.42K + \dfrac{1}{1 + 1.50K + 0.95K^2} & \text{for a helical undulator,} \\ \\ 1.70K + \dfrac{1}{1 + 1.88K + 0.80K^2} & \text{for a planar undulator.} \end{cases} \tag{6.22}$$

The dependence of this quantum fluctuation on the energy is rather strong, increasing with the fourth power. Therefore it can be expected that the amplification is significantly reduced

above an energy threshold. Based on the undulator parameter of the TESLA undulator the FEL performance for higher energies are simulated until the degradation by the quantum fluctuation is seen. The results for the saturation power and length are shown in Fig. 6.11. While the case of excluding the increase of the energy spread agrees with the analytic estimation with $L_{sat} \propto 1/\rho \propto \gamma$ and $P_{sat} \approx \rho P_{beam} \approx const$, the saturation power is even reduced for 25 GeV when quantum fluctuation is included. The saturation length is affected for electron beam energies above 35 GeV. Beyond 50 GeV, which corresponds to a radiation wavelength of 0.6 Å, the operation of the FEL becomes highly inefficient.
In theory the simplest method to reach shorter wavelengths is to decrease the gain length in particular providing a higher electron density. If the FEL saturates faster the energy spread is less accumulated. In practice it is difficult to realize because the transverse and longitudinal emittances and therefore energy spread, peak current, beam size and transverse divergence are defined by the electron source. Improving one parameter is accompanied by degrading others such as an increased energy spread for a shorter bunch length or a stronger axial velocity spread for a stronger focusing. The improvement of the electron sources is a major task of ongoing research and development.
A more efficient way is to use the higher harmonics of the FEL radiation. At saturation the longitudinal phase space distribution consists of many higher harmonics. Due to the symmetry in the distribution the even numbered harmonics are often suppressed.
The prebunched beam can be injected into a second undulator being resonant on a harmonic wavelength. Because the initial degree of bunching is already high the length to reach saturation is significantly reduced.

Chapter 7

Conclusion

In the near future many Free-Electron Lasers will become operational or are going to be constructed. These high brilliant radiation sources will cover a wide span of radiation wavelengths from the microwave region up to X-ray radiation. The tunability of the wavelength makes them superior over the conventional lasers. This opens new regions of experimental physics based on these new radiation sources.

For the design and operation of a Free-Electron Laser as well as the analysis of the radiation properties of existing devices a detailed knowledge of the amplification process is necessary. Although analytical results provide fast results they cannot cover the entire complex system of the Free-Electron Lasers. Simulation codes have to fill this gap.

In the frame of this thesis the code GENESIS 1.3 has been developed. Compared to other codes GENESIS 1.3 solves the self-consistent FEL equations with the least restrictions and approximations. It is capable of simulating any arbitrary transverse distribution or motion of the electron beam and radiation field because all differential equations are based on the same three dimensional Cartesian coordinate system. In addition the unique feature of describing any longitudinal variation of the electron beam and radiation field parameters allows to study time-dependent effects such as the degradation of the FEL performance by wake fields. The code has been tested with analytical results, other simulation codes and the experimental results of the UCAL/LANL/RRCKI/SSRL experiment on high gain SASE FEL [16].

Using the capabilities of GENESIS 1.3 simulations have been performed for the VUV-FEL at the TESLA Test Facility [17] as well as the X-ray FEL at the TESLA linear collider [20] to estimate the tolerance in the parameters of the electron beam and the undulator.

Any kind of transverse motion, either caused by off axis injection, quadrupole misalignment or field errors in the undulator, degrades the amplification mainly by the missing overlap between the electron beam and radiation field. If the separation is not larger than the electron beam size the correlated variation in the longitudinal velocity is almost negligible for the TTF-FEL.

Wake fields have a significant impact on the FEL performance for short bunch lengths as is the case for the TTF-FEL and, in particular, the TESLA-FEL. One critical parameter is the surface roughness of the beam pipe, which might be the dominating component of the wake fields. The degradation due to wake fields is about 30% for the TTF-FEL. If no special care is taken to avoid large wake field amplitudes the situation is even more severe for X-ray FELs.

The X-ray FEL at TESLA will operate at a beam energy of up to 50 GeV. Due to the fluctuation of hard X-ray photons in the incoherent spectrum of the undulator radiation the growth of the bean energy spread becomes noticeable and denies any reasonable FEL operation at a beam energy above 50 GeV. this effect was predicted based on an analytical approximation and could be investigated in more detail with the new code.
Another critical effect at this energy is the variation in the transverse velocity. Under certain conditions a large beam emittance can be compensated by a correlation between electron energy and amplitude of the betatron oscillation. One case is a quadrupole lattice with a small phase advance of the betatron phase per FODO cell length as it is planned for the undulator of the TESLA FEL. Due to the electron trajectory in this lattice the variation in the longitudinal velocity is fast oscillating on the scale of the gain length, and the compensation scheme is highly efficient.
Summarizing, it has been shown that the new 3D time-dependent simulation code GENESIS 1.3 is a valuable tool. The code makes it possible to study important effects which were not accessible by numerical simulation before. For this reason GENESIS 1.3 has been written.
The question remains whether GENESIS 1.3 is suitable for all kinds of FELs. For the current version of GENESIS 1.3 the answer is 'No', because the code is focused on the single pass high gain Free-Electron Laser. Nevertheless GENESIS 1.3 might be extended to cover other FEL concepts such as FEL oscillators, two-stage FELs or higher harmonic generation.

Appendix A

The Paraxial Approximation of the Wave Equation

In this appendix the paraxial approximation of the Maxwell's equations for the free space propagation is derived as well as a fundamental set of radiation modes which solve the equations. Due to their transverse dependence, which is dominated by a Gaussian function, these types of beams are called 'Gaussian beams'. Laser radiation trapped in an optical cavity as well as beam transport are a typical application regions of these modes.
For the amplification process the fundamental modes of a FEL might differ from those of free space due to the source term of the electron beam. Nevertheless most of the results presented in this appendix can be applied to the physics of an FEL.

A.1 The Scalar Paraxial Wave Equation in Free Space

To find the solution of the Maxwell's equations in free space it is a common practice to transform them into the frequency domain. The resulting wave equation is the Helmholtz equation. Using the dispersion relation of free space $k = \omega/c$, where k is the wavenumber and ω is the frequency, the scalar Helmholtz equation is

$$\left[\vec{\nabla}^2 + k^2\right] \hat{a}(x, y, z, k) = 0 \quad , \tag{A.1}$$

with the Laplace operator $\vec{\nabla}^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$. For a complete representation of the Maxwell's equations in the frequency domain the Helmholtz equation has four independent components: three for the vector potential $\vec{A}$ and one for the scalar potential ϕ.
Without solving the wave equations, it is known in advance that the propagation for many problems shows a dominant direction. This includes the propagation of the light beam with small divergence produced by conventional lasers as well as radiation of Free-Electron Lasers. The coordinate system is chosen in the way that the z-axis points along the main direction. Due to $k_z \gg k_x$, k_y the leading term of a solution of the Helmholtz equation, which is of interest, is $\hat{a} \propto e^{-ik_z z} \approx e^{-ikz}$. Decomposing into factors the field by $\hat{a} = ae^{-ikz}$ Eq. A.1 becomes the paraxial equation [42]

$$\left[\nabla_\perp^2 - 2ik\frac{\partial}{\partial z}\right] a = 0 \quad , \tag{A.2}$$

where $\vec{\nabla}_\perp^2 = \partial^2/\partial x^2 + \partial^2/\partial y^2$ is the Laplace operator restricted to the transverse (x, y)-plane because the second order derivative $\partial^2 a/\partial z^2$ is much smaller than all other remaining terms. For an estimation of the application region of the paraxial equation the exact solution of a plane wave propagating under an angle φ with respect to the z-axis is inserted into Eq. A.2. The slow varying amplitude is $a = a_0 \exp[-ik\sin\varphi x + ik(1 - cos\varphi)z]$. Inserting a into the differential equation and expanding the trigonometric function into Taylor series, the transverse Laplace operator yields the term $-k^2\varphi^2$ and the first order derivative with respect to z the term $k^2\varphi^2$. The omitted second order derivative has an value of $-k^2\varphi^4/4$. As long as the constrain $\varphi < 1/2$ is valid the resulting term is smaller by at least one order of magnitude. For all problems discussed in this thesis the angle is much small and the paraxial approximation is well justified. If only propagation in free space is taken into account all solutions of the Maxwell's equations can be expressed by a superpositions of plane waves. The time dependence is pure harmonic ($\propto e^{i\omega t}$) and the scalar Helmholtz equation is identical to the Maxwell's equation in the time domain.
With

$$\vec{E} = -\frac{\partial}{\partial t}\vec{A}$$

and

$$\vec{E} = \vec{\mathcal{E}} e^{i\vec{k}\vec{r} - i\omega t} \approx \vec{\mathcal{E}} e^{ikz - i\omega t}$$

any component of $\vec{\mathcal{E}} = (\mathcal{E}_x, \mathcal{E}_y, \mathcal{E}_z)$ must satisfy Eq. A.2. In the paraxial approximation the longitudinal component of the electric field $\mathcal{E}_z$ is negligible compared to the transverse field components due to $\vec{k} \cdot \vec{\mathcal{E}} = 0$. For convenience it is set to zero and the radiation has only transverse electric and magnetic field components ($\vec{B} \propto \vec{e}_z \times \vec{E}$).
If the two transverse components of the electric field have the same solution of the paraxial equation the polarization is conserved. Therefore it is sufficient to find only the solution a of the scalar equation (Eq. A.2), where the polarization is given by the initial conditions. In the following section a circular polarization is assumed ($a = \mathcal{E}_x + i\mathcal{E}_y$).
In general the polarization does not remain constant if the general solution of the paraxial equation is a superposition of different, independent solutions with different polarization.

A.2 Properties and Solutions of the Paraxial Equation

The paraxial equation

$$\left[\vec{\nabla}_\perp^2 - 2ik\frac{\partial}{\partial z}\right] a = 0 \quad , \tag{A.3}$$

is very similar to the time dependent Schrödinger equation of a free particle [128]

$$i\hbar\frac{\partial}{\partial t}\Psi = -\frac{\hbar^2}{2m}\vec{\nabla}^2\Psi \quad ,$$

where $\hbar = h/2\pi$, h is the Planck's constant, and m the mass of the particle.
Indeed most of the properties of the quantum mechanics are valid for the paraxial equation. Using the L_2-norm [70] for the radiation field

$$||a||_2 \equiv \int |a|^2 dxdy \quad ,$$

the norm of any solution should be finite. In particular the L_2-norm is proportional to the radiation power, because the absolute square of the radiation field amplitude is related to the intensity I by

$$|a|^2 = a^* a = |\mathcal{E}_x|^2 + |\mathcal{E}_y|^2 = Z_0 I \quad , \tag{A.4}$$

where $Z_0 = 377\Omega$ is the vacuum impedance and a^* is the complex conjugated of the radiation field a. The integration over the transverse plane yields the radiation power. For a finite power the transverse extension of the solution a of the paraxial equation must be limited in the sense that the field amplitude converges faster towards zero as $1/r$ in the limit of an infinite transverse radius $r \to \infty$. These solutions are said to be 'quadratic integrable'.
Again using the method of quantum mechanics it can be shown that the power (L_2-norm) is conserved:

$$\frac{\partial}{\partial z}P = \frac{1}{Z_0}\int\left[a^*\frac{\partial}{\partial z}a + (\frac{\partial}{\partial z}a^*)a\right]dxdy = \frac{1}{Z_0}\int\left[a^*\vec{\nabla}_\perp a - (\vec{\nabla}_\perp a^*)a\right]d\vec{s} = 0 \quad ,$$

by using Greens identity, where the area integration is converted to an integration along the boundary $\int d\vec{s}$ of that area. If the integration area is extended to the whole transverse plane the integral becomes zero because the gradient $\vec{\nabla}_\perp a$ converges even faster towards zero than a itself for $r \to \infty$.
To find the general solution of the paraxial equation the first step is to expand $a(x, y, z)$ into a series of functions $\{u_n(x, z)\}$,

$$a(x, y, z) = \sum_n c_n(y, z)u_n(x, z) \quad , \tag{A.5}$$

where the coefficients c_n still depend on y and z. This expansion exists if the set of functions $\{u_n\}$ fulfills the constraint of being an orthonormal set of functions

$$\int_{-\infty}^{\infty} u_n^*(x, z)u_m(x, z)dx = \delta_{n,m} \quad ,$$

which covers all points of x (complete relation)

$$\sum_n u_n^*(x, z)u_n(x', z) = \delta(x - x') \quad .$$

One possible set of functions are the solutions of the harmonic oscillator $h_n(x)$ in quantum mechanics, which are orthonormal and complete. These Hermite-Gaussian polynomials are defined as

$$h_n(x) = \left(\sqrt{\pi}2^n n!\right)^{-\frac{1}{2}} e^{-\frac{x^2}{2}} H_n(x) \quad , \tag{A.6}$$

with the Hermite polynomials

$$H_n(x) = (-1)^n e^{x^2} \left(\frac{d^n}{dx^n} e^{-x^2}\right) \quad . \tag{A.7}$$

To include a z-dependence the argument x can be scaled by a function $f(z)$ without loss of completeness. A different function $g(z)$ has to be multiply to the Hermite-Gaussian polynomials $h_n(f(z)x)$ to keep the norm independent of z. A phase shift $\exp(it(z)x^2/2)$ is added to allow transverse variation of the radiation phase. The so far arbitrary function values of $f(z)$ and $t(z)$ must be real values while this is not necessarily valid for $g(z)$. A complex value of $g(z)$ means a global phase shift of the radiation field depending on z. The most general approach for $u_n(x,z)$ is

$$u_n(x,z) = g(z) H_n(f(z)x) e^{i\frac{t(z)}{2}x^2} \quad . \tag{A.8}$$

By using properties of the Hermite-Gaussian polynomials the insertion of Eq. A.5 and Eq. A.8 into the paraxial equation Eq. A.3 yields

$$\sum_n \Bigg[\left[\left(\frac{\partial^2}{\partial y^2} - 2ik\frac{\partial}{\partial z}\right) c_n(y,z)\right] u_n(x,z) + \ldots$$
$$c_n(y,z) u_n(x,z) \left[\hat{F}_1(z,n) + \hat{F}_2(z) f(z) x \frac{H_n'(f(z)x)}{H_n(f(z)x)} + \hat{F}_3(z) f^2(z) x^2\right]\Bigg] = 0 \quad , \tag{A.9}$$

where the functions $\hat{F}_1$, $\hat{F}_2$ and $\hat{F}_3$ depend on f, g, t and their derivation with respect to z. Due to the properties of the Hermite-Gaussian polynomials [128] all these three terms are independent. Therefore $\hat{F}_1$, $\hat{F}_2$ and $\hat{F}_3$ must be zero resulting in the differential equations of the unknown functions f, g and t.

$$\hat{F}_1(z,n) = it(z) - 2ik\frac{g'(z)}{g(z)} - f^2(z)(2n+1) = 0 \tag{A.10}$$

$$\frac{1}{2i}\hat{F}_2(z) = t(z) - k\frac{f'(z)}{f(z)} = 0 \tag{A.11}$$

$$\hat{F}_3(z) = k\frac{t'(z)}{f^2(z)} + f^2(z) - \frac{t^2(z)}{f^2(z)} = 0 \tag{A.12}$$

Inserting $t(z) = kf'(z)/f(z)$ and substituting $\tilde{f}(z) = 1/f(z)$ Eq. A.12 has the general solution

$$\tilde{f}(z) = \sqrt{\pm az + b} \quad ,$$

with $a = 2i/k$ and any arbitrary, complex value for b. It is convenient for further calculation to write the resulting solution of $f(z)$ as

$$f(z) = \sqrt{\frac{ik}{2}\left[\frac{1}{q(z)} - \frac{1}{q^*(z)}\right]} \quad , \tag{A.13}$$

where $q(z) = z + q_0$ is linear in z. The complex constant q_0 depends on the initial conditions. The imaginary part of q_0 must be positive to get only real values of $f(z)$ for any z.
The function $t(z)$, directly derived from Eq. A.11, is

$$t(z) = -\frac{k}{2}\left[\frac{1}{q(z)} + \frac{1}{q^*(z)}\right] \quad . \tag{A.14}$$

Eq. A.10 is the only differential equation, which depends on the mode-number n and contains complex numbers. The complex solution causes an additional phase shift of the radiation field while propagating along z.
Insertion of Eqs. A.13 and A.14 into Eq. A.10 yields

$$g(z) = \left(\frac{q^*(z)}{q(z)}\right)^{\frac{n}{2}} \frac{1}{\sqrt{q(z)}} \quad . \tag{A.15}$$

Gathering all solutions the set of basic functions to expand the radiation field in the x-direction is found. After some simple algebra and including initial conditions Eq. A.8 becomes

$$u_n(x,z) = \left(\frac{f_0}{\sqrt{\pi}2^n n!}\right)^{\frac{1}{2}} \left(\frac{q_0}{q(z)}\right)^{\frac{1}{2}} \left(\frac{q_0}{q_0^*}\frac{q^*(z)}{q(z)}\right)^{\frac{n}{2}} H_n(f(z)x)\exp\left[-i\frac{kx^2}{2q(z)}\right] \quad , \tag{A.16}$$

with $f_0 = f(0)$.
For a complete solution of Eq. A.9 the coefficient $c(y,z)$ can be expanded into a series of orthonormal functions $\tilde{u}_m(y,z)$. These functions have to satisfy the same 1 dimensional paraxial differential equation as $u_n(x,z)$ and the expansion into Hermite-Gaussian modes seems to be reasonable. The two sets of orthonormal functions $\{u_n(x,z)\}$ and $\{\tilde{u}_n(y,z)\}$ might differ in their initial conditions q_0 and $\tilde{q}_0$, respectively. This will result in astigmatic modes which show different diffraction in x- and y-direction (see next section). However the fact that the set of functions $\{\tilde{u}_m(y,z)\}$ is always complete and independent of the initial conditions an arbitrary function $u(y,z)$ can be described as a superposition of $\tilde{u}_m(y,z)$ as well as a superposition of $\hat{u}_m(y,z)$, where the set of functions $\{\hat{u}(y,z)\}$ is based on the different initial condition $\hat{q}_0$. Therefore an unitary transformation $B = (b_{ij})$ must exist with

$$\hat{u}_i(y,z) = \sum_j b_{ij}\tilde{u}_j(y,z) \quad .$$

By choosing $\hat{q}_0 = q_0$ the set of functions to expand the radiation field in the x- and y-direction is the same.
The final solution of the paraxial equation becomes

$$u(x,y,z) = \sum_{n,m} \hat{c}_{nm} u_n(x,z) u_m(y,z) \quad . \tag{A.17}$$

The coefficients $\hat{c}_{nm}$ are independent of any variable. The calculation of the total radiation power yields

$$P = \frac{1}{Z_0} \sum_{n,m} |\hat{c}_{nm}|^2 \quad .$$

It has to be mentioned that several other sets of orthonormal function exists. Another common approach is to expand the field distribution in a dependence on the azimuthal angle ϕ and the radius ρ. This approach will give the Laguerre-Gaussian modes [129].
The advantage of using Hermite-Gaussian or Laguerre-Gaussian modes is that the shape of a single mode is conserved and can directly be calculated for any value of z (see next section). This is not the case for expansion into sine and cosine functions.

A.3 Hermite-Gaussian Modes

For the fundamental mode $(m, n = 0)$ the radiation field follows a Gaussian distribution although the argument of the exponential function might be complex

$$u(x,y,z) = \frac{f_0 q_0}{\sqrt{\pi} q(z)} \exp\left(-i \frac{k}{2q(z)} (x^2 + y^2) \right) \quad . \tag{A.18}$$

The root-mean-square radius $w(z)$ for the fundamental mode is calculated by

$$w^2(z) = 2 \int u^*(x,y,z)(x^2+y^2) u(x,y,z) dx dy \quad .$$

The factor '2' is added to compensate the effect of the absolute square of the Gaussian distribution. Inserting Eq. A.18 shows that the real part of $q^{-1}(z)$ drops out and that the beam size is given by the identity

$$\frac{1}{w(z)^2} = -\frac{k}{2} \Im\mathrm{m} \left[\frac{1}{q(z)} \right] \quad ,$$

With the linear, complex function $q(z) = z - q_0 = z - z_0 + i z_R$ the variance $w^2(z)$ becomes

$$w^2(z) = \frac{2 z_R}{k} \left(1 + \frac{(z - z_0)^2}{z_R^2} \right) \quad . \tag{A.19}$$

While the radiation field of the fundamental mode propagates along z-axis it has its minimum size at $z = z_0$. The root-mean-square size of the waist at that position is $w_0 = \sqrt{2 z_R / k}$. The parameter z_R defines the Rayleigh length and it is a measure for the diffraction of the radiation field. A large value of z_R means less diffraction. In the limit $z \gg z_R + z_0$ the radiation size

grows nearly linear with z. In this limit the diffraction angle is given by $\tan\theta = \sqrt{2/kz_R}$. Note that the factor f_0 in Eq. A.18 is equivalent to $\sqrt{2}/w(0)$.
The Hermite-Gaussian modes of the radiation field are closely related to the solution of the 2 dimensional harmonic oscillator in quantum mechanics. As long as only one single radiation mode is regarded the calculation of the variance $< x^2 + y^2 >$ is equivalent to the expectation value of the energy for eigenfunctions of the harmonic oscillator [128]. For a single (n,m)-Hermite Gaussian mode it is

$$\begin{aligned} < x^2 + y^2 >_{nm} &= 2\int u_n^*(x,z)u_m^*(y,z)(x^2+y^2)u_n(x,z)u_m(y,z)dxdy \\ &= (n+m+1)w^2(z) \quad , \end{aligned} \tag{A.20}$$

where $w^2(z) =< x^2 + y^2 >_{00}$ is the variance of the fundamental mode. The beam size as well as the diffraction angle scales with $\sqrt{n+m+1}$ for higher modes.
While the imaginary part of $q^{-1}(z)$ yields a finite transverse beam size the real part gives a transverse dependence of the radiation phase. It is fruitful to compare the fundamental Hermite Gaussian mode to the paraxial approximation of a spherical wave. The source of the wave is at $\vec{r}_0 = (x_0, y_0, z_0)$. For large $z \gg z_0$ the spherical wave becomes

$$\frac{\exp[-ik|\vec{r}-\vec{r}_0|]}{|(\vec{r}-\vec{r}_0)|} \approx \frac{e^{-ik(z-z_0)}}{z-z_0}\exp\left[-ik\frac{(x-x_0)^2+(y-y_0)^2}{2(z-z_0)}\right] \quad . \tag{A.21}$$

It is identical to the fundamental Hermite-Gaussian mode in the limit $z/z_R \to \infty$. In a more general way the term $R(z) = 1/\Re\mathrm{e}[q^{-1}(z)]$ can be regarded as the radius of the wavefront curvature. For large values of z it converges towards the linear dependence of a spherical wave ($R(z) \approx z - z_0$). At the waist ($z \to z_0$) $R(z)$ has a singularity and the phase front is uniform as a plane wave. The minimum radius of curvature for the fundamental Hermite-Gaussian mode is $2z_R$ at $z = z_R$.
Beside this transverse variation of the phase an axial phase shift occurs, called Guoy phase shift when the radiation goes through its waist. It is caused by the complex factor $f_0q_0/q(z)$ in Eq. A.18.
To estimate this phase shift it is convenient to split $q^{-1}(z)$ into its real and imaginary part

$$\frac{i}{q(z)} = \frac{2}{kw^2(z)} + i\frac{1}{R(z)} = \frac{e^{i\Psi(z)}}{|q(z)|} \quad , \tag{A.22}$$

defining the Guoy phase $\Psi(z)$

$$\tan\Psi(z) = \frac{kw^2(z)}{2R(z)} = \frac{z-z_0}{z_R} \quad . \tag{A.23}$$

After some algebra and using the identity $q(z) - q^*(z) = q_0 - q_0^*$ the factor $f_0q_0/q(z)$ can be rewritten as

$$\frac{f_0q_0}{q(z)} = \frac{\exp[i(\Psi(z)-\Psi_0)]}{w(z)} \quad . \tag{A.24}$$

It can be seen that in contrast to a plane wave the fundamental Hermite-Gaussian modes see an extra phase shift of $\Delta\Psi = \pi$ while going through its waist. This Guoy phase shift is even enhanced for higher modes. The extra term for higher modes is given by

$$\left[\frac{q_0}{q_0^*}\frac{q^*(z)}{q(z)}\right]^{\frac{n}{2}} = \exp[in(\Psi(z) - \Psi_0)] \quad .$$

In conclusion of this section the total phase shift for a (n, m)-Hermite Gaussian mode with respect to a plane wave is $\Delta\Psi = (n + m + 1)\pi$.

A.4 The Far Field

The radiation field propagation for rather long distances can be done in two different ways. Either the radiation field $u(\vec{r}_0)$ at position $\vec{r}_0 = (x_0, y_0, z_0)$ is expanded into a sum of orthonormal functions u_{nm} as shown in the previous section. The z-dependence is well known and the field distribution can be calculated at any other position.
Or the field is directly transformed by solving Huygens integral. This integration is based on Huygens principle, where each point of a given wavefront $u(\vec{r}_0)$ is regarded as a source of a spherical wave. To evaluate the radiation field at $\vec{r} = (x, y, z)$ the convolution of the radiation field $u(\vec{r}_0)$ and a spherical wave $\exp(-ikr)/r$ has to be calculated. In the paraxial approximation this convolution is

$$u(x, y, z) = \frac{e^{-ik(z-z_0)}}{z - z_0} \int u(x_0, y_0, z_0) \exp\left[-ik\frac{(x - x_0)^2 + (y - y_0)^2}{2(z - z_0)}\right] dx_0 dy_0 \quad . \tag{A.25}$$

or after some rearrangement

$$u(x, y, z) = \frac{\exp\left[-ikL - i\frac{k}{2L}(x^2 + y^2)\right]}{L} \int \tilde{u}(x_0, y_0, z_0) \exp\left[i\frac{k}{L}(x_0 x + y_0 y)\right] dx_0 dy_0 \quad , \tag{A.26}$$

with $L = z - z_0$ and the modified radiation field

$$\tilde{u}(x_0, y_0, z_0) = u(x_0, y_0, z_0) e^{-i\frac{2k}{L}(x_0^2 + y_0^2)} \quad . \tag{A.27}$$

In the far field limit $L \gg A_0 k$, where A_0 is the spot size of the radiation field at z_0, the phase factor in Eq. A.27 is close to unity over the whole beam size and can be neglected. The Huygens integral (Eq. A.26)is reduced to the transverse Fourier transformation of the near field multiplied by a phase factor depending on the observation angle. It is worth to mention that the root-mean-square diffraction angle $\theta_{rms} = \sqrt{< x^2 + y^2 >}/L$ for Hermite-Gaussian modes in the far field limit agrees with the diffraction angle derived in the previous section.

Appendix B

The Initial Value Problem of the 3D FEL Model

This appendix treats the special problem of the 3D model to find the radiation field evolution depending on the seeding field $A_0(x, y)$ of the FEL amplifier. The method to solve this problem is closely related to the eigenvalue problem in Section 2.6. It differs in the use of Laplace transformation techniques [33], which includes automatically the initial radiation field A_0 as a source term of the differential equation to be solved.
The motion of an electron within an undulator and radiation field is described by the equations

$$\dot{\theta}_j = ck_U - \omega\frac{1 + p_\perp^2 + K^2}{2\gamma_j^2} \quad , \tag{B.1}$$

and

$$\dot{\gamma}_j = -\omega\frac{f_c K}{2\gamma_j}(ue^{i\theta_j} + \text{c.c.}) - i\frac{e^2\mu_0}{km}\left[\sum_l \delta(\vec{r} - \vec{r}_l)e^{-i\theta_l} - c.c.\right] \quad , \tag{B.2}$$

where $\theta = (k + k_U)z - kct$ is the electron phase of the ponderomotive wave, k is the radiation wavenumber, k_U is the undulator wavenumber, γ is the Lorentz factor of the electron energy, $u = -i(e\hat{E}_0/mc^2k)\exp(i\Psi)$ is the slow varying radiation field amplitude, $\hat{E}_0$ is the root-mean-square electric field at the undulator axis, Ψ is the radiation phase, $K = e\hat{B}_0/mck_U$ is the undulator parameter, $\hat{B}$ is the root-mean-square undulator field at the undulator axis, f_c is the coupling factor, e is the electron charge, m is the electron mass and c is the speed of light. The index j denotes the jth electron and '$c.c.$' indicates the complex conjugated of the previous term.
Defining the 3D FEL parameter

$$\hat{\rho} = \left[\frac{I}{\gamma_0 I_A}\frac{K^2}{1 + K^2}\right]^{1/2} \quad , \tag{B.3}$$

where γ_0 is the mean energy, I is the peak current of the beam and $I_A \approx 17$ kA is the Alfven current, the differential equations Eqs. B.1 and B.2 are normalized to obtain the quantities $\Phi \equiv \theta$, $\eta \equiv (\gamma - \gamma_0)/\hat{\rho}\gamma_0$, $A \equiv f_c K k u / 4\gamma_0^2 k_U \hat{\rho}^2$ and the normalized position $\hat{z} = 2k_U\hat{\rho}ct$.

Assuming a small energy spread the differential equations are linearized in η and A yielding

$$\begin{aligned}\Phi_j' &= \eta_j + \delta \quad , &\text{(B.4)}\\ \eta_j' &= -\left[\left(A + i\hat{\sigma}^2 \left\langle e^{-i\Phi_j}\right\rangle\right) e^{i\Phi_j} + c.c\right] \quad . &\text{(B.5)}\end{aligned}$$

Two new parameters have been introduced to reduce the number of constants in the differential equation. The detuning parameter $\delta = (\gamma_0^2 - \gamma_R^2)/2\hat{\rho}\gamma_R^2$ describes the average phase slippage of the electron beam relative to the radiation field if the electron beam energy differs from the resonant energy $\gamma_R = \sqrt{k(1+K^2)/2k_U}$.
The other parameter arises due to the longitudinal space charge field inhibiting a periodic bunching of the electrons in θ. It is defined by $\hat{\sigma}^2 = 2\hat{\rho}(1+K^2)/r_0^2 k k_U K^2$, where r_0 is the electron beam radius.
The problem to solve Eqs. B.4 and B.5 is transformed into a problem to find a longitudinal phase space distribution f fulfilling the equation (see Section 2.6)

$$\left[\frac{\partial}{\partial t} + \Phi' \frac{\partial}{\partial \Phi} + \eta' \frac{\partial}{\partial \eta}\right] f = 0 \quad . \tag{B.6}$$

In the linear regime of the FEL the field distribution can be restricted to the fundamental and first harmonic of the Fourier series expansion in Φ with $f = f_0 + f_1 e^{i\Phi}$.
To simplify the model an axi-symmetric electron beam with fixed transverse electron positions is assumed. Inserting Eqs. B.4 and B.5 into Eq. B.6 and collecting all terms resonant in $\exp[i\Phi]$ the formal solution is given by

$$f_1 = \int_0^{\hat{z}} d\hat{z}' \left[A + i\hat{\sigma}^2 \left\langle e^{i\Phi}\right\rangle\right] \frac{\partial f_0}{\partial \eta} e^{i(\eta+\delta)(\hat{z}'-\hat{z})} \quad . \tag{B.7}$$

The wave equation in the paraxial approximation of the normalized radiation field is

$$\left[\hat{\nabla}_\perp^2 + 2iB\frac{\partial}{\partial \hat{z}}\right] A = 2i \int f(\eta, \phi, \hat{r}) e^{-i\Phi} d\Phi d\eta \quad , \tag{B.8}$$

defining the diffraction parameter $B = 2r_0^2 k k_U \hat{\rho}$ and normalized transverse position $\hat{r} = r/r_0$, where r_0 is radius of the electron beam. The integration of the source term over Φ is non-zero only for the first harmonic f_1 of the phase space distribution.
The remaining integration is identical with the bunching factor and allows to replace it in Eq. B.7 by the left hand side of Eq. B.8 divided by $2i$. Inserting f_1, including the substituted bunching factor, into Eq. B.8 results in the integro-differential equation for the radiation field

$$\left[\hat{\nabla}_\perp^2 + 2iB\frac{\partial}{\partial \hat{z}}\right] A = \int_0^{\hat{z}} d\hat{z}' \left(2A + \hat{\sigma}^2 \left[\hat{\nabla}_\perp^2 + 2iB\frac{\partial}{\partial \hat{z}'}\right] A\right) \int i d\eta \frac{\partial f_0}{\partial \eta} e^{i(\eta+\delta)(\hat{z}'-\hat{z})} \quad . \tag{B.9}$$

Eq. B.9 can be integrated over $\hat{z}'$ if a Laplace transformation

$$\tilde{A}(\hat{p}, \hat{r}) = \int_0^\infty A(\hat{z}, \hat{r}) e^{-\hat{p}\hat{z}} d\hat{z}$$

is applied to it. The real part or $\hat{p}$ must be positive to guarantee that the integration of the Laplace transformation converges. Some lengthy but straight forward calculations give the transformed wave equation

$$\left[\hat{\nabla}_\perp^2 - \frac{2D}{1-\hat{\sigma}^2 D} + 2iB\hat{p}\right]\tilde{A}(\hat{p},\hat{r}) = 2iBA_0(\hat{r}) \tag{B.10}$$

with

$$D = i\int \frac{\partial f_0}{\partial \eta}\frac{d\eta}{\hat{p}+i\eta+i\delta} \quad . \tag{B.11}$$

The initial radiation $A_0(\hat{r})$ at the entrance of the undulator is the source term of the wave equations. If $\hat{p}$ is substituted with $i\Lambda$ the homogeneous differential equation is identical with the eigenvalue equation Eq. 2.109 in Section 2.6. Despite the difference in the source term the approach to solve Eq. B.11 is the same by decomposing $\tilde{A}(\hat{p},\hat{r})$ and $A_0(\hat{r})$ into a Fourier series of the azimuthal angle ϕ.

The general solution is the sum of all solutions of the homogeneous differential equation (see Section 2.6) and one solution of the inhomogeneous one. Assuming an axi-symmetric electron beam with a stepped radial profile and using the notation

$$\mu^2 = \frac{-2D}{1-\hat{\sigma}^2 D} - g^2 \quad , \tag{B.12}$$

$$g^2 = -2iB\hat{p} \quad , \tag{B.13}$$

$$s_m(\hat{r}) = 2iB\int_0^{2\pi} A_0(\hat{r})e^{-im\phi}d\phi \tag{B.14}$$

the radiation field is given for $\hat{r} < 1$ by

$$A = c_1 J_m(\mu\hat{r}) + \frac{\pi}{2}Y_m(\mu\hat{r})\int_0^{\hat{r}} d\hat{r}'\hat{r}'J_m(\mu\hat{r}')s_m(\hat{r}') + \frac{\pi}{2}J_m(\mu\hat{r})\int_{\hat{r}}^1 d\hat{r}'\hat{r}'Y_m(\mu\hat{r}')s_m(\hat{r}') \tag{B.15}$$

and for $\hat{r} > 1$ by

$$A = c_2 K_m(g\hat{r}) - K_m(g\hat{r})\int_1^{\hat{r}} d\hat{r}'\hat{r}'I_m(g\hat{r}')s_m(\hat{r}') - I_m(g\hat{r})\int_{\hat{r}}^1 d\hat{r}'\hat{r}'K_m(g\hat{r}')s_m(\hat{r}') \quad . \tag{B.16}$$

The continuity conditions at $\hat{r} = 1$ give two identities to evaluate the amplitudes c_1 and c_2 by

$$\begin{aligned} c_1 &= [\mu J_{m+1}(\mu)K_m(g) - gK_{m+1}(g)J_m(\mu)]^{-1}\left(\int_1^\infty d\hat{r}'\hat{r}'K_m(g\hat{r}')s_m(\hat{r}')\right. \\ &+ \left.\frac{\pi}{2}[gK_{m+1}(g)Y_m(\mu) - \mu K_m(g)Y_{m+1}(\mu)]\int_0^1 d\hat{r}'\hat{r}'J_m(\mu\hat{r}')s_m(\hat{r}')\right) \end{aligned} \tag{B.17}$$

$$\begin{aligned} c_2 &= [\mu J_{m+1}(\mu)K_m(g) - gK_{m+1}(g)J_m(\mu)]^{-1}\left(\int_0^1 d\hat{r}'\hat{r}'J_n(\mu\hat{r}')s_m(\hat{r}'\right. \\ &+ \left.[gJ_m(\mu)I_{m+1}(g) + \mu J_{m+1}(\mu)I_m(g)]\int_1^\infty d\hat{r}'\hat{r}'K_m(g\hat{r}')s_m(\hat{r}')\right) \quad . \end{aligned} \tag{B.18}$$

The two amplitudes c_1 and c_2 are functions of $\hat{p}$ due to the dependence on g and μ.
For values of $\hat{p}$, which solve the dispersion relation of the eigenvalue problem (Section 2.6)

$$\mu J_{m+1}(\mu) K_m(g) = g K_{m+1}(g) J_m(\mu) \quad , \tag{B.19}$$

these amplitudes have singularities. They are important for the inverse Laplace transformation

$$A(\hat{z}, \hat{r}) = \frac{1}{2\pi i} \int_{\hat{p}_0 - i\infty}^{\hat{p}_0 + i\infty} A(\hat{p}, \hat{z}) e^{\hat{p}\hat{z}} d\hat{p} \quad ,$$

where the integration path is closed in the left side of the complex plane. The value of $\hat{p}_0$ is chosen in such a way that all singularities are enclosed by the path of integration. Evaluating the integration by the Cauchy-Riemann theorem [130] the radiation field for the m azimuthal mode is

$$A_m(\hat{z}, \hat{r}) = \sum_n \begin{pmatrix} c_1(\hat{p}_n) J_m(\mu_n \hat{r}) \\ c_2(\hat{p}_n) K_m(g_n \hat{r}) \end{pmatrix} e^{\hat{p}_n \hat{z}} \quad \text{for} \quad \begin{cases} \hat{r} < 1 \\ \hat{r} > 1 \end{cases} \quad , \tag{B.20}$$

where the sum includes all eigenvalues $\hat{p}_n$ of the dispersion relation.
By using properties of the Bessel functions [58], the amplitudes c_1 and c_2, corresponding to the eigenvalue of the nth eigenfunction, can be simplified to

$$\begin{aligned} c_1 &= \left[\frac{K_m(g_n)}{J_m(\mu_n)} \int_0^1 d\hat{r}' \hat{r}' J_m(\mu_n \hat{r}') s_m(\hat{r}') + \int_1^\infty d\hat{r}' \hat{r}' K_m(g_n \hat{r}') s_m(\hat{r}') \right] \\ &\times \left(\frac{d}{d\hat{p}} [\mu J_{m+1}(\mu) K_m(g) - g J_m(\mu) K_{m+1}(g)] \Big|_{\hat{p} = \hat{p}_n} \right)^{-1} \end{aligned} \tag{B.21}$$

$$c_2 = \frac{J_m(\mu_n)}{K_m(g_n)} c_1 \quad . \tag{B.22}$$

According to Eq. B.21 the initial amplitude of an exponentially growing radiation mode is proportional to the overlap integral of the corresponding eigenfunction and the initial radiation field [62]. If the eigenfunctions are known then the initial problem can be solved by integrating Eq. B.21.

Bibliography

[1] J.M.J. Madey, J. App. Phys. **42** (1971) 1906

[2] W.B. Colson, Phys. Lett. **59A** (1976) 187

[3] L.R. Elias *et al.*, Phys. Rev. Lett. **36** (1976) 717

[4] J.S. Frases, R.L. Sheffield, E.R. Gray, Nucl. Instr. & Meth. **A250** (1986) 71

[5] A.M. Kondratenko and E.L. Saldin, Par. Acc.**10** (1980) 207

[6] R. Bonifacio, C. Pellegrini and L.M. Narducci, Opt. Comm. **50** (1984) 373

[7] T.J. Orzechowski *et al.*, Phys. Rev. Lett. **54** (1985) 889

[8] A.L. Throop *et al.*, Nucl. Inst. & Meth. **A272** (1988) 15

[9] D. Kirkpatrick *et al.*, Nucl. Inst. & Meth. **A285** (1989) 43

[10] S. Okuda *et al.*, Nucl. Inst. & Meth. **A331** (1993) 76

[11] D. Bocek *et al.*, Nucl. Inst. & Meth. **A375** (1996) 13

[12] R. Prazeres *et al.*, Phys. Rev. Lett. **78** (1997) 2124

[13] M. Hogan *et al.*, Phys. Rev. Lett. **80** (1998) 289

[14] D.C. Nguyen *et al.*, Phys. Rev. Lett. **81** (1998) 810

[15] M. Babzien *et al.* Phys. Rev. **E57** (1998) 6093

[16] M. Hogan *et al.*, Phys. Rev. Lett. **81** (1998) 4867

[17] J. Rossbach *et al.*, Nucl. Inst. & Meth. **A375** (1996) 269

[18] S.V. Milton *et al.*, Nucl. Inst. & Meth. **A407** (1998) 210

[19] R. Tatchyn *et al.*, Nucl. Inst. & Meth. **A375** (1996) 274

[20] Linear Collider Conceptual Design Report, DESY print 97-48, Hamburg, DESY, 1997

[21] J.M.J. Madey, Il Nuovo Cimen. **50B** (1979) 64

[22] T.M. Tran and J.S. Wurtele, Comp. Phys. Comm. **54** (1989) 263

[23] R.A. Jong, W.M. Fawley and E.T. Scharlemann, SPIE 1045 (1989) 18

[24] S.G. Biedron *et al.*, "The APS SASE FEL: Modeling and Code Comparison", Proc. of the PAC Conf., New York, 1999

[25] K. Halbach, J. Phys. (Paris) **C1 44** (1983) 211

[26] E.T. Scharlemann, J. Appl. Phys. **58** (1985) 2154

[27] H. Goldstein, Classical Mechanics, (Addison-Wesley, Reading, MA, 1980)

[28] D.A. Edwards and M.J. Syphers, An Introduction to the Physics of High Energy Accelerators, (John Wiley and Sons, New York, 1993)

[29] B. Diviacco and P. Walker, Nucl. Inst. & Meth. **A292** (1990) 517

[30] K. Wille, Physik der Teilchenbeschleuniger und Synchrotronstrahlungsquellen, (B.G. Teubner Verlag, Stuttgart, 1992)

[31] G. Rowlands, J. Phys. A: Math. Gen. **13** (1980) 2839

[32] J.P. Blewett and R. Chasman, J. Appl. Phys. **48** (1977) 2692

[33] M. Abramowitz and I.A. Stegun, Handbook of Mathematical Function, (Dover Publication, New York, 1972)

[34] D. Noelle, Ph.D. Thesis, DELTA Int. Rep. 91-31

[35] M.J. Schmitt and C.J. Elliott, Phys. Rev. **A41** (1990) 3853

[36] J.B. Murphy, C. Pellegrini and R. Bonifacio, Opt. Comm. **53** (1985) 197

[37] F. Kuypers, Klassische Mechanik, (VCH, Weinheim, 1990)

[38] T.M. Tran and J.S. Wurtele, LRP 392/90, CRPP Report, Lausanne, Switzerland, 1990

[39] C. Pellegrini, in: Free Electron Lasers, eds. S. Martellucci and A.N. Chester (Plenum Press, New York, 1983)

[40] B.M. Kincaid, J. Appl. Phys. **48** (1977) 2684

[41] W.B. Colson, IEEE J. Quantum Electron. **QE17** (1981) 1417

[42] A.E. Siegman, Lasers, (University Science Books, Mill Valley, CA, 1986)

[43] J.D. Jackson, Classical Electrodynamics, (John Wiley and Sons, New York, 1975)

[44] A. Doria *et al.*, Phys. Rev. Lett. **80** (1998) 2841

[45] R. Bonifacio, F. Casagrande and C. Pellegrini, Opt. Comm. **61** (1987) 55

[46] J.B. Murphy and C. Pellegrini, "Introduction to the Physics of the FEL", Proc. of the South Padre Island Conf., Springer (1986) 163

[47] J.M. Wang and L.H. Yu, Nucl. Instr. & Meth. **A250** (1986) 484

[48] K.J. Kim, Phys. Rev. Lett. **57** (1986) 1971

[49] R. Bonifacio and F. Casagrande, Opt. Comm. **50** (1984) 251

[50] E.L. Saldin *et al.*, Nucl. Inst. & Meth. **A381** (1996) 545

[51] S. Krinsky, AIP Proc. **153**, Physics of Particle Accelerators, SLAC, (1985) 1015

[52] R.A. Jong, E.T. Scharlemann and W.M. Fawley, Nucl. Instr. & Meth. **A272** (1988) 99

[53] T.I. Smith and J.M.J. Madey, Appl. Phys. **B27** (1982) 195

[54] J. Rossbach *et al.*., "Interdependence of Parameters of an X-ray FEL", DESY print TESLA-FEL 95-06, Hamburg, DESY, 1995

[55] E.L. Saldin, E.A. Schneidmiller and M.V. Yurkov, Opt. Comm. **97** (1993) 272

[56] E.L. Saldin, E.A. Schneidmiller and M.V. Yurkov, Phy. Rep. **260** (1995) 187

[57] G.T. Moore, Opt. Comm. **52** (1984) 46

[58] G.N. Watson, A Treatise on the Theory of Bessel Functions, (Cambridge University Press, Cambridge 1966)

[59] I.N. Bronstein and K.A. Semendjajew, Taschenbuch der Mathematik, (Harri Deutsch, Frankfurt/Main, 1987)

[60] M. Xie, D.A.G. Deacon, J.M.J. Madey, Phys. Rev. **A41** (1990) 1662

[61] A.E. Siegman, J. Opt. Soc. Am. **63** (1973) 1093

[62] S. Krinsky and L.H. Yu, Phys. Rev. **A35** (1987) 3406

[63] J. Goodman, Statistical Optics, (John Wiley and Sons, New York, 1985)

[64] R. Bonifacio *et al.*, Phys. Rev. Lett. **73** (1994) 70

[65] E.L. Saldin, E.A. Schneidmiller and M.V. Yurkov, Nucl. Inst. & Meth., **A407** (1998) 291

[66] E.L. Saldin, E.A. Schneidmiller and M.V. Yurkov, Nucl. Inst. & Meth., **A407** (1998) 285

[67] K.J. Kim, Nucl. Instr. & Meth. **A250** (1986) 396

[68] R. Bonifacio, B.W.J. McNeil and P. Pierini, Phys. Rev. **A40** (1989) 4467

[69] L. Mandel and E. Wolf, Optical Coherence and Quantum Optics, (Cambridge University Press, Cambridge, 1995)

[70] O. Forster, Analysis 3, (Vieweg, Braunschweig, 1984)

[71] R. Bonifacio *et al.*, Nuc. Instr. & Meth. **A296** (1990) 358

[72] R.H. Dicke, Phys. Rev. **93** (1954) 99

[73] E.T. Scharlemann *et al.*, Nucl. Instr. & Meth. **A250** (1986) 150

[74] E.T. Scharlemann and W.M. Fawley, SPIE 642 (1986) 1

[75] B.D. McVey, Nucl. Inst. & Meth. **A250** (1986) 449

[76] R.J. Dejus, O.A. Shevchenko and N.A. Vinokurov, "An Integral Equation Based Computer Code for High-Gain Free-Electron Lasers", Proc. of the 20th FEL conference, Williamsburg, VA, 1998

[77] E.L. Saldin, E.A. Schneidmiller and M.V. Yurkov, submitted to Nucl. Inst. & Meth.

[78] H.P. Freund, Phys. Rev. **E52** (1995) 5401

[79] G. Travish, AIP Proc. **413**, Towards X-Ray Free Electron Lasers, Lake Garda, (1997) 37

[80] C.K. Birdsall and A.B. Langdon, Plasma Physics via Computer Simulations, (McGraw-Hill, New York, 1985)

[81] W.H. Press *et al.*, Numerical Recipes in Fortran 77, (Cambridge University Press, Cambridge, 1992)

[82] J.H. Halton, Numerische Mathematik **2** (1960) 84

[83] J.M. Hammersley and D.C. Handscomb, Monte Carlo Methods, (Methuen, London, 1964)

[84] L. Devroye, Non-Uniform Random Variate Generation, (Springer, New York, 1986)

[85] C. Penman and B.W.J. McNeil, Opt. Comm. **80** (1992) 82

[86] C.W. Gear, Numerical Initial Value Problems in Ordinary Differential Equations (Englewood Cliffs, Prentice-Hall, 1971)

[87] J. Stoer and R. Bulirsch, Introduction to Numerical Analysis (Springer, New York, 1980)

[88] P. Deuflhard, Numerische Mathematik, **41** (1983) 399

[89] M. Berz and H. Wollnik, Nucl. Instr. & Meth. **A267** (1988) 25

[90] P. Jha and J.S. Wurtele , Nucl. Inst. & Meth. **A331** (1993) 477

[91] W.F. Ames, Numerical Methods for Partial Differential Equations (Academic Press, New York, 1977)

[92] A. Goldberg, H.M. Schey and J.L. Schwartz, Am. J. Phys. **35** (1967) 177

[93] R.D. Richtmyer and K.W. Morton, Difference Methods for Initial Value Problems (John Wiley and SOns, New York, 1967)

[94] A.R. Mitchell, Computational Methods in Partial Differential Equations, (John Wiley and Sons, New York, 1969)

[95] I. Galbraith, Y.S. Ching and E. Abraham, Am. J. Phys. **52** (1984) 60

[96] W. Joppich and S. Mijalkovic, Multigrid Methods for Process Simulation, (Springer, New York, 1993)

[97] A. Brandt, Math. Comp. **31** (1977) 333

[98] W.M. Fawley, "An Informal Manual for GINGER and its Postprocessor XPLOTGIN", LBID-2141, CBP Tech Note-104, UC-414, 1995

[99] "A VUV Free Electron Laser at the TESLA Test Facility at DESY: Conceptual Design Report", DESY print TESLA-FEL 95-03, Hamburg, DESY, 1995

[100] E.L. Saldin, E.A. Schneidmiller and M.V. Yurkov, "Calculation with FS2R Code of 6nm and 70nm Options of SASE FEL at the TESLA Test Facility", DESY print TESLA-FEL 95-02, Hamburg, DESY, 1995

[101] E.L. Saldin, E.A. Schneidmiller and M.V. Yurkov, "Numerical Simulation of the UCLA/LANL/RRCKI/SLAC Experiment on a High Gain SASE FEL", DESY print TESLA-FEL 98-05, Hamburg, DESY, 1998

[102] S.V. Milton *et al.*, "The APS SASE FEL: Initial Commissioning Results", Proc. of the PAC Conf., New York, 1999

[103] P. Castro, "Orbit Correction by Dispersion Minimization in an Undulator with Superimposed FODO Lattice", Proc. of the EPAC, Stockholm, 1998

[104] B. Faatz *et al.*, "Regenerative FEL Amplifier at the DESY TESLA Test Faciliy as a Fully Coherent VUV Laser", DESY print TESLA-FEL 97-07, Hamburg, DESY, 1997

[105] J. Feldhaus *et al.*, Opt. Comm. **140** (1997) 341

[106] P. Pierini and W.M. Fawley, Nucl. Instr. & Meth. **A375** (1996) 332

[107] B.M. Kincaid, J. Opt. Soc. Am. **B2** (1985) 1294

[108] H.D. Shay and E.T. Scharlemann, Nucl. Inst. & Meth. **A272** (1988) 601

[109] L.H. Yu and S. Krinsky, Phys. Lett. **A129** (1988) 463

[110] L.H. Yu *et al.*, Phys. Rev. **A45** (1992) 1163

[111] B.L. Bobbs *et al.*, Nucl. Inst. & Meth. **A296** (1990) 574

[112] B. Faatz, J. Pflüger and Y.M. Nikitina, Nucl. Phys. & Meth. **A393** (1997) 380

[113] B. Faatz, J. Pflüger and P. Pierini, Nucl. Phys. & Meth. **A375** (1996) 441

[114] P. Castro, private communication

[115] A.W. Chao, Physics of Collective Beam Instabilities in High Energy Accelerators, (John Wiley and Sons, New York, 1993)

[116] K. Bane, SLAC Report AP-87 (1991)

[117] O. Henry and O. Napoly, Part. Accel. **35** (1991) 235

[118] J.D. Lawson, Part. Accel. **25** (1990) 107

[119] K. Bane and M. Sands, Part. Accel. **25** (1990) 73

[120] P.B. Wilson, AIP Proc. **87**, Phys. High Energy Accel., Fermilab, (1981) 450

[121] A. Novokhatski and A. Mosier, Proc. of the PAC97 Conference, Vancouver, 1997

[122] G. Stupakov *et al.*, Physica Reviw Special Topics – Accelerators & Beams **2** (1999) 060701

[123] A. Novokhatski, M. Timm and T. Weiland, Proc. of the PAC99 Conference, New York, 1999

[124] Y. Derbenev, "RF Compensation of Aberrational Velocity Spread of Electrons in FEL", reported at: Light Source Dept. Seminar, BNL, 1990

[125] A.N. Sessler, D.H. Whittum and L.H. Yu, Phys. Rev.Lett. **68** (1992) 309

[126] E.D. Courant and H.S. Snyder, Ann. Phys. **3** (1958) 1

[127] M. Sands, Phys. Rev. **97** (1955) 470

[128] A. Messiah, Quantenmechanik Vol. 1, (Walter de Gruyter, Berlin, 1991)

[129] R.W.B. Best and B. Faatz, J. Phys. D: Appl. Phys. **23** (1990) 1337

[130] W. Fischer and I. Lieb, Funktionentheorie, (Vieweg, Braunschweig, 1988)

Acknowledgment

I want to thank DESY for the opportunity to work on the interesting field of accelerator physics. During these three years I got all the financial and technical support to finish successfully my Ph.D. thesis.

First of all I gratefully thank my advisors Dr. Jörg Roßbach and Prof. Bernd Sonntag. They introduced me to the exciting field of Free-Electron Lasers. My work had many aspects and covered numerical and analytic calculation as well as the commissioning and operation of the linac and FEL at the TESLA Test Facility.

I want to express my gratitude to Prof. Dr. Claudio Pellegrini and Prof. Dr. James Rosenzweig, who invited me to join their group at the University of California, Los Angeles. It was very delightful to work at their experiment on SASE FELs. I gained enormous amount of knowledge and experience.

I'm very grateful to Bart Faatz for his profound knowledge about numerics and simulation. His deep insight in the physics and numerics was very helpful for solving the problems I encountered during my work on GENESIS 1.3.

I always found help on the theory of FEL physics by asking Evgeni Saldin, Evgeni Schneidmiller and Mikail Yurkov. I appreciate the stimulating and helpful discussion with them.

In case of problems Bernhard Holzer was always willing to help me with beam optics. Georg Hoffstätter was the right person I could ask concerning general problems with theoretical physics. For the realization of GENESIS 1.3 on the DESY solar cluster I highly appreciate the help of Susan Wipf and Carsten Kluth.

I am indebted to Martin Dohlus, Klaus Flöttmann, Susan Wipf, John Maidment, Mark Lomperski, Holger Schlarb, Winfried Decking, Desmond Barber and Bart Faatz, who have read my manuscript and helped me to eliminate any error or doubtful statement both in content and in the English language.

During my stay at UCLA I met a lot of new people. Without them I would miss thrilling experiences such as tubing or Marathon, to name a few. Thanks to Mark Hogan, Pedro Frigola, Alex Murokh, Scott Anderson, Ding Xiaodong, Soren Telfer, Pietro Musumeci, Jason Wingo, Bryant Burkhart, and Matteo Pellegrini. In particular, I thank Aaron Tremaine for teaching me the most about the American way of life, as well as Gil Travish and Judith Hoye, who have always offered me a place to stay.

I have to thank my colleagues and friends at DESY. They supported my work with a lot of helpful discussion or just be making the time at DESY one of the most enjoyable in my life. Among these, which have not been mentioned so far, are: Michaela Marx, Ingrid Reyzl, Gisela Lüscher, Nicolette Baboi, Marie Berglund, Mari Paz Zorzano, Pedro Castro, Christoph Montag,

Mathias Vogt, Gerald Schmidt, Andreas Kabel, Andreas Jöstingmeier, Andreas Ullrich, Marco Bahte, Klaus Hanke, Marc Geitz and Thomas Schilcher.