Bunch Shaping in Electron Linear Accelerators

G. Ha, K.-J. Kim, J.G. Power,[*] and Y. Sun (孙银娥)

Argonne National Laboratory,
Argonne, Illinois 60439,
USA

P. Piot

Northern Illinois University,
Department of Physics and Northern Illinois Center for Accelerator & Detector Development,
DeKalb Illinois 60115,
USA
Argonne National Laboratory,
Argonne, Illinois 60439,
USA

Modern electron linear accelerators are often designed to produce smooth bunch distributions characterized by their macroscopic ensemble-average moments. However, an increasing number of accelerator applications call for finer control over the beam distribution, e.g., by requiring specific shapes for its projection along one coordinate. Ultimately, the control of the beam distribution at the single-particle level could enable new opportunities in accelerator science. This review discusses the recent progress toward controlling electron beam distributions on the "mesoscopic" scale with an emphasis on shaping the beam or introducing complex correlations required for some applications. This review emphasizes experimental and theoretical developments of electron-bunch shaping methods based on bounded external electromagnetic fields or via interactions with the self-generated velocity and radiation fields.

CONTENTS

---

[*] jp@anl.gov

I. INTRODUCTION

The ability to control the electron beam distribution produced by an linear accelerator (linac) evolved during the 20th century. The first generation of electron linacs (e-linacs) that begun with Wideroe's invention (Wideröe, 1928) produced continuous streams of electrons by placing a cathode in an electrostatic gap and therefore had no control over the longitudinal distribution and only modest control over the transverse distribution, based on the size of the hole in the anode plate. Control over the longitudinal distribution began when RF power generators, developed for radar applications, became available after World War II. At that time, Luis Alvarez (Alvarez, 1946)

proposed an accelerator based on a linear array of drift tubes enclosed in resonant cavities, and the second generation of e-linacs (operating with DC electron guns) was born (Ginzton et al., 1948). This generation of e-linacs culminated in the construction of the 100-GeV electron-positron SLAC linear collider (Dupen et al., 1968). Electron bunches in these linacs can be approximated by a Gaussian distribution in phase space which is characterized by its second order moments. The next significant progression in the control over the bunch distribution in e-linacs took place in the early 1990s with the widespread adaptation of the RF photocathode gun (Fraser et al., 1985) and development of the magnetic chicane compressor (Carlsten and Russell, 1996). This progress facilitated, for example, the development of X-ray free-electron lasers (FELs); for a review see (Kim et al., 2017).

Despite the progress that was made during the 20th century, ever more demanding accelerator applications continued to appear. The e-linac community responded to these challenges and is now on the verge of taking the next step in the evolution of control over the bunch distribution. There are two aspects in this effort: Towards a control at a level finer than the macroscopic scale but coarser than the microscopic scale, which will be referred to as *mesoscopic* level, and towards multi-dimensional beam shaping for distributions that can no longer be characterized by the second order moments; see Fig. 1(b). The next step in beam shaping will in general involve both of these aspects.

The ultimate challenge for beam control (Nagaitsev et al., 2021) is to produce interesting and useful distribution at the finest level – the *microscopic scale* where the distribution is described by a "granular" Klimontovich-distribution function (Klimontovich, 1995). Such an ultimate degree of control would open the path toward producing structured beams, e.g., Wigner-crystal beams (Wigner, 1934) with arbitrary shapes; see Fig. 1(c).

The development of multi-dimensional shaping began in the late 1990s and early 2000s. The flat-beam generation (Brinkmann et al., 2001; Derbenev, 1998) and emittance exchange between transverse and longitudinal phase spaces (Cornacchia and Emma, 2002; Kim and Sessler, 2006) were introduced in this period. Experimental demonstration followed; the flat beam generation by (Edwards et al., 2000; Piot et al., 2006) and emittance exchange by (Ruan et al., 2011). The emittance exchange opened up the possibility of shaping a beam in transverse dimension and transferring it to the temporal dimension, and vice versa. The technique was used to produce a train of sub pico-second bunchlets (Sun et al., 2010a) and bunches with linearly-ramped current profile (Ha et al., 2017a)

Advanced beam phase-space shaping is needed to enable many accelerator applications, e.g., improving the efficiency of beam-driven advanced acceleration tech-
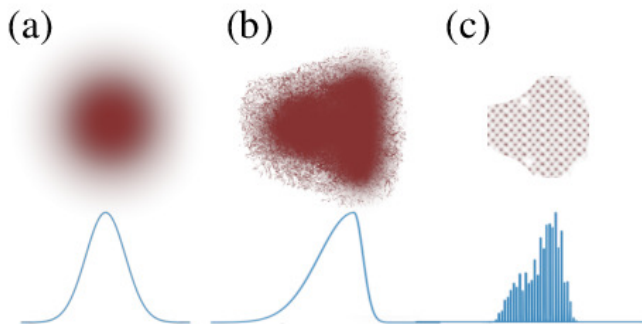
Figure 1 Definition of the level of control on a beam distribution: macroscopic (a), mesoscopic (b), and microscopic (c) scales. The shaded distributions depict the projected beam distribution in a plane defined by any pair of coordinates associated with the beam. The blue traces describe the projection of the distribution along one arbitrary axis.

niques (Bane *et al.*, 1985), improved X-ray free-electron laser interaction (Emma *et al.*, 2006), or the development of compact accelerator-based radiation sources (Gover *et al.*, 2019). Tailored electron beams are also used as a tool to manipulate hadron beams by exerting nonlinear focusing (Shiltsev, 2016) and cooling (Blaskiewicz, 2014). Finally beam distributions assuming known continuous function can also be used to mitigate beam degradation arising from collective effects, such as space-charge force (Kapchinskij and Vladimirskij, 1959; Kellogg, 1967) and self-interaction via radiative effects (Derbenev *et al.*, 1995; Seeman, 1992)

The present review is mainly devoted to phase-space shaping techniques employing bounded external electromagnetic fields or via interactions with self-generated velocity and radiation fields. Techniques that couple lasers with electron beams are not included here since they been have reviewed in (Hemsing *et al.*, 2014).

We categorize these techniques into three broad classes. The first category includes techniques that control the electron distribution from the electron gun. This includes shaping the distribution of the emission-triggering laser in photoemission electron sources or engineering the cathode properties or surface to control the emitted electron distribution. The second class of manipulation consists of shaping systems that operate within one degree of freedom, e.g., those based on the use of external and internal fields to control the distribution in one of the three phase-space planes. Finally, the third category involves shaping techniques that also use the external and internal fields to introduce correlations between two degrees of freedom.

This article is organized as follows. Section II introduces the fundamental concepts necessary to the understanding of phase-space manipulations, and the subsequent Sections III, IV, V, and VI discuss the various classes of manipulation mentioned above. Finally, Section VII offers some perspective on likely research directions motivated by recent developments.

II. GENERAL PRINCIPLES

A. Hamiltonian formalism

1. Equation of motion under EM fields

We consider the motion of a charged particle under electromagnetic field $\mathbf{E}$ and $\mathbf{B}$ satisfying Maxwell equations:

$$\nabla \cdot \mathbf{E} = \frac{\rho_S}{\varepsilon_0}, \quad \nabla \times \mathbf{E} = -\frac{\partial}{\partial t}\mathbf{B}, \tag{1}$$

$$\nabla \cdot \mathbf{B} = 0, \quad \nabla \times \mathbf{B} = \mu_0 \mathbf{J}_S + \frac{1}{c^2}\frac{\partial}{\partial t}\mathbf{E}.$$

Here ρ_S and $\mathbf{J}_S$ are the charge density and current density of the sources, respectively, consisting of the external sources and the beam itself. External sources are not present within the beam pipes. Thus $\rho_S = 0$ and $\mathbf{J}_S = 0$ until we consider the beam generated fields in II.D. We will use the MKS units throughout this paper. Introducing the vector and the scalar potential, $\mathbf{A}$ and ϕ, respectively, the electromagnetic fields can be written as follows:

$$\mathbf{B} = \nabla \times \mathbf{A} \text{ and } \mathbf{E} = -\frac{\partial}{\partial t}\mathbf{A} - \nabla\phi. \tag{2}$$

The Hamiltonian H for a particle of mass m and charge e is

$$H(\mathbf{x}, \mathbf{p}, t) = \sqrt{m^2 c^4 + c^2(\mathbf{p} - e\mathbf{A}(\mathbf{x}, t))^2} + e\phi(\mathbf{x}, t). \tag{3}$$

Here c is the velocity of light, $\mathbf{x}$ the coordinate vector, and $\mathbf{p}$ is the canonical momentum conjugate to $\mathbf{x}$.

The Hamiltonian equations of motion are

$$\frac{d\mathbf{x}}{dt} = \frac{\partial H}{\partial \mathbf{p}}, \text{ and } \frac{d\mathbf{p}}{dt} = -\frac{\partial H}{\partial \mathbf{x}}. \tag{4}$$

The Hamilton's equations reproduce the Lorentz force equation:

$$\frac{d\mathbf{p}_{kin}}{dt} = e\left(\mathbf{E} + \mathbf{v} \times \mathbf{B}\right). \tag{5}$$

Here, $\mathbf{p}_{kin}$ is the kinetic momentum related to the canonical momentum $\mathbf{p}$ as follows:

$$\mathbf{p}_{kin} \equiv m\gamma\frac{d\mathbf{x}}{dt} = \mathbf{p} - e\mathbf{A}, \tag{6}$$

where γ = relativistic kinetic energy$/mc^2 = 1/\sqrt{1 - \beta^2}$, $\beta = v/c$, $v = |d\mathbf{x}/dt|$.

nian is ;

$$\mathcal{H}\left(\mathbf{x}_{new}, \mathbf{p}_{new}; s\right) = -eA_s - \left(1 + \frac{x}{\rho(s)}\right) \times$$

$$\sqrt{\left(\frac{U - e\phi}{c}\right)^2 - m^2c^2 - (p_x - eA_x)^2 - (p_y - eA_y)^2}. \tag{9}$$

Here $\rho(s)$ is the radius of curvature, $A_s = (1 + x/\rho(s)) \mathbf{A} \cdot \mathbf{e}_s$, where $\mathbf{e}_s$ is the unit vector in the tangential direction. Since the transformation is canonical, the equation of motion is the same as in Eq.(4) with the replacement $t \to s$, $\mathbf{x} \to \mathbf{x}_{new}$, $\mathbf{p} \to \mathbf{p}_{new}$, and $H \to \mathcal{H}$.

The second transformation is to *deviation variables* relative to the trajectory of the reference particle $(x_0, y_0, -t_0; p_{0x}, p_{0y}, U_0) = (0, 0, -t_0; 0, 0, U_0)$. The deviation variables are therefore $(x, y, -t+t_0; p_x, p_y, U - U_0)$. The change to deviation variables is also a canonical transformation, in which the Hamiltonian $\mathcal{H}_D$ is the same as Eq. (9) except the zeroth and the linear terms in the power series expansion of $\mathcal{H}$ are excluded (Duffy and Dragt, 2016).

When the electric field is absent and the magnetic fields are static and perpendicular to $\mathbf{e}_s$, we can choose $A_x = A_y = \phi = 0$. Also, the energy U and the curvature ρ_0 are constant. In this case, $\mathcal{H}_D/p_0$, where p_0 is the momentum of the reference particle, can be chosen as a scaled Hamiltonian with canonically conjugate variables $(x, y, -v_0(t - t_0); p_x/p_0, p_y/p_0, (U - U_0)/(v_0 p_0))$ (MacKay and Conte, 2012; Ruth, 1986). Note we introduced an additional scale factor v_0, the velocity of the reference particle, for the longitudinal variables. For the highly relativistic case, $v \approx c$ if we ignore the correction factor of $1/2\gamma^2$, then we can approximate

$$\frac{p_x}{p_0} \approx \frac{dx}{ds} = x', \; \frac{p_y}{p_0} \approx \frac{dy}{ds} = y', \; \frac{U - U_0}{v_0 p_0} \approx \frac{p_s - p_0}{p_0} = \delta. \tag{10}$$

We will introduce the following notation for the longitudinal deviation variable:

$$- c(t(s) - t_0(s)) = z(s) \tag{11}$$

The quantity $z(s)$ is the distance (in general the arc distance) *ahead* of the reference particle along s, thus arriving there at an *earlier* time. Thus, the canonical variables in the present case are

$$(x, y, z; p_x, p_y, \delta). \tag{12}$$

The corresponding scaled Hamiltonian is

$$\mathcal{H}_S(x, y, z; x', y', \delta; s)$$
$$\approx \left[-e\frac{A_s}{p_0} - \left(1 + \frac{x}{\rho}\right)\sqrt{(1 + \delta)^2 - x'^2 - y'^2}\right]$$
$$\approx -e\frac{[A_s]}{p_0} - \frac{x}{\rho}\delta + +\frac{1}{2}\left(x'^2 + y'^2\right) + ... \tag{13}$$
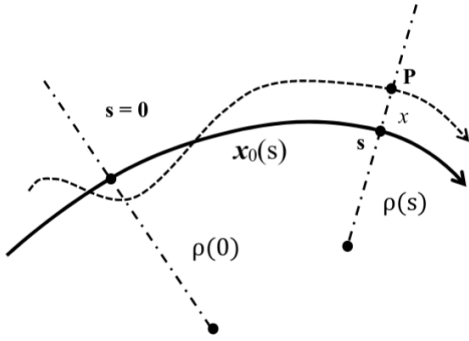

Figure 2 Curvilinear coordinate system for particle motion. See the text for details.

2. Curvilinear coordinates

In a beam, particles are bunched in a small region of space and stay together while moving. Therefore, it makes sense to introduce a "reference" particle as the one that is at the beam "center". Its trajectory is referred to as the reference orbit, which may be curved but will be assumed to lie on a plane referred to as the horizontal plane. The reference orbit is parametrized as $\mathbf{x}_0(s)$ where s is the arc length along the reference orbit. Then the position of any particle in the beam can be represented as:

$$\mathbf{x} = \mathbf{x}_0(s) + x\mathbf{e}_\rho + y\mathbf{e}_y, \tag{7}$$

where $\mathbf{e}_\rho$ is the unit vector normal to the reference orbit at s on the horizontal plane, $\rho(s)$ the radius of curvature, and $\mathbf{e}_y$ is the unit vector normal in the vertical direction. The set (x, y, s) constitutes the curvilinear coordinate system shown in Fig. 2. The figure is drawn on the horizontal plane on which the reference trajectory $\mathbf{x}_0(s)$ (solid line) lies. The unit vector $\mathbf{e}_\rho$ is in the horizontal plane and perpendicular to the reference trajectory and the unit vector $\mathbf{e}_y$ is directed towards the reader. The line perpendicular to the reference trajectory at s represents the *transverse plane* extending in the vertical direction. The dotted line represents the projection in the horizontal plane of an arbitrary particle, intersecting the transverse plane at s at (x, y).

We now make two canonical transformations (Goldstein, 2002; Landau and Lifshitz, 1969) to change the variables to ones convenient for studying beams in accelerators: First, we adopt s as the independent variable and use $(x, y, -t)$ as the coordinates (Courant and Snyder, 1958), where t is the time the particle arrives at the transverse plane at s. The new canonically conjugate variables are:

$$\mathbf{x}_{new} = (x, y, -t), \; \mathbf{p}_{new} = (p_x, p_y, U), \tag{8}$$

where U is the energy of the particle The new Hamilto-

In the above [..] implies removing the zeroth and the first order terms in the power series expansion in the scaled variables of the expression inside the square brackets.

When acceleration is present, then we need to go back to the original Hamiltonian equation of motion Eq.(4) or the Lorentz force equation Eq.(5), as will be done in Section II.B.2. The presence of longitudinal magnetic field can be treated by working in a rotating frame, as discussed Section II.B.4.

3. Symplecticity

Our goal in this section is to illustrate the special property enjoyed by a mechanical system that can be described by the Hamiltonian equation of motion. We will mostly consider the cases in which the variables Eq.(12) and the Hamiltonian Eq.(13) are applicable. This is not valid when acceleration is present as in Section II.B.2, where we will revert to the variables $(\mathbf{x}, \mathbf{p})$.

We rearrange the 6D canonical variables into a column vector $\boldsymbol{\mathcal{Z}}$ as follows:

$$\boldsymbol{\mathcal{Z}} = \begin{pmatrix} x \\ x' \\ y \\ y' \\ z \\ \delta \end{pmatrix} = \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \zeta_3 \\ \zeta_4 \\ \zeta_5 \\ \zeta_6 \end{pmatrix}. \quad (14)$$

It is often useful to introduce the subspace as follows:

$$\boldsymbol{\mathcal{Z}} = \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \end{pmatrix}; \mathbf{X} = \begin{pmatrix} x \\ x' \end{pmatrix}, \mathbf{Y} = \begin{pmatrix} y \\ y' \end{pmatrix}, \mathbf{Z} = \begin{pmatrix} z \\ \delta \end{pmatrix}. \quad (15)$$

Note these are the canonical deviation variables in the curvilinear coordinates introduced in Section II.A.2, *not* the usual Cartesian variables of the laboratory frame.

Introducing the gradient vector ∇ in 6D phase space

$$\nabla_j = \frac{\partial}{\partial \zeta_j}, \quad (16)$$

the equation of motion with the scaled Hamiltonian can be written as

$$\frac{d}{ds}\boldsymbol{\mathcal{Z}} = J\nabla \mathcal{H}_{\mathcal{S}}(\boldsymbol{\mathcal{Z}}; s). \quad (17)$$

Here, we have introduced the unit symplectic matrix:

$$J = \begin{pmatrix} J_{2D} & 0 & 0 \\ 0 & J_{2D} & 0 \\ 0 & 0 & J_{2D} \end{pmatrix}, \quad J_{2D} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}. \quad (18)$$

By solving Eq.(17), the map M $\boldsymbol{\mathcal{Z}} \to \bar{\boldsymbol{\mathcal{Z}}}$ corresponding to a section of the accelerator from s to $\bar{s}$ can be found :

$$\bar{\boldsymbol{\mathcal{Z}}} = M\left(\boldsymbol{\mathcal{Z}}\right); \quad \bar{\zeta}_i = M_i\left(\boldsymbol{\mathcal{Z}}\right). \quad (19)$$

The inverse map is

$$\boldsymbol{\mathcal{Z}} = M^{-1}\left(\bar{\boldsymbol{\mathcal{Z}}}\right). \quad (20)$$

We introduce the Jacobian matrix R whose components are:

$$R_{ij}\left(\boldsymbol{\mathcal{Z}}\right) = \frac{\partial \bar{\zeta}_i}{\partial \zeta_j}. \quad (21)$$

For Hamiltonian dynamics, the Jacobian matrix is symplectic:

$$R^T J R = R J R^T = J. \quad (22)$$

It follows from Eq. (22) and the continuity of R as $\bar{s} \to s$ that its determinant is unity:

$$\det(R) = 1. \quad (23)$$

All 2×2 matrices with unit determinant are symplectic. For higher dimensions, Eq. (22) imposes significant restrictions on the matrix.

In the following we will mostly consider the case where the transformation in Eq. (19) is linear:

$$\bar{\boldsymbol{\mathcal{Z}}} = R\boldsymbol{\mathcal{Z}}. \quad (24)$$

Here R is the Jacobian matrix given by Eq. (21) whose elements are independent of $\boldsymbol{\mathcal{Z}}$.

B. Single particle motion in external field

In this sub-section, we present some important examples of the transformation matrix R relevant for beam shaping.

1. Free space, bending magnets, and quadrupole magnets

With no electric fields, and the components of the static magnetic fields corresponding to a bending magnet and quadrupole are given by

$$B_x = B_1(s)y, \ B_y = -B_0(s) + B_1(s)x. \quad (25)$$

Here $B_0 = p_0/(e\rho)$ is the strength of the dipole magnet bending the particle horizontally. In computing the vector potential A_s one finds the scaled Hamiltonian Eq.(13) up to the quadratic terms:

$$\mathcal{H}_S = -\delta\frac{x}{\rho} + \frac{x^2}{2\rho^2} + \frac{K_q}{2}\left(-x^2 + y^2\right) + \frac{1}{2}\left(x'^2 + y'^2\right) \quad (26)$$

Here $K_q = eB_1/(p_0c)$ is the quadrupole strength. When $K_q = 1/\rho = 0$, this will be free space. The equation of motion is obtained from the Hamiltonian equation Eq.(17). The equations in the transverse directions, after

reducing first order differential equations to second order ones, become:

$$\frac{d^2x}{ds^2} + \left(-K_q(s) + \frac{1}{\rho(s)^2}\right)x = \frac{1}{\rho(s)}\delta, \quad (27)$$

$$\frac{d^2y}{ds^2} + K_q(s)y = 0. \quad (28)$$

In deriving these equations, we are using the symbol x' both as the canonical momentum and as the slope dx/ds. We indicated that ρ and K_q are functions of s. However, we assume the functions are piece-wise constant and neglect the transition effects. Equation(27) shows that the motion in the x direction is influenced both the quadrupole force as well as the centripetal force due to the curvature. For historical reason, the motion described by Eqs.(27) and (28) is known as the betatron motion. Although the sign of the quadrupole focusing strength K_q in the x direction is opposite to that in the y-direction, focusing in both direction can be achieved either by the centripetal focusing (weak focusing) or by arranging the adjacent quadrupoles to have opposite sign (strong focusing). See text books for details, for example (Wiedemann, 1999).

The solution of Eq. (28) can be written in the following form (Courant and Snyder, 1958):

$$y = \sqrt{2J_y\beta_y}\cos(\psi_y),$$
$$y' = \frac{\beta_y'}{\beta_y}y - \sqrt{\frac{2J_y}{\beta_y}}\sin(\psi_y) \quad (29)$$

Here J_y is a constant, and

$$\psi_y = \int_0^s \frac{d\bar{s}}{\beta_y(\bar{s})} + \psi_{y0}, \quad (30)$$

where ψ_{y0} is another constant. The function $\beta_y(s)$ is one of the Courant-Snyder amplitudes, commonly referred to as the beta function. The other two are:

$$\alpha_y = -\frac{1}{2}\beta_y', \quad \gamma_y = \frac{1+\alpha_y^2}{\beta_y}. \quad (31)$$

The oscillatory motion described by Eq.(29) is known as the betatron motion, which has an invariant known as the Courant-Snyder invariant given by:

$$\gamma_y y^2 + 2\alpha_y yy' + \beta_y y'^2 = J_y. \quad (32)$$

The pair (ψ_y, J_y) are known as the angle-action variables in classical mechanics (Goldstein, 2002; Ruth, 1986). The beta function is determined by the following nonlinear, second order differential equation and appropriate boundary conditions:

$$2\beta_y\beta_y'' - \beta_y'^2 + 4\beta_y^2 K_q = 4. \quad (33)$$

The advantage of writing the solution in the form of Eq.(29) is that the betatron motion is specified by two distinct characteristics, the initial conditions associated with each particle via J_y and ψ_{y0} and the magnet arrangement of the beamline via the beta function.

The homogeneous part of the Eq. (27) has the same structure as that of Eq.(28). Thus, we have the betatron motion in the x-direction and its associated Courant-Snyder invariant as well. By forming linear combinations of solutions of the form given in Eq.(29) with appropriate constants J_x and ψ_{x0}, we can construct two independent solutions of the homogeneous part of the equation, $C_{s\tau}$ (cosine-like) satisfying $C_{\tau\tau}=1$, $C'_{\tau\tau}=0$ and $S_{s\tau}$ (sine-like) with $S_{\tau\tau}=0$, $S'_{\tau\tau}=1$. Then the full solution can be written as follows (Brown, 1968; Wiedemann, 1999):

$$x_s = C_{s\tau}x_\tau + S_{s\tau}x'_\tau + \eta_{s\tau}\delta,$$
$$x'_s = C'_{s\tau}x_\tau + S'_{s\tau}x'_\tau + \eta'_{s\tau}\delta. \quad (34)$$

Here

$$\eta_{s\tau} = S_{s\tau}\int_\tau^s d\zeta\, C_{\zeta\tau}/\rho(\zeta) - C_{s\tau}\int_\tau^s d\zeta\, S_{\zeta\tau}/\rho(\zeta). \quad (35)$$

The longitudinal motion is given by

$$\frac{dz}{ds} = -\frac{x}{\rho(s)},$$
$$\frac{d\delta}{ds} = 0, \quad (36)$$

with the solution

$$z_s = z_\tau - \int_\tau^s d\zeta\frac{x_\zeta}{\rho(\zeta)}, \quad \delta_s = \delta_\tau. \quad (37)$$

Inserting Eqs.(34) into Eq.(37) and collecting results so far, we obtain the R matrix for transformation from τ to s in $\mathbf{X}$, $\mathbf{Z}$ space:

$$R_{s\tau} = \begin{pmatrix} C_{s\tau} & S_{s\tau} & 0 & \eta_{s\tau} \\ C'_{s\tau} & S'_{s\tau} & 0 & \eta'_{s\tau} \\ R_{51s\tau} & R_{52s\tau} & 1 & R_{56s\tau} \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (38)$$

The $R_{5,j}$ elements the above are

$$(R_{51s\tau}, R_{52s\tau}, R_{56s\tau}) = -\int_\tau^s d\zeta\,(C_{\zeta\tau}, S_{\zeta\tau}, \eta_{\zeta\tau})/\rho(\zeta), \quad (39)$$

The matrix given by Eq. (38) satisfies the symplectic condition, Eq. (22).

For free space (or drift space), we have $\rho \to \infty$, $\eta = \eta' = R_{5j} = 0$ and the 2×2 upper-left bloc of $R_{s\tau}$ becomes

$$D_\ell = \begin{pmatrix} 1 & \ell \\ 0 & 1 \end{pmatrix}. \quad (40)$$

Here $\ell = s-\tau$ is the length of the free space. If we have a thin quarupole of focal length f_C, then the bock becomes

$$F_{f_C} = \begin{pmatrix} 1 & 0 \\ -1/f_C & 1 \end{pmatrix}. \quad (41)$$

Using Eq. (39), the matrix for a sector bending magnet of constant ρ and deflection angle θ can be found, see Eq.(77) in (Brown, 1968). The matrix for the corresponding rectangular magnet is obtained by multiplying a defocusing lens of focal length $\rho/\tan(\theta/2)$ on both sides of the sector magnet matrix with the result :

$$R_{Brect}(\rho) =$$
$$\begin{pmatrix} 1 & \rho\sin(\theta) & 0 & \rho(1-\cos(\theta)) \\ 0 & 1 & 0 & 2\tan(\theta/2) \\ -2\tan(\theta/2) & -\rho(1-\cos(\theta)) & 1 & -\rho(\theta-\sin(\theta)) \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$
$$(42)$$

This matrix is also given in (Brown, 1999).[1]

A *dogleg* consisting of a rectangular bending magnet, a drift of length ℓ , and a reverse rectangular bending magnet has the following matrix:

$$R_{DL} = \begin{pmatrix} 1 & \ell & 0 & \eta \\ 0 & 1 & 0 & 0 \\ 0 & \eta & 1 & \xi \\ 0 & 0 & 0 & 1 \end{pmatrix}. \qquad (43)$$

Here

$$\ell = d + 2\rho\sin(\theta),$$
$$\eta = 2\left(d + \rho\sin(\theta)\right)\tan\left(\frac{\theta}{2}\right), \qquad (44)$$
$$\xi = -2\rho\left(\theta + \sin(\theta)\right) + 8\rho\tan\left(\frac{\theta}{2}\right) + 4d\tan^2\left(\frac{\theta}{2}\right).$$

The dogleg is often employed in beam shaping, such as for bunch compression and emittance exchange (Sections II.C.3 and VI.C.1).

2. RF photo-cathode cavity

An RF photo-cathode cavity is an important device for generating bright beams (Fraser *et al.*, 1985). The cavity will be assumed to have cylindrical symmetry and shown schematically in Fig. 3 . Since acceleration is involved, we need to use the canonical variables introduced in Eq.(8) in Section II.A.1 with the independent coordinate as the distance along the axis of the cavity s. The discussion here follows closely that in reference (Kim, 1989).

The non-vanishing components of the EM field , keeping terms quadratic in radial coordinate r, take the fol-

---

[1] However, the sign convention in this reference such that the z coordinate has the opposite sign to ours. Thus, the $(5,j)$ and $(j,5)$ elements there should be multiplied by -1. Otherwise, the matrix is not symplectic.
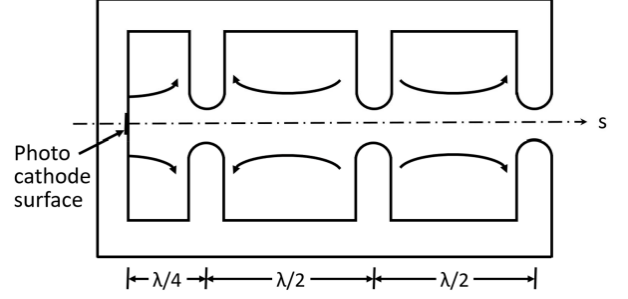
Figure 3 A schematics of an RF photo-cathode cavity. The first cell where photocathode is located is a half cell. From Ref. (Kim, 1989).

lowing form:

$$E_s = \left(\mathcal{E}(s) - \frac{r^2}{4}\left(\mathcal{E}''(s) - k^2\mathcal{E}(s)\right)\right)\sin(\omega t),$$
$$E_r = -\frac{1}{2}r\mathcal{E}'(s)\sin(\omega t),$$
$$B_\varphi = \frac{k}{2c}r\mathcal{E}(s)\cos(\omega t). \qquad (45)$$

Here $k = 2\pi/\lambda = \omega/c$, λ is the RF wavelength, and

$$\mathcal{E}(s) = E_0\Theta(s)\cos(kz). \qquad (46)$$

Here, E_0 is the peak on-axis electric field and $(') = (d/ds)$. The function $\Theta(s)$ accounts for the cavity exit at $s = s_f = (n+1/2)\lambda/2$; it is unity inside the cavity and decreases rapidly to zero as s leaves the cavity at s_f. Therefore Θ is similar to the delta function. Inside the cavity $0 \leq s < s_f$, the transverse field E_r and B_φ are linear in r while E_s is independent of r. These properties are crucial in minimizing the emittance growth, especially near the cathode located at $s = 0$. The parameters of the finite profile as shown in Fig. 3 can be chosen to produce fields close to those in Eq. (45) (McDonald, 1988).

a. Longitudinal motion For the longitudinal motion on the axis $r = 0$, the canonical variables can be taken as (ψ, γ) where

$$\psi = \omega t - ks = k\left(ct - s\right) = \psi_0 - kz. \qquad (47)$$

Here ψ_0 is the reference phase. Note z is the same as that in Eq.(11) for particles moving with relativistic velocity c. The equations of motion are

$$\frac{d\psi}{ds} = k\left(\frac{\gamma}{\sqrt{\gamma^2-1}} - 1\right)$$
$$\frac{d\gamma}{ds} = \frac{eE_0}{2mc^2}\left[\sin(\psi) + \sin(\psi + 2ks)\right]. \qquad (48)$$

For photo-cathode cavity operation, we need to solve Eq.(48) with the initial condition $(\psi, \gamma) = (\psi_i, 1)$. Noting that the RHS of the phase equation is appreciable only near the cathode, $0 \leq ks \ll 1$, an approximate solution for the entire range $0 \leq s \leq n\lambda/2$ can be found:

$$\psi = \frac{1}{2\alpha \sin(\psi_i)} \left[\sqrt{\tilde{\gamma}^2 - 1} - (\tilde{\gamma} - 1) \right] + \psi_i,$$
$$\gamma = 1 + \alpha \left[ks \sin(\psi_i) + \frac{1}{2} (\cos(\psi) - \cos(\psi + 2ks)) \right] \quad (49)$$

Here ψ_i is the initial phase and

$$\tilde{\gamma} = 1 + 2\alpha \sin(\psi_i) ks \quad (50)$$

and we introduced the acceleration strength parameter

$$\alpha = \frac{eE_0}{2mc^2 k}. \quad (51)$$

The approximate solution is fairly accurate for $3+1/2$ cell cavity with $\alpha = 1$ and $30° \leq \psi_i \leq 70°$. Improvement of the approximation is discussed in reference (Flöttmann, 2015). At the exit, the particle would have accelerated to a high energy $\gamma_f \gg 1$. The final phase ψ_f becomes

$$\psi_f = \frac{1}{2\alpha \sin(\psi_i)} + \sin(\psi_i). \quad (52)$$

Thus the ratio of the final to initial phase spread, or the the bunch length is

$$\frac{\Delta \psi_f}{\Delta \psi_i} = 1 - \frac{\cos(\psi_i)}{2\alpha \sin^2(\psi_i)}. \quad (53)$$

The bunch at the photo-cathode cavity exit will be compressed if $0 < \psi_0 < \pi/2$. Equation (48) can be cast in Hamiltonian form by using the longitudinal Hamiltonian

$$\mathcal{H}_L(\psi, \gamma; s) = \frac{eE_0}{2mc^2} [\cos(\psi) + \cos(\psi + 2ks)] + \sqrt{\gamma^2 - 1} + \arctan\left(\frac{1}{\sqrt{\gamma^2 - 1}}\right). \quad (54)$$

Therefore the map connecting the initial to the final points in the (ψ, γ) phase space is symplectic and area-preserving. However, the transformation is nonlinear and will lead to an increase in the effective longitudinal emittance.

b. Transverse motion   The Lorentz force given by Eq.(5) is purely radial with the magnitude

$$F_r = e(E_r - v_s B_\varphi). \quad (55)$$

Here $v_s = ds/(dt)$. Using Eq.(45), the force can be written in the following form:

$$F_r = eE_0 \frac{1}{2}\rho \left[-\frac{1}{c}\frac{d}{dt} (\Theta \sin(ks) \cos(\omega t)) - \Theta' \sin(\psi) + (1-\beta)\Theta' \sin(ks) \cos(\omega t) \right] \quad (56)$$

The change of the radial momentum from the cathode to exit is the time integral of the above equation. The contribution of the last term is negligible for relativistic velocities and the first term integrates to zero. The contribution of the middle term comes only from the exit region with the result

$$\Delta p_r = eE_0 \frac{r}{2c} \sin(\psi_f). \quad (57)$$

Equation(57) gives rise to a bow-tie shaped (r, p_r) phase space at the cavity exit due to the different exit phase ψ_f of each particle, leading to an increase in the projected emittance in the (r, p_r) space.

The radial force due to the space charge fields, Eq.(135), since it also depends on the longitudinal position z, also leads to an emittance increase. This effect dominates the transverse emittance growth since the space-charge force is not localized at the cavity exit but persistent from the beam creation at the cathode. Fortunately, the increase can be corrected by the emittance compensation technique (Carlsten, 1989; Ferrario *et al.*, 2007, 2000; Flöttmann, 2017; Miginsky, 2009; Serafini and Rosenzweig, 1997; Wang, 2006).

3. Transverse deflecting cavity

In the RF photo-cathode cavity considered in previous section, the transverse deflection occur only at the cavity exit. To see how a sustained deflection can be induced, write the transverse part of the third and fourth of Maxwell equations, Eq.(1), as follows:

$$\frac{\partial}{\partial s}\mathbf{E}_\perp - \mathbf{e}_s \times \frac{\partial}{\partial t}\mathbf{B}_\perp = -\mathbf{e}_s \times \nabla_\perp E_s,$$
$$\frac{\partial}{c^2 \partial t}\mathbf{E}_\perp - \mathbf{e}_s \times \frac{\partial}{\partial s}\mathbf{B}_\perp = -\mathbf{e}_s \times \nabla_\perp B_s. \quad (58)$$

The first of the above is the differential form of Panofsky-Wenzel theorem (Panofsky and Wenzel, 1956) and the second was noted by (Paramonov and Flöttmann, 2019). Let's consider a traveling wave with fields in the following form:

$$(\mathbf{E}_\perp, \mathbf{B}_\perp) = \left(\tilde{\mathbf{E}}_\perp(x,y), \tilde{\mathbf{B}}_\perp(x,y)\right) \sin(\psi),$$
$$E_s = \tilde{E}_s(x,y) \cos(\psi). \quad (59)$$

Assuming that the phase velocity of the wave and the particle velocity are both equal to c, we can show that the LHS of both equations are proportional to the transverse force. From the expressions on the RHS, it then follows
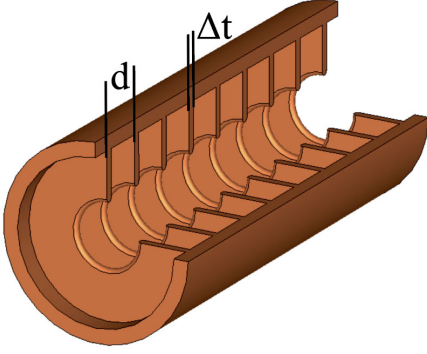
Figure 4 Iris-loaded structure producing HEM modes for transverse deflection. Δt is the thickness of the discs and d the structure periodicity. From Ref (Paramonov and Flöttmann, 2019).

that the transverse gradient of both E_s and B_s should not vanish for a sustainable transverse force . An iris-loaded structure, such as shown in Fig. 4, supports hybrid electromagnetic (HEM) modes, that can provide a sustained transverse force. The mode was studied by many authors since 1963 (Flöttmann and Paramonov, 2014; Garault, 1964; Hahn, 1963; Paramonov and Flöttmann, 2019).

When the iris thickness Δt is much smaller than the structure period d, which in turn is much smaller than the mode wavelength, the electric field should be in the radial direction at $r = a$. The field components up to the second order of polynomials in x and y are found to be

$$E_x = -\hat{E}\frac{k^2}{4}\left(a^2 + x^2 - y^2\right)\sin\left(\psi\right),$$

$$E_y = -\hat{E}\frac{k^2}{2}xy\sin\left(\psi\right),$$

$$E_s = \hat{E}kx\cos\left(\psi\right),$$

$$cB_x = \hat{E}\frac{k^2}{2}xy\sin\left(\psi\right),$$

$$cB_y = -\hat{E}\frac{k^2}{4}\left(a^2 - 4/k^2 + x^2 - y^2\right)\sin(\psi),$$

$$cB_s = -\hat{E}ky\cos\left(\psi\right). \tag{60}$$

The Lorentz force is then

$$\mathbf{F} = (F_x, F_y, F_s) = e\hat{E}\left(-\sin\left(\psi\right), 0, kx\cos\left(\psi\right)\right) \tag{61}$$

For beam shaping applications, we wish the transverse deflection changes sign as z varies across a bunch so that the head and tail receive an opposite kick. Thus, we choose $\psi_0 = 0$ in Eq.(47). Assuming $kz \ll 1$, the force to first order in the deviation variables (x, z) becomes

$$\mathbf{F} \approx e\hat{E}\left(kz, 0, kx\right) \tag{62}$$

The equation of motion inside the HEM structure is then

$$\frac{dx}{ds} = x', \quad \frac{dx'}{ds} = \frac{\kappa_T}{\ell}z,$$
$$\frac{dz}{ds} = 0, \quad \frac{d\delta}{ds} = \frac{\kappa_T}{\ell}x. \tag{63}$$

Here ℓ is the length of the cavity and

$$\kappa_T = \frac{e\hat{E}\ell k}{mc^2\gamma} \tag{64}$$

is a parameter characterizing the deflection strength. In Eq.(63), dz/ds vanishes since particles are highly relativistic. The solution of Eq.(63) can be written by the following matrix connecting the entrance and exit values of the vector $(\mathbf{X}, \mathbf{Z}) = (x, x', z, \delta)$ (Cornacchia and Emma, 2002):

$$R_{TDC} = \begin{pmatrix} 1 & \ell & \kappa_T\ell/2 & 0 \\ 0 & 1 & \kappa_T & 0 \\ 0 & 0 & 1 & 0 \\ \kappa_T & \kappa_T\ell/2 & \kappa_T^2\ell/6 & 1 \end{pmatrix}. \tag{65}$$

The matrix R_{TDC} has the desirable property that $(R_{TDC})^n$ is the same as the matrix for a cavity n times longer, corresponding to the matrix obtained by substituting $\ell \to n\ell$ and $\kappa_T \to n\kappa_T$ in Eq.(65).

The $(6, 5)$ element in R_{TDC} wreaks havoc in the transverse-to-longitudinal emittance exchange discussed in Section VI. However, it can be removed by suitable accelerating cavities (Zholents and Zolotorev, 2011).

4. Axial magnetic field

A cathode may be immersed in an axial magnetic field to guide and focus electron beams (Reiser, 1994) and for certain beam shaping purpose. Assuming cylindrical symmetry around the s-axis perpendicular to the cathode surface, the vector potential can be written as

$$\mathbf{A}(r, s) = \frac{1}{2}rB\left(s\right)\mathbf{e}_\varphi. \tag{66}$$

Here we are using the cylindrical coordinate (r, φ, s) with origin at the cathode center, and s is the distance away from the cathode. The magnetic field is

$$\mathbf{B}(r, s) = \nabla \times \mathbf{A} = B(s)\mathbf{e}_s - \frac{1}{2}rB'\mathbf{e}_\rho. \tag{67}$$

The second term involving $B' = dB/ds$ gives the magnetic flux spreading out radially outside the solenoidal coil. We are assuming the variation in s is slow, $B'' \approx 0$, so that $\nabla \times \mathbf{B}$ vanishes.

The transverse motion is given by

$$\frac{d^2\mathbf{x}}{ds^2} = -\frac{eB\left(s\right)}{p_s}\mathbf{e}_s \times \frac{d\mathbf{x}}{ds} - \frac{eB'\left(s\right)}{2p_s}\mathbf{e}_s \times \mathbf{x}. \tag{68}$$

The motion looks simpler in a special rotating frame. The coordinate vector $\mathbf{x}$ in the laboratory frame is related to the coordinate vector $\mathbf{x}_R$ in a frame rotating at a rate κ (Kim and Wang, 2000) as follows :

$$\mathbf{x} = M_\varphi \mathbf{x}_R, \quad M_\varphi = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix},$$
$$\text{with } \varphi(s) = \int_0^s \kappa(\tau)\, d\tau. \tag{69}$$

By differentiating, we obtain

$$\frac{d\mathbf{x}}{ds} = \left(\frac{d}{ds} M_\varphi \right) \mathbf{x}_R + M_\varphi \frac{d}{ds} \mathbf{x}_R. \tag{70}$$

The first term is

$$\left(\frac{d}{ds} M_\varphi \right) \mathbf{x}_R = -\kappa \begin{pmatrix} \sin(\varphi)\, x_R - \cos(\varphi)\, y_R \\ \cos(\varphi)\, x_R + \sin(\varphi)\, y_R \end{pmatrix}$$
$$= -\kappa M_\varphi \left(\mathbf{e}_s \times \mathbf{x}_R \right). \tag{71}$$

We thus obtain the well-known relationship between the space-frame (laboratory frame) and body-frame (moving with the body) differentiation (Goldstein, 2002):

$$\frac{d}{ds}\mathbf{x} = M_\varphi \left(\frac{d}{ds} - \kappa \mathbf{e}_s \times \right) \mathbf{x}_R. \tag{72}$$

The second derivative is

$$\frac{d^2\mathbf{x}}{ds^2} = M_\varphi \left(\frac{d}{ds} - \kappa \mathbf{e}_s \times \right) \left(\frac{d}{ds} - \kappa \mathbf{e}_s \times \right) \mathbf{x}_R. \tag{73}$$

Using Eq. (73) in Eq. (68), and choosing

$$\kappa(s) \to \kappa_L(s) \equiv \frac{eB(s)}{2p_s}, \tag{74}$$

we find that the equation of motion in the rotating frame becomes very simple:

$$\frac{d^2\mathbf{x}_R}{ds^2} = -\kappa_L^2(s)\, \mathbf{x}_R. \tag{75}$$

The spatial frequency κ_L given by Eq. (74), is one half of the cyclotron frequency, and is known as the *Larmor* frequency (Brillouin, 1945). In this frame the particle appears to be under a cylindrically symmetric focusing force. Equation (72) can also be written as

$$\frac{d}{ds}\mathbf{x}_R = R^{-1} \left(\frac{d}{ds} + \kappa_L \mathbf{e}_s \times \right) \mathbf{x}. \tag{76}$$

Noting that $\mathbf{e}_s \times \mathbf{x} = \rho \mathbf{e}_\varphi$ and in view of Eq. (66), we have

$$\mathbf{p}_R \equiv p_s \mathbf{x}' = R^{-1} \left(\mathbf{p} + e\mathbf{A} \right). \tag{77}$$

Therefore, the transverse momentum in the instantaneous rotating frame is simply the canonical momentum. The canonical angular momentum is

$$\boldsymbol{L} = p_s \mathbf{x}_R \times \mathbf{x}'_R = 2p_s \mathcal{L} \mathbf{e}_s. \tag{78}$$

The magnitude of $\boldsymbol{L}$ in laboratory frame quantities is

$$L = p_s \left(xy' - x'y + \kappa_L \left(x^2 + y^2 \right) \right) = \gamma m r^2 \dot{\varphi} + \frac{e}{2\pi} \Phi. \tag{79}$$

Here the term proportional to $\dot{\varphi} = d\varphi/dt$ is the kinetic part and the term proportional to Φ, which is the magnetic flux enclosed inside a circle of radius r at a given location in s, is the field part. L is constant due to the conservation of canonical angular momentum (Noether, 1971). This is also referred to as the *Busch's theorem* (Busch, 1926; Reiser, 1994). The conversion of field part to kinetic part occurs as the particle exits the solenoid when it receives an azimuthal kick from the radial magnetic field in the transition region.

The relation between the 4D phase-space vector in the rotating frame and that in the laboratory frame is

$$\begin{pmatrix} \mathbf{X}_R \\ \mathbf{Y}_R \end{pmatrix} = \begin{pmatrix} x_R \\ x'_R \\ y_R \\ y'_R \end{pmatrix} = R_B \begin{pmatrix} x \\ x' \\ y \\ y' \end{pmatrix}. \tag{80}$$

Here

$$R_B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & -\kappa_L & 0 \\ 0 & 0 & 1 & 0 \\ \kappa_L & 0 & 0 & 1 \end{pmatrix}. \tag{81}$$

Note M_B is *not* simplectic. However, its determinant is unity:

$$\det[R_B] = 1. \tag{82}$$

C. Beam transformation under external field

1. Liouville's theorem and manipulation of phase-space distribution

A beam consists of many particles and can be specified by its phase-space distribution function $f(\boldsymbol{\mathcal{Z}}; s)$. Let the phase-space volume Ω in $\boldsymbol{\mathcal{Z}}$ at s_1 transform to $\bar{\Omega}$ in $\bar{\boldsymbol{\mathcal{Z}}}$ at s_2. Here, we are going back to the general nonlinear map M, Eq.(19). Assuming that there are no obstructions leading to particle loss, the distribution function can be normalized to unity as the beam goes through an accelerator beamline:

$$\int_\Omega d^6\boldsymbol{\mathcal{Z}} f(\mathcal{Z}, s_1) = \int_{\bar{\Omega}} d^6\bar{\boldsymbol{\mathcal{Z}}} f(\bar{\boldsymbol{\mathcal{Z}}}, s_2) = 1. \tag{83}$$

The left-hand side of Eq. (83) can, noting Eq. (23), be written as

$$\int_{\bar{\Omega}} d^6\bar{\boldsymbol{\mathcal{Z}}} \det\left(R^{-1}\right) f\left(M^{-1}\left(\bar{\boldsymbol{\mathcal{Z}}}\right), s_1\right)$$
$$= \int_{\bar{\Omega}} d^6\bar{\mathcal{Z}} f\left(M^{-1}\left(\bar{\boldsymbol{\mathcal{Z}}}\right), s_1\right). \tag{84}$$

Here R is the Jacobian matrix given by Eq.(21). Comparing this with RHS of Eq. (83), we find that the transformation of the distribution function corresponding to the transform of the phase-space variable in Eq. (19):

$$f\left(\boldsymbol{\mathcal{Z}}, s_2\right) = f\left(M^{-1}\left(\boldsymbol{\mathcal{Z}}\right), s_1\right). \tag{85}$$

This is the celebrated Liouville's theorem (Liouville, 1838), fundamental to many beam shaping schemes involving external fields.

Section III discusses how $f\left(\boldsymbol{\mathcal{Z}}; s\right)$ at the cathode leads to different distribution subsequently. If the 6D distribution function $f\left(\boldsymbol{\mathcal{Z}}, s\right)$ is integrated over all components of $\boldsymbol{\mathcal{Z}}$ except the variable ζ_j, we obtain the 1D density distribution in ζ_j. Section IV discusses various ways of obtaining interesting 1D distributions by suitable map M.

We now return to the case where the map is linear and can be represented by matrix R.

To give a simple example, consider the distribution in the (z, δ) subspace:

$$f\left(z, \delta; s_1\right) = \frac{1}{\ell\sqrt{2\pi}\sigma_\delta} \exp\left[-\frac{\delta^2}{2\sigma_\delta{}^2}\right]. \tag{86}$$

Here ℓ is the bunch length, over which the distribution is regarded as constant.

The corresponding line density in z is constant:

$$\lambda\left(z; s_1\right) = \int d\delta f\left(z, \delta; s_1\right) = \frac{1}{\ell}. \tag{87}$$

The matrix for the bunch compression section consisting of chirping (energy change linear in z) followed by an R_{56} is

$$R_C = \begin{pmatrix} 1 & R_{56} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ h & 1 \end{pmatrix} = \begin{pmatrix} 1 + hR_{56} & R_{56} \\ h & 1 \end{pmatrix}. \tag{88}$$

Its inverse is

$$R_C{}^{-1} \begin{pmatrix} z \\ \delta \end{pmatrix} = \begin{pmatrix} 1 & -R_{56} \\ -h & 1 + hR_{56} \end{pmatrix} \begin{pmatrix} z \\ \delta \end{pmatrix}$$
$$= \begin{pmatrix} z - R_{56}\delta \\ -hz + \delta/C \end{pmatrix}. \tag{89}$$

Here we introduced the compression factor C:

$$C = 1/\left(1 + hR_{56}\right). \tag{90}$$

Therefore, the distribution after the beamline, applying Eq. (85), becomes

$$f\left(z, \delta; s_2\right) = \frac{1}{\ell\sqrt{2\pi}\sigma_\delta} \exp\left[-\frac{\left(-hz + \delta/C\right)^2}{2\sigma_\delta^2}\right]. \tag{91}$$

The corresponding line density is

$$\lambda\left(z; s_2\right) = \int d\delta f\left(z, \delta; s_2\right) = \frac{C}{\ell}. \tag{92}$$

The line density is increased by C since the bunch length is compressed by the same factor.

2. Beam matrix and emittance

A beam can also be completely specified by all of its beam moments, which are:

$$\langle\zeta_i\zeta_j\zeta_k...\rangle_s = \int d^6\boldsymbol{\mathcal{Z}}\, f\left(\mathcal{Z}; s\right) \zeta_i\zeta_j\zeta_k..., \tag{93}$$

with $i, j, k, .. = 1, 2, ..6$. The moments transform accordingly to

$$\langle\zeta_{i2}\zeta_{j2}\zeta_{k2}..\rangle_{s2} = R_{i2,i1}R_{j2,j1}R_{k2,k1}.. \langle\zeta_{i1}\zeta_{j1}\zeta_{k1}..\rangle_{s1}. \tag{94}$$

The first order moments can be made to vanish noting that ζ's are deviation variable, that is variables relative to the trajectory of the reference particle. The reference particle is at the "center" of the beam in the sense that the first order moments of the deviation variables vanish:

$$\langle\zeta_i\rangle = \int d^6\boldsymbol{\mathcal{Z}}\, \zeta_i f(\boldsymbol{\mathcal{Z}}, s) = 0. \tag{95}$$

The second order moments are

$$\Sigma = \langle\boldsymbol{\mathcal{Z}}\boldsymbol{\mathcal{Z}}^T\rangle;\ \Sigma_{ij} = \int d^6\boldsymbol{\mathcal{Z}}\, f\left(\mathcal{Z}\right)\ \zeta_i\zeta_j. \tag{96}$$

Since $f\left(\boldsymbol{\mathcal{Z}}\right) \geq 0$ and cannot vanish identically, the *beam matrix* Σ is symmetric and positive definite.

Note that the R matrix is symplectic, satisfying Eq. (22). The first moments vanish by suitably choosing the coordinate frame. The second moments are elements of the beam matrix, which can be written as

$$\Sigma = \langle\boldsymbol{\mathcal{Z}}\boldsymbol{\mathcal{Z}}^T\rangle = \begin{pmatrix} \langle\mathbf{X}\mathbf{X}^T\rangle & \langle\mathbf{X}\mathbf{Y}^T\rangle & \langle\mathbf{X}\mathbf{Z}^T\rangle \\ \langle\mathbf{Y}\mathbf{X}^T\rangle & \langle\mathbf{Y}\mathbf{Y}^T\rangle & \langle\mathbf{Y}\mathbf{Z}^T\rangle \\ \langle\mathbf{Z}\mathbf{X}^T\rangle & \langle\mathbf{Z}\mathbf{Y}^T\rangle & \langle\mathbf{Z}\mathbf{Z}^T\rangle \end{pmatrix}. \tag{97}$$

Here $\langle XX^T\rangle$, etc., are 2×2 matrices and $\langle\rangle$ is the averaging operation as in Eq. (93). The beam matrix transforms as

$$\Sigma(s_2) = R\Sigma(s_1)R^T. \tag{98}$$

We introduce the quantities called *projected emittances* in each $\mathbf{X}, \mathbf{Y}, \mathbf{Z}$ subspace:

$$\varepsilon_x^{proj} = \sqrt{\det\left(\langle\mathbf{X}\mathbf{X}^T\rangle\right)} = \sqrt{\langle x^2\rangle\langle x'^2\rangle - \langle xx'\rangle^2},$$
$$\varepsilon_y^{proj} = \sqrt{\langle y^2\rangle\langle y'^2\rangle - \langle yy'\rangle^2}, \tag{99}$$
$$\varepsilon_z^{proj} = \sqrt{\langle z^2\rangle\langle \delta^2\rangle - \langle z\delta\rangle^2}.$$

The projected emittances are in general not invariant.

Williamson (Williamson, 1936) has proved that a positive definite, symmetric matrix such as Σ can be transformed to a 2×2 block diagonal form by a *symplectic* matrix A:

$$A\Sigma A^T = \begin{pmatrix} \langle\mathbf{X}_1\mathbf{X}_1{}^T\rangle & 0 & 0 \\ 0 & \langle\mathbf{X}_2\mathbf{X}_2{}^T\rangle & 0 \\ 0 & 0 & \langle\mathbf{X}_3\mathbf{X}_3{}^T\rangle \end{pmatrix}. \tag{100}$$

Here the 2-vectors $\mathbf{X}_j$, $j=1,2,3$ represent a new partition of the 6D space into three, decoupled, 2D phase spaces. The emittances in each decoupled space

$$\varepsilon_j = \sqrt{\det \langle \mathbf{X}_j \mathbf{X}_j{}^T \rangle}. \qquad (101)$$

will be referred to as the *principal emittance*[2]. The elements of the 2×2 beam matrix can be parametrized as follows:

$$\langle \mathbf{X}_j \mathbf{X}_j{}^T \rangle \equiv \begin{pmatrix} \langle x_j^2 \rangle & \langle x_j x_j' \rangle \\ \langle x_j x_j' \rangle & \langle x_j'^2 \rangle \end{pmatrix} = \varepsilon_j \begin{pmatrix} \beta_j & -\alpha_j \\ -\alpha_j & \gamma_j \end{pmatrix}. \qquad (102)$$

From Eq.(101), it follows that the parameters are related by $\beta_j \gamma_j - \alpha_j^2 = 1$. Note that this relation is the same as the second of Eq.(31). The parametrization here is directly connected to that introduced in the Courant-Snyder form of betatron motion (29) (identifying the y-subspace with the j-subspace); If we compute the beam matrix using Eq.(29) as the beam trajectory, assuming that the betatron phase ψ_{j0} is uniformly distributed, then we find that the principal emittance in jth subspace is the statistical average of particles' action variable in that subspace:

$$\varepsilon_j = \langle J_j \rangle. \qquad (103)$$

Therefore, the parameters α_j, β_j and γ_j here are in fact identical to those introduced in Eq.(29).

The beam matrix Eq.(102) can be diagonalized with the matrix

$$B^j = \begin{pmatrix} 1 & -\alpha_j/\gamma_j \\ 0 & 1 \end{pmatrix}. \qquad (104)$$

Indeed

$$B_j \langle \mathbf{X}_j \mathbf{X}_j{}^T \rangle B_j{}^T = \begin{pmatrix} \varepsilon_j \beta_j{}^* & 0 \\ 0 & \varepsilon_j/\beta_j{}^* \end{pmatrix}. \qquad (105)$$

Here we used the notation

$$\beta_j{}^* \equiv 1/\gamma_j. \qquad (106)$$

The transformation represented by matrix B^j is a free-space translation going back a distance $\ell = \alpha_j/\gamma_j$ to the *waist* location of the beam where the correlation α_j vanishes. In this sense, $\beta_j{}^*$ is known as the beta function at the waist. Since angular divergence does not change under a free-space translation, the value of γ_j does not change while translating to the waist. Note the second of Eq.(31) can be written as

$$\beta_j = \left(1 + \alpha_j{}^2\right)/\gamma_j = \beta_j{}^* \left(1 + (\ell/\beta_j{}^*)^2\right) \qquad (107)$$

—————

[2] In reference (Dragt, 2011), this emittance is referred to as *eigen-emittance*.

This is the well-known and useful equation describing how the beta function changes away from the waist.

Performing diagonalization in other dimensions, we obtain

$$BA\Sigma A^T B^T =$$
$$\mathrm{diag}\left(\varepsilon_1 \beta_1{}^*, \varepsilon_1/\beta_1{}^*, \varepsilon_2 \beta_2{}^*, \varepsilon_2/\beta_2{}^*, \varepsilon_3 \beta_3{}^*, \varepsilon_3/\beta_3{}^*\right). \qquad (108)$$

The beta functions at the waist, $\beta_j{}^*$s, are also positive but not invariant. With an additional transformation, Eq. (108) can be reduced to Williamson's normal form $\mathrm{diag}\left(\varepsilon_1, \varepsilon_1, \varepsilon_2, \varepsilon_2, \varepsilon_3, \varepsilon_3\right)$. However, this last transformation is not physical since it changes the dimensions of the elements.

The principal emittance ε_j is invariant under any transformation that leaves the subspace j intact, *if there is no acceleration*. If $j = x$, a translation along z-axis is an example of such transformation. If particles are accelerated in the z direction, for example, then $x' = dx/ds$ is no longer a canonical variable, and we need to start from the correct canonical variable $p_x = m\gamma dx/dt = mc\gamma\beta x'$ and its conjugate x. When the particle velocities are nearly the same, the emittance that is invariant under acceleration also is

$$\varepsilon_j{}^n = \beta\gamma\varepsilon_j. \qquad (109)$$

Note that β and γ in the above equation are the velocity in the z-direction divided by c and the particle energy divide by mc^2, *not* the Courant-Snyder amplitudes. The emittance defined by Eq.(109) is referred to as *normalized* emittance, while that defined by Eq. (101) is known as the *un-normalized* emittance. The emittances in this review are un-normalized emittances unless specified otherwise.

The principal emittances can be found by adopting the standard eigenvalue problem (Dragt, 2011). They can also be obtained as follows (Courant, 1966; Neri and Rangarajan, 1990). First, note that the quantities

$$\Gamma^n = \mathrm{Tr}[(J\Sigma)^{2n}], \quad n = 1, 2, 3.. \qquad (110)$$

are, in view of Eq. (22), invariant under symplectic transformation. Note also that Σ inside the trace of Eq.(110) can be replaced by the diagonal matrix Eq.(108). Thus, we obtain

$$\Gamma^n = 2(-1)^n \left((\varepsilon_1)^{2n} + (\varepsilon_2)^{2n} + (\varepsilon_3)^{2n}\right). \qquad (111)$$

The above gives three equations for three principal emittances. We also have

$$\det\left(\Sigma\right) = \left(\varepsilon_1 \varepsilon_2 \varepsilon_3\right)^2. \qquad (112)$$

Any of the three equations from Eqs. (111) and (112) can be solved for the three principal emittances.

The square root of Eq.(112), or the product of three principal emittances, are known as the 3D emittance.

3. Emittance exchange (EEX) and phase-space exchange (PSE)

An important corollary of Williamson's theorem is that *an arbitrary re-partition of the emittances is not possible by means of symplectic transformation.* Thus, for example, we exchange $(\varepsilon_x, \varepsilon_y, \varepsilon_z)$ to $(\varepsilon_z, \varepsilon_y, \varepsilon_x)$ but cannot re-partition it to $(\varepsilon_x/4, 4\varepsilon_y, \varepsilon_z)$. This fact appears to be first noted in the accelerator physics context by Courant (Courant, 1966).

Although the set of three principal emittances does not change, the ordering in the set could be changed. For example, the emittance in $\mathbf{X}$ space can be exchanged to the emittance in $\mathbf{Z}$ space. To explain the meaning of the emittance exchange (EEX), consider a beam matrix of the form:

$$\Sigma_I = \begin{pmatrix} \Sigma_x & 0 \\ 0 & \Sigma_z \end{pmatrix};$$

$$\Sigma_x = \begin{pmatrix} \langle x^2 \rangle & \langle xx' \rangle \\ \langle xx' \rangle & \langle x'^2 \rangle \end{pmatrix}, \Sigma_z = \begin{pmatrix} \langle z^2 \rangle & \langle z\delta \rangle \\ \langle z\delta \rangle & \langle \delta^2 \rangle \end{pmatrix}. \quad (113)$$

The matrix Σ_I is uncoupled, that is, the off-diagonal blocks vanish. Therefore, the determinants of the 2×2 matrices Σ_x and Σ_z are, respectively, $\varepsilon_x{}^2$ and $\varepsilon_z{}^2$, where ε_x and ε_z are the principal emittances . Suppose a beamline gives rise to the following transformation:

$$R \begin{pmatrix} \Sigma_x & 0 \\ 0 & \Sigma_z \end{pmatrix} R^T = \begin{pmatrix} \Sigma_z{}' & 0 \\ 0 & \Sigma_x{}' \end{pmatrix}. \quad (114)$$

Here, $\Sigma_x{}'$ is a 2×2 beam matrix symplectically connected to Σ_x and similarly $\Sigma_z{}'$ to Σ_z. It can be shown that R is of the following form:

$$R = \begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix}, \quad (115)$$

where B and C are 2×2 symplectic matrices. Thus

$$R \begin{pmatrix} \mathbf{X} \\ \mathbf{Z} \end{pmatrix} = \begin{pmatrix} B\mathbf{Z} \\ C\mathbf{X} \end{pmatrix}. \quad (116)$$

Therefore, an EEX transformation is much more than an exchange of the magnitude of the sub-space area—it exchanges the whole sub-space $\mathbf{X}$ with the whole sub-space $\mathbf{Z}$. A phase-space shape in $\mathbf{X}$ will be transformed to a corresponding shape in $\mathbf{Z}$, making the exchange transformation useful for beam shaping purposes. Although the transformation was named EEX since it was considered in the context of the emittance exchange (Cornacchia and Emma, 2002; Emma *et al.*, 2006; Kim and Sessler, 2006), it would be more proper to refer to it as *phase-space exchange*(PSE).

Examples of emittance and phase-space exchange methods are discussed in Sections VI.C.1 and VI.C.2 , respectively.

4. Emittance repartitioning

Although emittances can only be exchanged wholly in a symplectic transformation, an emittance *re-partitioning* is possible if a non-symplectic step occurs at some point during the transformation. An example is provided by a beam produced from a photocathode immersed in an axial magnetic field B, as discussed in Section II.B.4. We construct the beam matrix of the rotating frame-vector Eq. (80):

$$\Sigma_B = \begin{pmatrix} \left\langle \mathbf{X}_R \mathbf{X}_R{}^T \right\rangle & \left\langle \mathbf{X}_R \mathbf{Y}_R{}^T \right\rangle \\ \left\langle \mathbf{Y}_R \mathbf{X}_R{}^T \right\rangle & \left\langle \mathbf{Y}_R \mathbf{Y}_R{}^T \right\rangle \end{pmatrix}. \quad (117)$$

The subscript B indicates the presence of the magnetic field. Let ε_{R1} and ε_{R2} be the two principal emittances associated with the beam matrix Σ_B, which will in general be different from the principal emittance of the beam in the absence of the magnetic field, ε_{01} and ε_{02}. In view of Eq. (82), we have

$$\det \left[\Sigma_B \right] = \det \left[R_B \Sigma_0 R_B{}^T \right] = \det \left[\Sigma_0 \right]. \quad (118)$$

Here, Σ_0 is the beam matrix for $B = 0$. Thus,

$$\varepsilon_{B1}\varepsilon_{B2} = \varepsilon_{01}\varepsilon_{02}. \quad (119)$$

Invoking the Busch's theorem for beam, it is shown in Section VI that the emittances ε_{B1} and ε_{B2} can be determined and their ratio arbitrarily adjusted by varying the magnetic field . Is this a violation of the Williamson-Courant theorem? No, since the matrix R_B in Eq. (81) is not symplectic. Since the beam is really born in a magnetic field with the beam matrix $R_B \Sigma_0 R_B{}^T$, one may object calling this example an emittance re-partition. Indeed, if the beam is first produced from a cathode in a field-free region and then encounters an axial magnetic field, its principal emittances will not change. Emittance re-partitioning schemes *always* involve non-symplectic elements, such as beam masks or tapered absorbing blocks. A generalization of Busch's theorem to a non-cylindrically symmetric system was discussed by (Groening *et al.*, 2018). A general emittance re-partitioning was investigated in Ref. (Carlsten *et al.*, 2011a).

Details of experimental setup and results are discussed later in Sections VI.B.1 and VI.B.2.

5. Nonlinear case

If the Hamiltonian contains polynomials of order higher than quadratic in the scaled deviation variables, the variables in Eq. (14) are not canonical and the map M in Eq. (19) becomes nonlinear. If the nonlinearity is small, the canonical variables Eq. (8) can be expressed in Taylor series in $\boldsymbol{\mathcal{Z}}$. The map from $\boldsymbol{\mathcal{Z}} \to \bar{\boldsymbol{\mathcal{Z}}}$ can be found by solving the equation of motion using the original Hamiltonian Eq. (9) in the following form:

$$\bar{\zeta}_i = R_{ij}\zeta_j + T_{ijk}\zeta_j\zeta_k + ... \quad (120)$$

The coefficients $T_{ijk},..$ were worked out in detail in (Brown, 1968) (However, see footnotes 1 for the sign convention). The constraints on these coefficients from symplectic property was discussed in (Wollnik and Berz, 1985).

Higher order solutions can be obtained in classical mechanics by canonical perturbation theory, which provides a procedure for finding canonical transformations in which the new canonical momenta become constants of motion (Arnold, 1978). The procedure was applied to accelerator beam dynamics, see for example (Ruth, 1986). A powerful method using Lie canonical transformation with polynomial generators has been developed that can handle very high order polynomial terms by numerical computation (Dragt, 2011).

If the transformation is nonlinear, Liouville's theorem still applies microscopically. However, there can be an apparent increase in the macroscopic phase-space volume due to filamentation (Sørensen, 1988). For a weakly nonlinear system, one can still introduce adiabatic invariants, which is the phase-space area following the physical orbit (Landau and Lifshitz, 1969).

D. Beam-generated fields

1. Wakefield and impedance

A particle moving uniformly at a speed βc, $\beta \approx 1$ carries a Coulomb field with it, squeezed to an angular width of $\sim 1/\gamma = \sqrt{1 - \beta^2}$ due to Lorentz contraction. If there is a surface parallel to the particle trajectory at a distance b, a moving area of longitudinal length b/γ is under the influence of the Coulomb field. If the surface is perfectly conducting and smooth, the boundary condition at the surface will be maintained. If there are interruptions in the surface , however, then the field interacts with the surface and produces fields behind the particle (thus, the term wake) that can influence the motion of the trailing particles.

The longitudinal wakefield $w(z)$ is defined as the EM field E_s on a test particle trailing a fixed distance z behind the drive particle (Chao, 1993; Heifets and Kheifets, 1991; Stupakov, 2001; Wilson, 1989):

$$w(z) = -\frac{1}{e} \left. E_s \right|_{z=s-ct}. \tag{121}$$

The minus sign in the above is to make a positive wake that corresponds to the test particle losing its energy. The sign of z is that it is ahead if positive. The dependence on the transverse coordinates is not important in most of the following and is thus neglected. The unit of the wake field is V per C per m. In the above, we are considering the case where the wake is uniform along the beam chamber. When the wake is localized, the wake function is defined by the integral over the passage of the local structure.

The energy loss per unit distance of a particle at position z due to other particles is[3]

$$\Delta E(z) = eQ \int_{-\infty}^{\infty} dz' \, w(z - z') \lambda(z'). \tag{122}$$

Here Q is the total charge and $\lambda(z)$ is the line-charge density of the particles in the beam normalized as $\int dz \lambda(z) = 1$.

Now we introduce the impedance per unit length and Fourier transform of the electric field and current profile as follows (Chao, 1993; Nielsen et al., 1959):

$$Z(k) = \frac{1}{c} \int_{-\infty}^{\infty} ds \, w(z) e^{-ikz}, \tag{123}$$

$$\tilde{E}(k) = \int_{-\infty}^{\infty} dz \, e^{-ikz} E(z), \tag{124}$$

$$I(k) = cQ \int_{-\infty}^{\infty} dz \, e^{-ikz} \lambda(z). \tag{125}$$

From Eq. (122) these quantities are related via

$$\tilde{E}(k) \equiv V(k) = -Z(k) I(k), \tag{126}$$

where $V(k)$ is the (negative) voltage applied per unit distance. In Eq. (87) the line density $\lambda(z)$ was given as an integral of the phase space distribution keeping only the **Z**. In the more general case, we should write

$$\lambda(z; s) = \int d\delta d^2\mathbf{X} d^2\mathbf{Y} f(\boldsymbol{\mathcal{Z}}; s). \tag{127}$$

In discussing coherent instabilities including free-electron lasers, one often uses the term bunching factor $b(k; s) = I(k)/cQ$. Note that $I(k)$ or $b(k)$ can be expressed as an integral in 6D phase space $\boldsymbol{\mathcal{Z}}$:

$$b(k; s) = \int dz e^{-ikz} \lambda(z; s) = \int d^6 \boldsymbol{\mathcal{Z}} e^{-ikz} f(\boldsymbol{\mathcal{Z}}; s). \tag{128}$$

The collective force due to a beam-generated field on a particle at z will change the longitudinal momentum and hence δ:

$$\frac{d\delta_{coll}}{ds} = eE(z) = -\frac{r_e}{\gamma} N \int dk e^{ikz} Z(k) b(z; s). \tag{129}$$

Here r_e is the the classical electron radius, and N is the total number of particles. The subscript coll is to emphasize that this is the beam-generated, collective force.

The meaning of equations(126) and (129) is that a current modulation at spatial frequency k impresses an energy modulation via the impedance $Z(k)$. Passing through the subsequent beamline, energy modulation can cause beam instabilities, as discussed in Section II.D.5 .

Next, we discuss three representative cases of impedance.

————

[3] Note that $w(z - z')$ is written as $w(z' - z)$ in some references, e.g., (Chao, 1993; Stupakov, 2001). Our choice is convenient since it ensures the same Fourier-transform convention for impedance as well as current.

2. Fields due to boundary perturbation

Since wakefields originate from the EM field in the moving disc scattered off by the surface interruptions, a wakefield satisfies the causality condition (Wilson, 1989)

$$w(z) = 0; \quad z > 0. \tag{130}$$

A wakefield device, referred to as a *dechirper*, with corrugated walls can be useful in correcting the energy chirp that may arise while compressing a bunch for high peak current (Bane and Stupakov, 2016, 2012; Deng *et al.*, 2014; Emma *et al.*, 2014b). A flat dechirper consisting of two opposing corrugated plates separated by a half gap a has been shown to be effective in removing the chirp in the bunches driving an X-ray FEL oscillator (Qin *et al.*, 2016). The wakefield of such a device was computed and can be represented approximately by the following form:

$$w(z) \approx \frac{\pi c Z_0}{16a^2} \quad \text{for } z < 0; \quad = 0 \quad \text{for } z > 0, \tag{131}$$

with $Z_0 = 377 \ \Omega$ being the free-space impedance. For a flat-top charge density, we see readily from Eq. (122) that the energy loss is linear in z, and the particles in the tail losing more energy than those at the head. The difference in energy correction from head to tail is found to be

$$\Delta E(z) = -\frac{\pi c Z_0 QL}{16a^2}. \tag{132}$$

The impedance per unit length corresponding to Eq. (131) can be computed from Eq. (123):

$$Z_{flat}(k) = i\frac{\pi Z_0}{16a^2 k}. \tag{133}$$

We note here that the impedance of a round pipe of radius a with random or periodic corrugations in the high frequency limit is given by (Gluckstern, 1989)

$$Z_{round}(k) = i\frac{Z_0}{\pi a^2 k}. \tag{134}$$

Impedances due to the interruption of perfectly conducting walls are sometimes referred to as *geometric* impedances.

3. Space-charge force

The longitudinal space-charge wake was computed for a bunch moving between two parallel conductors (Nielsen and Sessler, 1959) and moving inside a pipe (Neil and Sessler, 1965) with the following approximate method:

Recall the variable $z = s - c\beta t \approx s - ct$. Consider a beam of uniform charge density $\lambda(z) = $ constant and a uniform cross section, of radius b, travelling at velocity

βc, $\beta \approx 1$ along the axis of a circular, perfectly conducting pipe of radius a. The non-vanishing EM field components in cylindrical coordinates are the radial electric field E_r and the azimuthal magnetic field B_φ given by

$$r \leq b : E_r(r,z) = \frac{Q\lambda(z)r}{2\varepsilon_0 b^2}, B_\varphi(r,z) = \frac{\beta Q\lambda(z)r}{2\varepsilon_0 b^2 c},$$
$$b < r < a : E_r(r,z) = \frac{Q\lambda(z)}{2\varepsilon_0 r}, B_\varphi(r,z) = \frac{\beta Q\lambda(z)}{2\varepsilon_0 rc} \tag{135}$$

Now consider the density has a variation so that the z-dependence of $\lambda(z)$ needs to be taken account. If the dependence on z is slow, the main fields are still given by Eq. (135), but, in addition, there will be a longitudinal electric field on axis $E_z(0,z)$. This can be determined from Faraday's law: The line integral of electric field along the loop consists of straight lines connecting the points in cylindrical coordinates, $(0,z) \to (0,z+\Delta z) \to (a,z+\Delta z) \to (a,z) \to (0,z)$ should be equal to the time derivative of the magnetic flux into the loop. We obtain in this way

$$E_z(0,z) = -\frac{Q}{2\varepsilon_0\gamma^2}\left(1 + 2\log\frac{a}{b}\right)\frac{d\lambda}{dz}. \tag{136}$$

According to Eq. (136), the space-charge field tends to smooth away a density bump, as the particles repel each other. This is dramatically illustrated by Fig. 28 (b) in Section V.A.1. If there is a periodic modulation in the density, the other hand, then the modulation could become enhanced, as discussed in SectionV.A.3.

Equation (136) diverges logarithmically as $a \to \infty$. However, the derivation is valid only if the density bump is not too steep, implying [see problem of 1.5 of (Chao, 1993)]

$$a < \gamma\Delta z, \tag{137}$$

where Δz is the extent of the density variation. The corresponding impedance is

$$Z_{SC} = i\frac{Z_0 k}{4\pi\gamma^2}\left(1 + 2\log\left[\frac{a}{b}\right]\right). \tag{138}$$

In the absence of a vacuum chamber pipe, the inequality Eq. (137) is violated. In this case, one can use the squeezed electric field due to a single particle moving with uniform velocity $\beta c \approx c$ parallel to the z-axis (Jackson, 1998) as the Green's function, perform the Fourier transform in z variable, and observe that only the 0-th component in the azimuthal series will contribute. The impedance per unit length can then be computed with the result (Rosenzweig *et al.*, 1997; Venturini, 2008):

$$Z_{SC}(k) = i\frac{Z_0}{\pi b^2 k}\left(1 - \frac{kb}{\gamma}K_1\left(\frac{kb}{\gamma}\right)\right). \tag{139}$$

Expanding $K_1 (kb/\gamma)$ for small kb/γ (Huang *et al.*, 2005; Venturini, 2008):

$$Z_{SC} (k) = i\frac{Z_0 k}{4\pi\gamma^2} \left(2\log\left[\frac{\gamma}{kb}\right] + 2\log 2 - 2\gamma_E + 1 \right)$$
$$\approx i\frac{Z_0 k}{4\pi\gamma^2} \left(2\log\left[\frac{\gamma}{kb}\right] + 1.23 \right), \tag{140}$$

where $\gamma_E = 0.577$ is the Euler number.

4. Coherent synchrotron radiation

A particle on a circular path emits synchrotron radiation. The radiation from the tail of a bunch, proceeding in a path tangential to the circular path, can exert a force to the head of the bunch moving on an arc. Making use of the results known in 1912 (Schott, 1912), the wakefield at a distance s *ahead* of the emitting particle (note s is negative) can be written as follows (Derbenev *et al.*, 1995; Murphy *et al.*, 1997; Saldin *et al.*, 1997):

$$E_\parallel(z) = -\frac{e}{4\pi\varepsilon_0} \frac{2}{3^{1/3}\rho^{2/3}} \frac{d}{ds} G(z), \tag{141}$$
$$\text{with } G(z) = \frac{1}{(z)^{1/3}}, \text{for } z \gg \frac{2\rho}{3\gamma^3}.$$

In the above equation ρ is the radius of curvature. The function $G(s)$ vanishes for $s \geq 0$. The force due to density distribution $\lambda(s)$ will then be

$$E_{\parallel,\lambda}(z) = \frac{e}{2\pi\varepsilon_0} \frac{1}{3^{1/3}\rho^{2/3}} \int_{-\infty}^{s} dz' \frac{1}{(z-z')^{1/3}} \frac{d}{dz'} \lambda(z'). \tag{142}$$

In the above equation we used integration by parts and also assumed that $|s - s'|$ is about the bunch length, which is much greater than ρ/γ^3. The impedance associated with this force is referred to as *coherent synchrotron radiation* (CSR) impedance:

$$Z_{CSR}(k) = \frac{Z_0}{2\pi 3^{1/3}\rho^{2/3}} \int_0^\infty \frac{1}{(z)^{1/3}} e^{-ikz} dz$$
$$= \frac{Z_0 k^{1/3}}{2\pi 3^{1/3}\rho^{2/3}} e^{i\pi/6} \Gamma(2/3)$$
$$\approx (1.63 + 0.94i) \frac{Z_0 k^{1/3}}{4\pi\rho^{2/3}}. \tag{143}$$

The CSR impedance was first identified in connection with storage ring physics, in which the mode number $n = k\rho$, where ρ is the radius of the curvature (Faltens and Laslett, 1975; Iogansen and Rabinovich, 1960). The CSR impedance can be significant when bending magnets are used in linacs, e.g., in the final chicane magnet of a bunch compressor where the current is high (Heifets *et al.*, 2002; Huang and Kim, 2002).

Synchrotron radiation has also high frequency incoherent part, referred to as incoherent synchrotron radiation (ISR) (Saldin *et al.*, 1996; Sands, 1955, 1969). However, the ISR effects that may reduce shaping accuracy are not significant for low energy electron beams for shaping.

5. Collective motion

We now discuss how the force due to the beam-generated fields acting back on the beam, limiting to the 4D phase space $\mathbf{Z} = (\mathbf{X}, \mathbf{Z})$ for simplicity. The first question to ask is whether the Liouville's theorem, Eq.(85), is still valid in the presence of the beam-generated field. The answer is yes, as long as we can neglect the discrete particle aspects and regard the beam as a continuous fluid. The fluid description is valid if the Debye length λ_D is much shorter than the length scale of the collective disturbance:

$$k\lambda_D \ll 1, \tag{144}$$

where k is the wave number of the disturbance. The criteria in Eq. (144) were first derived for non-relativistic plasma (Pines and Bohm, 1952) and extended to relativistic beams (Kim and Lindberg, 2011; Rosenzweig *et al.*, 1997; Sørensen, 1988). For beam physics, we define the Debye length as the transverse spread of the particles during one plasma oscillation:

$$\lambda_D = \frac{c\sigma_{\Delta\beta}}{\omega_p}, \tag{145}$$

where $c\sigma_{\Delta_\beta}$ is the RMS velocity spread and

$$\omega_p = \sqrt{\frac{e^2 n}{\varepsilon_0 m\beta\gamma^3}}. \tag{146}$$

is the plasma frequency for a relativistic beam. The basic process behind collective motion is as follows: A region in phase space may develop higher spatial density when the beam goes through some part of the beamline, such as a compressor. This part of the phase space then exerts a collective force via Eq. (129). This force is added to the external force to modify the beam evolution. A classic example is the plasma oscillation when the density of a part of the beam is increased at the expense of another part. These two parts then oscillate against each other with the plasma frequency given by Eq. (146). Collective motion is often detrimental for beam shaping, but it can be useful in some particular cases, such as generation of ultrashort bunch trains via nonlinear plasma oscillation, as will be discussed in Section V.A.3.

Let's now discuss how a beam-generated field affects the evolution of the phase-space distribution $f(\mathbf{Z}; s)$ in a beamline beginning at $s = 0$ (Heifets *et al.*, 2002; Huang and Kim, 2002). In the absence of a beam-generated field, the beam distribution functions f_0 at two different locations are related by the Liouville's theorem

$$f_0(\mathbf{Z}_s; s) = f_0(\mathbf{Z}_\tau; \tau); \mathbf{Z}_\tau = R_{s\tau}^{-1}\mathbf{Z}_s. \tag{147}$$

This is the same as Eq. (85), $R_{\tau s}$ being a combination of any transformation matrix discussed in Section II.B. For notational clarity, a variable at location τ was given the same subscript, e.g., $\boldsymbol{\mathcal{Z}}_\tau$. Let $d\delta_{coll}(z_\tau;\tau)$ be the increase in the electron's relative energy due to the beam-generated field in a small interval between τ and $\tau + d\tau$. We obtain

$$f(\boldsymbol{\mathcal{Z}}_s, s) = f_0(\boldsymbol{\mathcal{Z}}_s; s) - \int_0^s d\tau \frac{\partial f_0(\boldsymbol{\mathcal{Z}}_\tau; \tau)}{\partial \delta_\tau} \frac{d\delta_{coll}(z_\tau; \tau)}{d\tau}. \tag{148}$$

From this equation the bunching factor at s, Eq.(128), can be obtained after some mathematical manipulation (Huang and Kim, 2002):

$$b(k, s) = b_0(k, s) + \\ \frac{ikr_e}{\gamma} \int d\tau R_{56,s\tau} \int \frac{dk_1}{2\pi} Z(k_1; \tau) b(k_1; \tau) \\ \times \int d\boldsymbol{\mathcal{Z}}_0 e^{-ikz_s(\boldsymbol{\mathcal{Z}}_0) + ik_1 z_\tau(\boldsymbol{\mathcal{Z}}_0)} f_0(\boldsymbol{\mathcal{Z}}_0). \tag{149}$$

Here

$$z_s(\boldsymbol{\mathcal{Z}}_\tau) = \sum_{j=1}^6 R_{5j,s\tau} \zeta_{j,\tau} \\ = z_\tau + R_{51,s\tau} x_\tau + R_{52,s\tau} x'_\tau + R_{56,s\tau} \delta_\tau \tag{150}$$

In the first part, $\zeta_{j,\tau}$ denote the components of $\boldsymbol{\mathcal{Z}}_\tau$, see Eq. (14). The second part of the above equation is for the case $\boldsymbol{\mathcal{Z}} = (\mathbf{X}, \mathbf{Z})$. Equation (149) is an integral equation for the evolution of the bunching factor. We consider the initial distribution in the following form:

$$f_0(\boldsymbol{\mathcal{Z}}_0) = \bar{f}_0(\boldsymbol{\mathcal{Z}}_0) + \hat{f}_0(\boldsymbol{\mathcal{Z}}_0). \tag{151}$$

The first part $\bar{f}_0$ is smooth and the second part $\hat{f}_0$ contains high-frequency modulation giving rise to the initial bunching factor b_0. This term is regarded to be small. Therefore, f_0 in Eq. (149) can be replaced by $\bar{f}_0$. Equation (149) can be solved iteratively.

To see the physical meaning of Eq. (149), we neglect the ($\mathbf{X}$) and ($\mathbf{Y}$) phase space, that is, $\boldsymbol{\mathcal{Z}} = \mathbf{Z} = (z, \delta)$. We assume the initial distribution given by

$$f_0(z, \delta; 0) = n_0 \frac{1}{\sqrt{2\pi}\sigma_\delta} e^{-(\delta - hz)^2/2\sigma_\delta^2}. \tag{152}$$

This is a Gaussian distribution, as in Eq. (86), but chirped with chirping coefficient h, with $n_0 = 1/\ell$ being the line density. The quantities in the exponent of Eq. (149) are

$$z_s(\boldsymbol{\mathcal{Z}}_0) = z + R_{56}(s)\delta, \\ z_\tau(\boldsymbol{\mathcal{Z}}_0) = z + R_{56}(\tau)\delta. \tag{153}$$

Here we simplified the notation by $R_{s0,56} \to R_{56}(s)$ and $(z_0, \delta_0) \to (x, \delta)$. The integral over δ and z can be performed to obtain

$$b(k; s) = b_0(k, s) + \int_0^s d\tau R_{56,s\tau} \frac{I(\tau)}{\gamma I_A} \\ \times Z[k(\tau); \tau] b[k(\tau); \tau] \xi(s, \tau). \tag{154}$$

Here, $I_A = ec/r_e$ is the Alfven current, $I = ecn_0 C(\tau)$ is the peak current at τ, and

$$k(\tau) = \frac{C(\tau)}{C(s)} k; C(\tau) = \frac{1}{1 + hR_{56}(\tau)}. \tag{155}$$

The compression factor C was introduced in Eq. (90) and

$$\xi(s, \tau) = e^{-k^2 \sigma_\delta^2 [C(s)R_{56}(s) - C(\tau)R_{56}(\tau)]^2/2}. \tag{156}$$

The meaning of Eq. (154) is clear: The density modulation at τ becomes an energy modulation through the impedance (that can be enhanced through compression), which in turn becomes density modulation through R_{56}. The process can lead to an instability detrimental for beam applications such as X-ray free-electron laser. Note the factor ξ becomes small if the energy spread σ_δ is large. Intentionally increasing σ_δ could therefore suppress the instability.

III. BUNCH CONTROL VIA THE ELECTRON GUN

In this Section, we present a variety of gun-based methods for controlling the electron bunch distribution. We classify these methods according to the distribution of the generated bunch: either *mesoscopically* shaped or *macroscopically* smooth bunches. (see Fig. 1 and the related discussion in Section I). In the shaped case, the electron gun is used to directly generate the desired shaped bunch. In the smooth case, the gun generates a smooth bunch which is subsequently shaped by the methods presented in the later Sections, IV and V, of this review. Gun-based methods presented will include those that have achieved control over the 2D transverse or 1D longitudinal distribution, as well as recent progress in controlling the complete 3D distribution.

A. Introduction

The electron gun consists of a cathode surface in a region of accelerating and focusing fields and is used to generate the initial electron bunch which is injected into the linac. The emission of the initial electron bunch distribution from the cathode surface and its evolution through the gun is complex and varied. The reader interested in understanding electron gun physics in-depth is pointed to (Dowell *et al.*, 2008; Dowell, 2010, 2016; Rao and Dowell, 2015) and the references therein. In the introduction to this Section, we only present high-level details needed for understanding gun-based shaping.

Electron bunch generation begins at the cathode surface and continues until the bunch exits the external fields (accelerating and focusing) of the gun. The initial electron distribution emitted from the cathode surface is affected by both the properties of the cathode material

and the fields at the cathode surface while the evolution of the bunch is affected by the self-generated fields of the bunch (e.g. space charge) as well as the external fields. However, in this Section, we focus on an idealized regime where the self-generated fields are negligible, and the external fields of the gun merely accelerate and guide the bunch to high energy while preserving the shape (although not the size) of initial cathode distribution. The later Sections of this review include the other effects.

Electron guns used in e- linacs can be classified by their electric fields (DC and RF) or by their cathode type. There are a wide variety of cathodes in use, but they can conveniently be classified by their emission mechanism: field, thermionic, and photo emission, FE, TE and PE, respectively. When these cathodes are operated in an electron gun, they go by the names of (DC or RF) field emission gun, (DC or RF) thermionic gun, and (DC or RF) photoemission cathode, which is usually shorted to "photocathode" gun. However, as stated above, the impact of the gun fields on shaping are being ignored (in this Section) so we will drop the DC and RF labels, unless noted otherwise. On the other hand, the classification according to the emission mechanism is crucial to understanding gun-based shaping methods since the methods presented in this section differ substantially in their capabilities for controlling the initial cathode distribution emitted from each cathode type.

Historically, all three cathode types have been used to generate smooth beams, but the same cannot be said of shaped beams. TE cathode guns are robust electron sources and they are the workhorse of storage rings for light sources around the world (LightSources, 2021). They are widely used in e- linacs for producing smooth bunches but have not been used in applications requiring either transverse or longitudinal shaping. FE cathode guns are still in the R&D phase and are not yet used in the e- linacs that are the topic of this review. Nonetheless, the FE cathode gun has great potential and it has demonstrated transverse shaping, although not longitudinal, in proof of principle experiments. The PE cathode gun is the workhorse for linacs in SASE FEL facilties and they are being used for both transverse and longitudinal shaping applications. The reason FE and TE cathodes are not used for longitudinal shaping is because the electron emission from these cathodes follows the time structure of the applied electric field (ranging from continuous in a DC gun to ns-scale in an RF gun) which is too long for the time scales of interest (fs-ps) in this review. In principle, TE cathodes could be used for generating transversely shaped bunches, but this avenue has not been pursued. For a thorough discussion of the various cathode types, we refer the interested reader to the tutorial paper from (Jensen, 2018).

The remainder of this Section is organized into three parts. In III.B, we present a review of cathodes focused on cathode properties relevant to bunch shaping. Second, we present demonstrated methods for generating smooth bunches (III.C) and end with methods for generating shaped bunches (III.D).

B. Cathode review

In this Section, we review different cathodes types and summarize their properties relevant to shaping (see Table I). These properties include the magnitude of the electric field at the cathode (E_c) and the work function (W) of the cathode, which is the minimum energy needed to extract an electron from the cathode surface. There are also several properties of the emitted electron bunch of relevance: current density (J), average kinetic energy emitted from the cathode (E_k), transverse and longitudinal rms spot sizes (σ_x, σ_z) and dimensionless rms momentum (σ_{p_x}, σ_{p_z}). We also include some parameters that are specific to the particular cathode type but those will be introduced in the corresponding Sections below.

1. Current density

There is great variation in the magnitude of the current density, J, generated by the three cathode types. Typical values of J for the three cathode types are shown in row *Longitudinal Parameters* of Table I. Note that while the current density does not play a direct role in bunch shaping we include this Section since it provides the necessary background to understand how we classify cathodes in this review.

A FE cathode consists of an arrangement of one or more sharply pointed tips, or emitters, located on a cathode immersed in an applied "macroscopic" electric field, E_c. The geometry of the tip is characterized by the field enhancement factor β which enhances E_c ($\sim$100 MV/m) to generate extremely high local fields, $F = \beta E_c$ ($\sim$10 GV/m) on the tip, from where the electrons are extracted. The local FE current density is given by the Fowler-Nordheim equation (Fowler and Nordheim, 1928)

$$j_{FE}(F) = A_{FE} F^2 e^{-\frac{w^{3/2}}{F}}, \qquad (157)$$

with material constant A_{FE}. Lower case letters, j_{FE} and w, are used to differentiate local from macroscopic parameters, J_{FE} and W. The field emission electron microscope (FEM) is an example of a DC FE gun based on a single FE tip. In a typical FEM, a tungsten tip ($W = 4.5$ eV) has a radius ranging from 100 nm to 1 mm. Small tips are capable of extremely high local current densities, $j_{FE} \sim 10^6$ A/mm^2, with correspondingly small emission areas $\sim$100 nm^2.

The challenge of operating an electron gun with a single FE tip is that the current emitted is only $\sim$0.1 mA per tip, which is inadequate for operation of the linacs

we consider in this review since they require larger current (at least $1A$ is needed) from the cathode. The solution is to use large surface area FE cathodes engineered to hold many microscopic emitters. Three such engineered FE cathodes have been developed and used in FE electron guns: field emission arrays (FEAs) (Jarvis et al., 2010), carbon nanotubes (CNTs) (Laszczyk, 2020), and ultra-nanocrystalline diamonds (UNCDs) (Baryshev et al., 2014). These surfaces are composed of a quasi-continuous distribution of electron emitters of macroscopic area ~ 1 mm^2 and achieve macroscopic current densities of $J_{FE} \sim 1$ A/mm^2, sufficient for e- linacs. The FE parameters shown in Table I are for large area engineered FE cathodes, not single tips.

TE is the liberation of electrons from a heated surface and are operated in guns with applied electric field on the cathode E_c (~ 100 MV/m). Dispenser cathodes have work functions of W (~ 1.6 eV) and are representative of TE cathode and are operated with a temperature of T (~ 1400 K) to allow electron emission. The TE macroscopic current density is given by the Richardson's Law (Richardson, 1913),

$$J_{TE}(T) = A_{TE}T^2 e^{-\frac{W}{k_b T}} \tag{158}$$

at temperature $T(K)$, where A_{TE} is a material constant, and k_b is the Boltzmann constant. Dispenser cathodes operated in electron guns with high electric fields, E_c, (~ 100 MV/m) can achieve current densities of $J_{TE} \sim 1$ A/mm^2 and are therefore sufficient for electron linacs.

PE occurs when the photons illuminating the photocathode surface have an energy in excess of the work function of the cathode material. The current density emitted from a PE cathode (Dowell and Schmerge, 2009) is

$$J_{PE}(h\nu) = n_{ph} A_{PE}(h\nu - W)^2, \tag{159}$$

where $h\nu$ is the photon energy, h is Plank's constant, A_{PE} is a material constant, n_{ph} is the number of photons per unit area, and $(h\nu - W)$ is known as the excess energy and is the kinetic energy of the emitted electrons. PE cathodes are operated in electron guns with high electric fields, E_c, (~ 100 MV/m) and achieve the highest macroscopic current density of the three cathode types, $J_{PE} \sim 1000$ A/mm^2.

2. Kinetic energy

E_k is the average, total kinetic energy of the electrons emitted from the cathode which are emitted isotropically into the half-sphere over the cathode. In this Section, we give the expression for E_k for each of the different cathode types. As will be seen in the next section, the intrinsic emittance is determined by E_k. Note that the total kinetic energy, E_k, is equipartitioned into each degree of freedom, x, y, z.

The average kinetic energy of electrons emitted from a FE cathodes (Forbes, 2015) is ,

$$E_k = \frac{e\hbar F}{\sqrt{8 m_e W}}, \tag{160}$$

where m_e is the electron mass and typical $E_k = 0.3$ eV for FE cathode tips. For electrons emitted from a TE cathode, their average kinetic energy is,

$$E_k = \frac{3}{2} k_b T, \tag{161}$$

where, for example, $E_k = 0.12$ eV for a dispenser cathode operating at $T=1400$ K. Finally, electrons emitted from PE cathodes have average kinetic energy given by,

$$E_k = \frac{h\nu - W}{2}, \tag{162}$$

where, for example, a copper cathode (work function of 4.65 eV) is illuminated by a laser of wavelength λ=248nm, has $E_k = 0.18$ eV.

An examination of the three expression for E_k reveal two things worth pointing out. First, the expressions for FE and PE both depend on the work function W. In practice, the work function W is usually replaced by the effective work function ϕ_{eff} to account for the Schottky effect (Dowell and Schmerge, 2009) which is the reduction of the work function by the applied field and it plays a role in all emission processes, especially field emission. However, we keep W in all expressions for simplicity. Second, note that a parameter associated with the emission mechanism appears in each expression for E_k: it depends on the local field F for FE cathodes, on temperature T for TE cathodes, and on photon energy $h\nu$ for PE cathodes.. Typical values of E_k are shown in *Excess Kinetic Energy* of Table I.

3. Emittance

In the first half of this Section, we explain the fundamental role that emittance plays in limiting the resolution of the shape and end with a discussion on the intrinsic emittance generated by the three cathode types.

a. Shaping resolution   We define the shaping resolution to be the smallest spot size that can be obtained at the end of a beamline, of transfer matrix R, for a fixed beam spot size at the beginning of the beamline. If we place a lens (i.e. a quadrupole magnet) at the beginning of this

beamline (Carey, 1987), then the lens can be varied to minimize the transverse spot size at end,

$$\sigma_{x,min} = \frac{R_{12}\varepsilon_x}{\sigma_{x,0}}, \tag{163}$$

where R_{12} is the $(1,2)$ element of R, $\sigma_{x,0}$ is the beam size at the beginning of the beamline (which is also at the lens) and ε_x is the horizontal beam emittance (assumed to be constant). Note, we are ignoring collective effects of the beam and higher-order magnetic optics of the beamline. (For the derivation of Eq. (163) see Eq. 6.45 in (Carey, 1987) and the discussion therein. Also note that smallest spot size is not necessarily the waist of Eq. (106).) A similar expression for the minimum longitudinal bunch length ($\sigma_{z,min}$) at the end of the beamline can be derived,

$$\sigma_{z,min} = \frac{R_{56}\varepsilon_z}{\sigma_{z,0}}, \tag{164}$$

where $\sigma_{z,0}$ is the longitudinal bunch length at the beginning of the beamline, ε_z is the longitudinal beam emittance, and R_{56} is an element of the transfer matrix. Note that in this case, we use a *longitudinal lens* (i.e. an RF cavity) at the beginning of the beamline to minimize the bunch length at the end of the beamline

The significance of the preceding two equations is that they give us three ways to minimize the spot size. or increase the resolution. Two ways can be easily controlled, increase the initial spot size ($\sigma_{x,0}, \sigma_{z,0}$) or decrease the elements (R_{12} and R_{56}). The third way of increasing the resolution is to decrease the emittance (ε_x, ε_z), however, this is not possible once the beam leaves the cathode since the emittance is an invariant. Therefore, the intrinsic emittance of the beam emitted from the cathode is the ultimate limit on the shaping resolution, assuming no collimation or cooling of the beam. We now turn to the intrinsic emittance generated at the cathode.

b. Intrinsic emittance Having established the fundamental importance of emittance in shaping, we now discuss the emittance of the electron bunch emitted from the cathode, i.e. the *intrinsic emittance*. Assuming there is no correlation between the three phase-space planes of the electrons emitted from the cathode, the normalized intrinsic emittance (see Eq. (109)) in 6D is given by the product of the three normalized intrinsic 2D emittances,

$$\varepsilon_{6D}^n = \varepsilon_x^n \varepsilon_y^n \varepsilon_z^n, \tag{165}$$

Assuming there is no correlation between the position of an emitted electron and its transverse momentum, then intrinsic rms normalized horizontal emittance at the cathode is,

$$\varepsilon_x^n = \sigma_x \sigma_{p_x}, \tag{166}$$

where $\sigma_x = \sqrt{<x^2>}$ and $\sigma_{p_x} = \sqrt{<p_x^2>}/m_0 c$. Corresponding expressions apply for each coordinate so we can rewrite, Eq. (166) to give the *intrinsic rms normalized emittance at the cathode* for each of the 2D phase space planes as,

$$\varepsilon_c^n = \sigma_c \sigma_{p_c}, \tag{167}$$

where σ_c and σ_{p_c} are the initial rms bunch size and the dimensionless rms momentum at the cathode, respectively, in the horizontal/vertical/longitudinal plane. The accelerator designer can easily control σ_c, but σ_{p_c} is an intrinsic property of the cathode. Developing cathodes with low initial σ_{p_c} is an active area of research for each of the emission mechanisms. The intrinsic rms normalized transverse emittance of all three cathode types (e.g., using the x-direction to be definite) (Flöttmann, 1997) can be written in terms of the kinetic energy of the electrons emitted from the cathode, E_k, and is given by,

$$\varepsilon_x^n = \sigma_x \sqrt{\frac{2E_k}{3m_e c^2}}, \tag{168}$$

where c is the speed of light, σ_x is the rms horizontal spot radius on the cathode. The dimensionless rms horizontal momentum at the cathode, σ_{p_x}, is given by the *square root term* in Eq. (168) and can also be thought of as the *intrinsic rms normalized emittance at the cathode per rms spot size* (i.e. $\sigma_{p_x} = \varepsilon_x^n/\sigma_x$) by virtue of Eq. (166) and has units of μm/mm. This equation shows that the intrinsic emittance of the cathode depends on E_k (of the emitted electrons) and implies that low emittance can be achieved if the electrons are emitted with small E_k. Unfortunately, as we can see in the PE cathode case, as E_k approaches 0, so does the emitted current J_{PE}, see Eq. (159). Thus there is no way to have both low intrinsic emittance and high current density.

The expression for the intrinsic rms normalized horizontal emittance at the cathode per rms spot size of any of the cathode types can be found by substituting the appropriate expression for E_k from III.B.2 into Eq. (168). As an example, the expression for σ_{p_x} for PE is found to be,

$$\varepsilon_x^n/\sigma_x = \sqrt{\frac{h\nu - W}{3m_e c^2}}, \tag{169}$$

with typical values shown in the row *Transverse Parameters* of Table I along with typical spot sizes on the cathode, σ_x. Note that it is becoming increasingly common to talk about the mean transverse energy ($MTE = 2/3\, E_k$) of the electrons emitted from the cathode. By equipartition of energy, we know the mean longitudinal energy is $1/3\, E_k$ and we can find an expression the intrinsic rms normalized longitudinal emittance at the cathode per rms bunch length as,

$$\varepsilon_z^n/\sigma_z = \sqrt{\frac{h\nu - W}{6m_e c^2}}, \tag{170}$$

4. Response time

The response time of a cathode is defined as the temporal lag between the excitation of the cathode and emission of electrons. While all types of cathodes have intrinsically short response times, in practice, the duration of the electron bunch emitted by FE and TE cathodes is controlled by the duration of the electric field, not the cathode response time. On the other hand, the emission of electrons from PE cathodes is gated by a laser pulse which can be much shorter than the electric field duration of the gun. Therefore, the PE cathode is the only one used for generating longitudinally shaped bunches.

The response time of PE cathodes varies from the fs-scale, for metallic photocathode, to 10's ps-scale, for certain semiconductor photocathodes, (Smedley and Poelker, 2012). Row *Longitudinal Parameters* of Table I shows response times of the cathodes. However, we do not list these times for FE and TE cathodes since they are not used for longitudinal shaping in electron guns.

5. Cathode selection guidance

Based on the above considerations, the source designer must choose one of the three electron cathodes for their shaping application. The choice between cathodes is made by considering the trade-offs between the various factors shown in Table I. The following factors must be considered by the electron source designer in choosing a cathode that meets the requirements of the shaping application.

1. *Longitudinal shaping.* If the application requires longitudinal shaping at the source then there is only one choice, choose PE.

2. *Charge per bunch required.* This is a product of the transverse spot size and current density. The larger the spot size on the cathode the more charge.

3. *Transverse and longitudinal emittance requirements.* These will determine the bunch shaping resolution. The smaller the spot size on the cathode the lower the emittance.

While PE cathodes have the most shaping capabilities, they also come with the most operational complexity as they require large laser systems that must be actively synchronized the linac. TE cathodes do not require laser systems but have the added complexity of operating at $T > 1000°C$. While FE cathodes do not require laser systems or heating so are the simplest of all.

We end this Section with the caveat that we left out the hybrid cathodes that combine two of the three basic emission mechanisms: photo-assisted field emission cathode (Mustonen *et al.*, 2011; Swanwick *et al.*, 2014) and

photo-assited thermionic emission cathode (Sun *et al.*, 2006). These are potentially important cathodes, since they open up the possibility of longitudinal shaping for FE and TE cathodes, however, we do not cover them here due to space limitations.

C. Smooth distributions

We now turn to the most common initial electron distribution emitted from an electron gun, a smooth bunch. The methods introduced in this section focus on generating smooth bunch profiles, and therefore, bunches with low intrinsic emittance, Eq.(168), in order to provide high shaping resolution. In this case, shaping is accomplished by shaping systems located downstream of the electron gun described in Sections IV, V, and VI of this review. In this Section, we describe how the smooth distributions are produce for each of the three cathode types (Jensen, 2018).

1. FE based smooth transverse distributions

FE cathodes are thought to have the potential of providing a robust source of low emittance electrons. As discussed in III.B.1, large area, engineered FE cathodes, can be based on FEAs, CNTs or UNCDs. All of their surfaces are composed of a quasi-continuous distribution of electron emitters (Fig. 5). In this Section, we describe demonstrated methods of producing smooth transverse distributions for each of the engineered FE cathode types.

The FEA is an arrangement of a large number of discrete tips on the cathode surface. The emission sites of the FEA, shown in Fig. 5(a), come from diamond tips on pyramid bases (Piot *et al.*, 2014) separated by 10 μm. The CNTs can be deposited into regular arrays, like the FEA, or randomly oriented on the cathode surface. In the later case, emission sites of the CNT, shown in Fig. 5 (top, center), comes from the randomly oriented fibers covering the CNT surface (Mihalcea *et al.*, 2015). The emission site of the UNCD cathode, shown in Fig. 5(top, right), is believed to come from the grain boundaries of the UNCD thin film deposited on the surface (Baryshev *et al.*, 2014). The emitter separation of the FEA, UNCD, and CNT cathodes are approximately 1-10 μm, 0.1-1 μm, and 10-100 nm, respectively (Fig. 5, top row). All engineered FE cathode types have experimentally produced smooth electron distributions (Fig. 5, bottom row). The smoothness of the transverse distribution from large-area FE cathodes is limited by two factors: the need to merge emission from discrete emission sites and non-uniformity of the emitters across the cathode. The discreteness of the emitters becomes less noticeable as the separation between sites decreases. Further, another limitation of the FE cathode RF gun is the large energy spread and long

Table I Cathode characteristics

|	(large area) Field Emission	Thermionic Emission (TE)	Photoemission (PE)	units
typical field at cathode (E_c)	100	100	100	MV/m
Local field, temperature, wavelength	$F \sim 10$ GV/m	$T = 1400$ K	$\lambda \simeq 260$ nm	
typical work function (W)	4.5 (tungsten)	1.6 (dispenser)	4.6 (copper)	eV
typical excess kinetic energy (E_k)	0.3	0.12	0.2	eV
typical spot size σ_x	1-3	1-3	0.1-10	mm
typical $\varepsilon_{x,y}$	0.3-1.0	0.12-0.36	0.02-2	μm
emission time	—	—	$10^{-3} - 10$	ps
current density (J)	≤ 1	≤ 1	$\leq 10^3$	A/mm^2
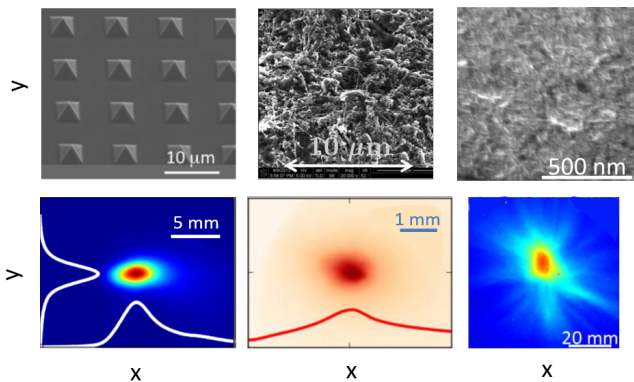
Figure 5 Three types of large area FE cathodes. Surface images (top row, left-right) of FEA, CNT, and UNCD cathodes. The distance between the emitters is largest for FEA (10 μm), smaller for CNT (0.1-1 μm) and smallest for UNCD (10-100 nm). Smooth beam images (bottom row, left-right) generated by FEA, CNT, and UNCD cathodes. Figure adapted from (Baryshev *et al.*, 2014; Mihalcea *et al.*, 2015; Piot *et al.*, 2014).

Figure 6 Comparison of inhomogeneous (a) and homogeneous (b) laser profiles at the plane of the PE cathode. From (Halavanau *et al.*, 2017).

duration of the electron bunches due to emission taking place over a large range of RF phases. This issue arises in TE cathode RF guns too and it is handled with an alpha magnet (Lewellen *et al.*, 1998). Even though these are longitudinal parameters, they adversely affect the transverse properties of the bunch. Most importantly, the large energy spread leads to strong chromatic aberrations.

2. TE based smooth transverse distributions

The TE cathode gun is one of the simplest and most robust electron sources; it is the workhorse of light source facilities around the world (LightSources, 2021). Historically, the TE cathode gun has not been used for the more demanding application of driving FELs, due to its lower transverse beam brightness and difficulty with achieving a short pulse. However, researchers at SACLA (Asaka *et al.*, 2017) have developed a low-emittance thermionic-gun-based injector. In the injector, electron beams are emitt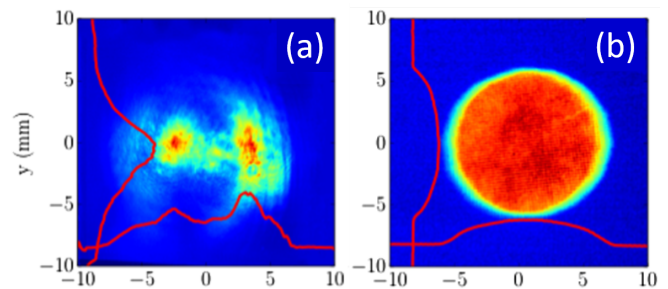ed from a CeB$_6$ thermionic cathode of 3-mm diameter located in a DC 500-kV gun followed by a beam chopper and a bunch compresso to produce an electron beam with high peak current (3–4 kA) and low transverse normalized-slice emittance (below 1 μm) sufficient to drive a compact free-electron laser.

3. PE based smooth transverse distributions

Many applications of the electron source require smooth transverse distributions, such as the uniform flat-top or the Gaussian profile. While the typical IR output from the photocathode laser (TEM$_{00}$) is very nearly Gaussian at the output, its shape can become distorted during the frequency upconversion process to generate UV. In the typical case, smoothing is usually needed due to inhomogeneities in the far-field image of the UV beam at the PE cathode plane (see Fig. 6(a)) to produce the desired homogeneous profile (see Fig. 6(b)). Methods to achieve smooth transverse electron distributions with PE cathodes use optical elements, inserted into the laser path, of which there are two types: passive (e.g., micro-lens array) and active (e.g., deformable mirror). In general, passive elements are simpler but active elements have greater capabilities for obtaining complicated transverse distributions as we describe below.

Microlens arrays (MLAs) are passive optical elements used to homogenize the laser's transverse profile (Bich *et al.*, 2008). Fig. 7(a) shows a schematic of homogenizing optics where a UV laser beam (248 nm) passes
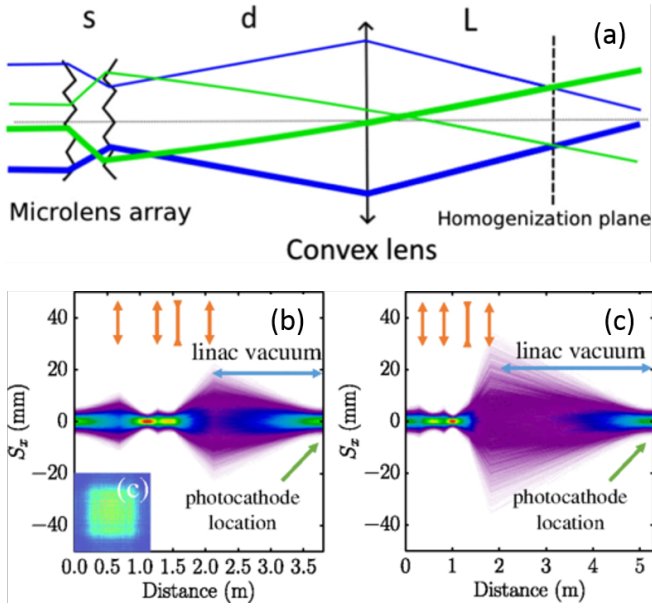
Figure 7 Microlens array (MLA) based homogenization system. (a) Diagram of several rays passing through the MLA pair onto the homogenization plane. (b,c) ray tracing from the homogenization plane, through imaging optics and to the photocathode plane for two different systems. From (Halavanau *et al.*, 2017).
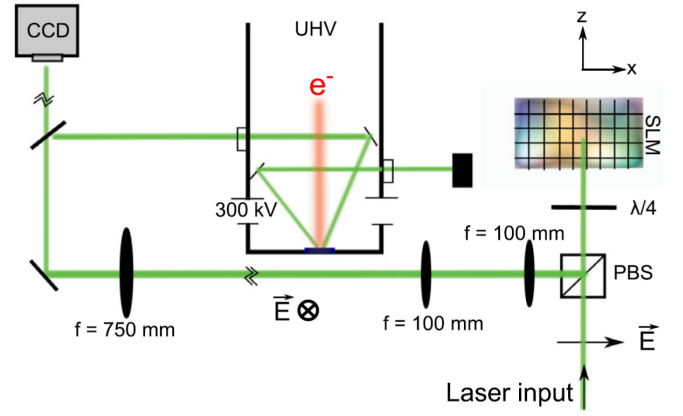

Figure 8 Spatial light modulator (SLM) homogenization system. An SLM is actively adjusted via a computer (not shown) in a feedback loop to homogenize the laser profile on the photocathode by monitoring the profile at the virtual cathode location with a camera (CCD). From (Maxson *et al.*, 2015) .

through a pair of MLAs followed by a convex lens resulting in a continuous and homogenized laser profile at the homogenization plane (see Fig. 6(b)). The MLA system is located outside the beam vacuum, and the laser emerging from the MLA system (at the homogenization plane) has large beam divergence, which makes transport to the cathode difficult. A solution was found in Ref. (Halavanau *et al.*, 2017) where the laser was imaged from the homogenization plane to the PE cathode plane located approximately 3.5 m away with an imaging system Fig. 7(b,c). Another advantage of passive homogenization systems (e.g. MLA) have over active ones is that they will homogenize a laser profile even if its profile fluctuates from shot to shot. Active homogenization systems cannot do this because they use feedback loops to transform an incoming inhomogeneous laser profile into a homogeneous one which requires the incoming profile to be stable on the time-scale of the feedback loop. On the other hand, the passive system will instantaneously transform an arbitrary inhomogeneous laser profile into a homogeneous.

Active optical elements provide a flexible, yet more complicated system for obtaining a homogenized transverse distribution at the phototcathode. The first systems were based on deformable mirrors (DMs) consisting of an array of electrically adjustable small mirrors. The intensity profile of the laser pulse is controlled by adjusting the angles of the small mirrors which are under computer control. After the laser beam reflects off the

DM, a beam splitter sends a small fraction of the beam to the CCD camera located at the *virtual cathode* while the majority continues to the PE cathode. Numerical optimization algorithms, such as the genetic algorithm (GA), are run on a computer to adjust the angles of the small mirrors to optimize the profile (Matsui *et al.*, 2008). The DM method is still actively underdevelopment by researchers at (Li *et al.*, 2017). A second active approach is based on spatial light modulators (SLMs) as shown in Fig. 8. While SLMs work on a different optical principal (birefringence) than DMs (reflection), their shaping functionality is the same. The SLM is placed in a feedback loop which monitors the transverse profile of the laser (again, a small fraction at the virtual cathode) while a computer running a GA is used to control the SLM element. In Fig. 8, laser light enters from the bottom (in the z direction) and is polarized along x. A quarter-wave plate and SLM act as a polarization rotator with spatial dependence, which shapes the light when used with a polarizing beam splitter (PBS). The surface of the SLM is then 4-f imaged (f =100mm lens pair) onto an intermediate plane to preserve the beam divergence, and then this intermediate plane is imaged with a single long focal length lens (f =750mm) onto either the photocathode or a CCD. An ultrahigh vacuum (UHV) mirror reflects light to the center of the photocathode. The SLM-based system was found to have greater capacity in handling poor input laser quality (Maxson *et al.*, 2015) than the DM-based system. On the other hand, the SLM only works with IR and visible light while DM-based methods can work in the UV (Li *et al.*, 2017) thus avoiding the distortions in the up-conversion process. In other words, this continues to be a lively area of research.
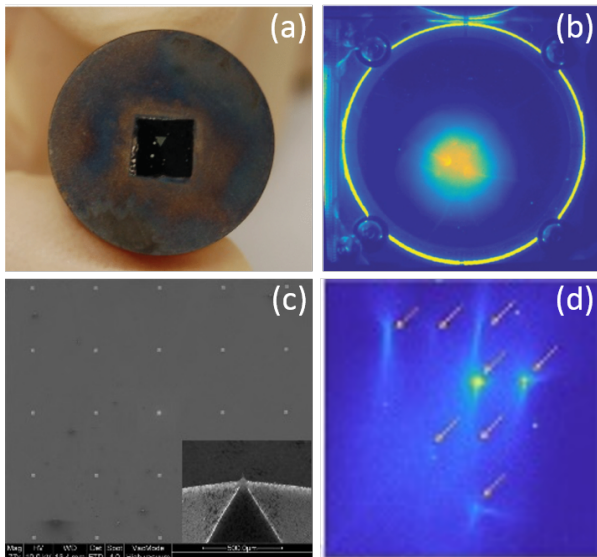
Figure 9 FEA cathodes for transverse shaping. Photograph of an FEA cathode with a triangular array inside a central square area of a round cathode plug (a) and its corresponding electron bunch image on a downstream screen showing a slight triangular shape (b). SEM image of rectangular array on a cathode plug (c) with diamond tips (inset), and its corresponding electron bunch image on a downstream screen showing a grid-shaped pattern (d). From (Andrews *et al.*, 2020; Nichols *et al.*, 2020).

D. Shaped distributions

Transverse bunch shaping methods have been demonstrated for both FE and PE cathodes. In this subSection, we present shaping methods developed for these two emission mechanisms.

1. FE based shaped transverse distributions

Transverse shaping of the electron distribution generated by large area FE cathodes is controlled by engineering the emitting surface. This is an active research area but is not yet capable of generating high quality electron bunches suitable for the modern electron linacs we consider here. None the less, due to the recent activity in this area coupled with its great potential to be used in electron linacs we present it here. In particular, FEA-based cathodes have been used to generate both continuous (e.g., triangular) and modulated (e.g., array of beamlets) transverse distributions.

As an example of a transversely shaped distribution, a triangular array of pyramid emitters was deposited into a 1-mm equilateral triangle with $\sim$ 10 μm spacing on a cathode plug (Andrews *et al.*, 2020), as shown in Fig. 9(a). The downstream electron bunch image captured on an electron imaging screen, as shown in Fig. 9(b), suggests a triangular shape, but space-charge

effects and the long phase emission period are suspected to have blurred the image. The emission period can be shortened with gated FEAs (Jarvis *et al.*, 2009). Modulated transverse distributions (e.g., an array of spots) have been generated at the source with FEA cathodes, see Fig. 9(c) (Nichols *et al.*, 2020). The downstream electron bunch image Fig. 9(d) shows that original modulation was maintained but degraded due to the non-uniformity of the emitters. Note that as spacing of the emitters gets closer (e.g., a nanoengineered FEA (Graves *et al.*, 2012), it becomes more difficult to maintain the modulation. In order to reproduce the initial source modulation out of the gun, the charge must be kept very low, to avoid space charge dilution, and the FEA must be gated to keep energy spread low.

2. PE based shaped transverse distributions

In this Section, we present methods for shaping the transverse distribution of the electron bunch generated by a PE cathode in an electron gun. This distribution is controlled by controlling the transverse profile of the photocathode laser beam.

Certain accelerator applications require shaped electron bunches with modulated transverse distributions, such as an array of beamlets [(Fig. 10(a,b)] or hollow beams [Fig. 10(c)]. To date, these patterns have only been generated with passive optical systems in PE cathode guns based on an optical mask (Rihaoui *et al.*, 2009; Wisniewski *et al.*, 2012) or an MLA (Halavanau *et al.*, 2017). In the former case, an optical mask is inserted into the laser path to block the unwanted part of the laser beam to create the desired laser pattern, which is then imaged onto the PE cathode to create the electron distribution on the cathode. Examples include: an aluminum plate with six holes was used to generate a *positive* array of laser spots [(Fig. 10(a)], and a painted quartz plate was used to create a *negative* hollow laser ring [(Fig. 10(c)]. In the latter case, an MLA system was used to create a large array of spots [(Fig. 10(b)] by changing the location of the convex lens (Fig. 7(top)) as described in the reference.

3. PE based shaped longitudinal distributions

As stated above, due to the difficulties in gating FE and TE cathodes at the picosecond (and shorter) timescale, longitudinal bunch shaping is the domain of PE cathode guns. The longitudinal bunch shape of the electron distribution generated by a PE cathode gun depends on both the temporal laser pulse shape and the response time of the PE material. Photocathode laser systems can generate laser pulse duration ranging from 10's fs to 10's of ps while PE cathode response times range
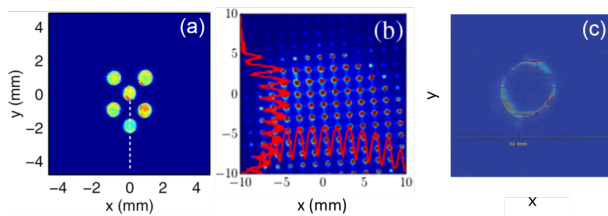
Figure 10 Transversely shaped electron bunch distributions from a PE cathode. An array of spots generated with an optical mask (a) and an MLA (b). A hollow beam generated with an optical mask (c). Images from (Halavanau *et al.*, 2017; Rihaoui *et al.*, 2009; Wisniewski *et al.*, 2012).
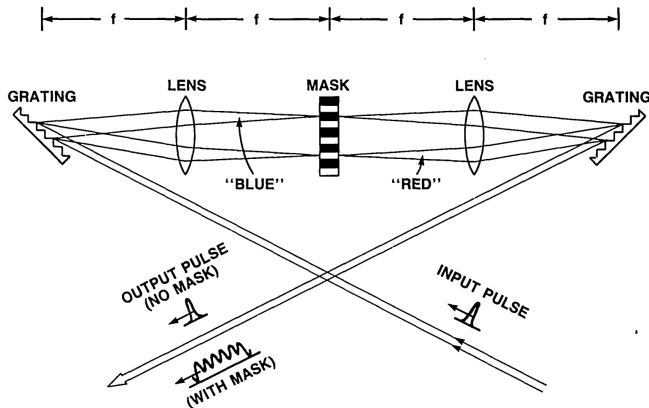

Figure 11 Fourier Transform laser pulse shaping system. An input laser pulse enters a 4-f grating stretcher and its frequency spectrum is exposed at the Fourier Transform Plane. A mask described by a frequency domain transfer function, $H(\omega)$, is placed at this plane to modify the spectrum of the input laser pulse resulting in a modified time structure of the output laser pulse. From (Weiner *et al.*, 1988).

from the fs-scale, for metallic cathodes, to 10's ps for some semiconductor cathodes (Dowell *et al.*, 2008). For longitudinal shaping applications via the electron gun, one chooses a PE cathode with a response time much less that the duration of the laser pulse. In this way, the electron bunch temporal shape will simply follow the laser temporal shape. Laser pulse shaping methods have been used to generate both single bunches (smooth and shaped) and bunch trains. Various methods used to control the longitudinal distribution are presented below and are categorized as either frequency domain or time domain.

a. Frequency domain laser shaping Frequency domain methods manipulate the frequency spectrum of the input laser pulse to control the time profile of the output laser pulse. There are two approaches to frequency domain-based laser shaping that have been applied to PE cathode guns: acousto-optic programmable dispersive filter and Fourier Transform pulse shaping.

The acousto-optic programmable dispersive filter (AOPDF) relies on a longitudinal interaction between a polychromatic acoustic wave and a polychromatic optical wave in the bulk of a birefringent crystal. It controls the IR spectrum by controlling the group delay versus wavelength with a programmable acoustic wave in the birefringent crystal. Optical signals in the hundreds of Terahertz range are controlled with RF signals in the tens of MHz range. This is a compact IR device, installed between the photocathode laser oscillator and amplifier, with high temporal resolution, which makes it suitable for the shaping of femtosecond IR pulses. A widely utilized commercial AOPDF—the DAZZLER™—can shape pulse length over a 6-ps duration at a maximum repetition rate close to a MHz (Tournois, 1997; Verluise *et al.*, 2000).

In Fourier Transform pulse shaping, a 4-f grating stretcher is used to expose the spectrum of the input laser pulse in the spectral Fourier Transform plane of the stretcher (see Fig. 11). The spectrum can be controlled with an optical element placed in this plane. Let the spectrum of the input laser pulse be given by $X(\omega)$ and the frequency-domain transfer function of the optical element at the Fourier plane be represented by, $H(\omega)$, then the spectrum of the output laser pulse is given by their convolution,

$$Y(\omega) = X(\omega)H(\omega), \qquad (171)$$

so that the output laser pulse in the time-domain is just the inverse Fourier Transform of this convolution,

$$y(t) = F^{-1}[Y(\omega)], \qquad (172)$$

The first systems used for longitudinal shaping with an RF photocathode gun (Neumann *et al.*, 2009, 2003) used a fixed mask, described by $H(\omega)$, located in the Fourier Transform plane to modify the amplitude of the input spectrum. However, this method can be based on the modification of any of the amplitude, phase, or polarization of the input spectrum. described by the appropriate $H(\omega)$. After the spectrum is modified, the second grating is used to bring the spectrum back to a line in the time domain to generate the desired temporal pulse shape of the output pulse. The laser community has achieved both laser pulse trains and temporal flattop laser pulses (Fig. 12) using masks to modify the amplitude of the spectrum (Weiner *et al.*, 1988). Recent Fourier Transform pulse shaping approaches are based on programmable SLMs due to their superior resolution and flexibility. The first beam physics applications used an SLM system to convert a Gaussian pulse of 9 ps FWHM into a flattop pulse of the same length to drive a photocathode RF gun (Yang *et al.*, 2002) in order to reduce the transverse emittance. Gaussian pulses were transformed into both triangular and super-triangular laser pulses with an SLM-based system (Kuzmin *et al.*, 2018)..

The advantage of the frequency domain method is its high resolution and flexibility to produce various pulse
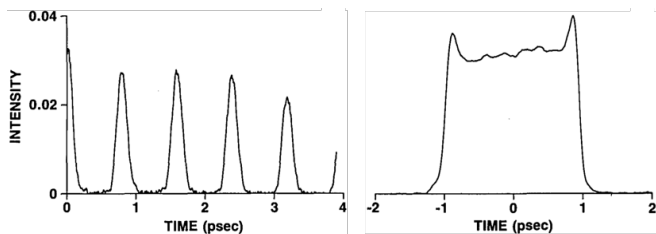
Figure 12 Demonstrated laser pulse temporal shapes generated with the Fourier Transform laser pulse shaping system of Fig. 11 using an amplitude mask. Laser pulse train (left) and quasi-flattop laser pulse (right). Courtesy (Weiner *et al.*, 1988).

shapes while the downside is its complexity and stability. In addition, the need to transmit the IR laser pulse shape through the laser amplifier and harmonic conversion crystals introduces non-linearities causing differences between the IR and UV laser pulse shapes. One solution to this issue would be the development of programmable Fourier Transform pulse shaping applied directly in the UV and another would be the development of PE cathodes that can respond to IR; this is another active area of research.

b. Time domain laser shaping Time domain methods are based on laser pulse stacking. This is where a series of short laser pulses are longitudinally combined (i.e., stacked) to form the desired longitudinal profile. There are a number of methods that have been tried in the past. An early method used for PE cathode guns (Siders *et al.*, 1998) was based on a Michelson interferometer in which n 50/50 beam splitters, embedded in a set of optical delay legs, were used to split a single input laser pulse into a series of 2^n output laser pulses. Their device was used to generate a train of 16 pulses at THz spacing with very low loss. More recently, pulse stacking with birefringent crystals is almost exclusively used. It exploits the group velocity mismatch between the ordinary and extraordinary axes of a birefringent crystal (e.g. α -BBO) (Power and Jing, 2009; Will and Klemz, 2008; Zhou *et al.*, 2007). When a single Gaussian input pulse passes through a uniaxial birefringent crystal rotated at an angle relative to the crystal axis, two output pulses will emerge projected onto the ordinary and extraordinary axes. The time delay difference between these two components is due to the differing group velocities along the ordinary and extraordinary axes

$$\Delta t = L_{Xtal} \left(\frac{1}{v_{g,e}} - \frac{1}{v_{g,o}} \right), \qquad (173)$$

where L_{Xtal} is the length of the crystal, and $v_{g,e}$ and $v_{g,o}$ are the group velocities associated with the extraordinary and ordinary optical axes, respectively. The crystal
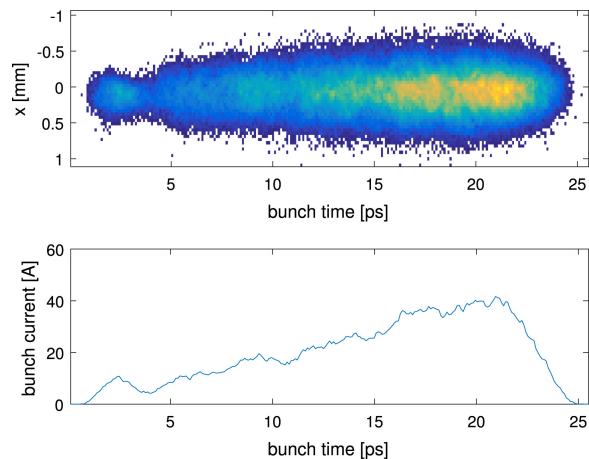

Figure 13 Measured double-triangular electron bunch generated with α -BBO laser pulse stacking. (top) x-z projection of the bunch on a screen. (bottom) corresponding current profile. From (Loisch *et al.*, 2018).

length controls the delay between the different polarizations while crystal angle controls the relative intensity of the different polarizations. If the angles are set to $45°$, a flattop laser pulse is obtained (Will and Klemz, 2008). However, it is also possible to obtain more complicated shapes (e.g., a double triangle) by passing the input laser pulse through a stack of crystals rotated at optimized angles relative to the incident laser pulse. In (Liu *et al.*, 2019; Loisch *et al.*, 2018), they used this method to generate a double-triangular laser pulse which was then used in PE cathode gun to generate the corresponding double-triangular electron bunch as shown in Fig.13.

The advantage of the time domain method lies in its simplicity. Moreover, if the pulse stacker is implemented in the UV (Power and Jing, 2009) the distortion of the pulse due to nonlinearities arising during the frequency-conversion process is completely averted. Its downside is its limited flexibility since there is no way to change the crystal length on the fly though some control over the pulse shape can be achieved by rotating the crystals.

In addition to producing shaped the "continuous" distributions shown above (i.e. triangle and flattop), birefringent crystals have also been used to produce bunch trains. Reference (Li and Kim, 2008) proposes a simple and more compact system specifically adapted to bunch train generation. The setup consists of a concentric stair-step echelon combined with a focusing lens. The echelon consists of a series of concentric flat zones with different thicknesses, so as to introduce a discrete delay correlated to the transverse radius. At the focal point of a downstream lens, such a configuration produces pulses that are delayed in time. The nature of the echelon design introduces small delays so that the method is well adapted to the generation of pulses with ps-scale temporal separation. In (Li and Kim, 2008) the technique is numerically investigated for the formation of a train comprising 20- to

100-fs bunches with an associated bunching factor peaked at ~ 0.5 THz for application to a coherent Smith-Purcell THz source based on a 50-kV electron beam.

4. PE based spatio-temporal (3D) shaping

A natural continuation of the progress with programmable 2D transverse and 1D longitudinal laser shaping are methods for shaping the complete 3D spatio-temporal distribution of the laser. This method gives the electron source designer complete control over the initial electron distribution in a PE cathode gun. Recall, however, that this does not mean complete control over electron distribution in general, since, once it leaves the cathode it is subject to nonlinear space-charge forces that will distort the initial distribution. Interestingly, one of the main motivations for this line of R&D is to generate electron bunches with uniform 3D ellipsoid distributions since these are the only distributions whose space-charge fields (i.e., internal force fields) are linear functions of position (Luiten *et al.*, 2004). This gives rise to particularly simple dynamical behavior: a uniform ellipsoid under the influence of its self-fields (electrostatic) will change its size but retain its shape; a uniform ellipsoid with linear internal fields. In addition to the 3D ellipsoid laser pulse shapes, other distributions of interest. For example, one could combine a 2D circular transverse distribution, for low emittance, with a triangular 1D longitudinal distribution, for high-transformer ratio, to simultaneously achieve different objectives.

Recent progress in 3D laser shaping with IR laser pulses has been achieved by several groups (Kuzmin *et al.*, 2020, 2019; Mironov *et al.*, 2016). For example, researchers used multiple programmable SLMs (Fig. 14) to generate both a 3D quasi-cylinder and 3D quasi-ellipsoidal IR laser distribution (Mironov *et al.*, 2016). In addition to SLM-based techniques, 3D shaping can also be accomplished via control of chromatic aberration. as discussed by (Li and Chemerisov, 2008), where a DAZZLER™ system is used to introduce a complex spectral structure that, in combination with a highly dispersive section, resulted in the generation of a 3D ellipsoidal bunch (Li and Lewellen, 2008).

Despite the successful demonstration of these 3D IR laser distributions, these have not yet been used to extract electrons from a photocathode gun. The next steps require converting the IR pulse shapes to the UV where they can be used to exicte PE cathodes. Nonetheless, this is an exciting and important area of research. Further, it is interesting to note that the demand for 3D laser pulse shaping, and its subsequent progress, has been lead by the PE cathode gun community, not the laser community.
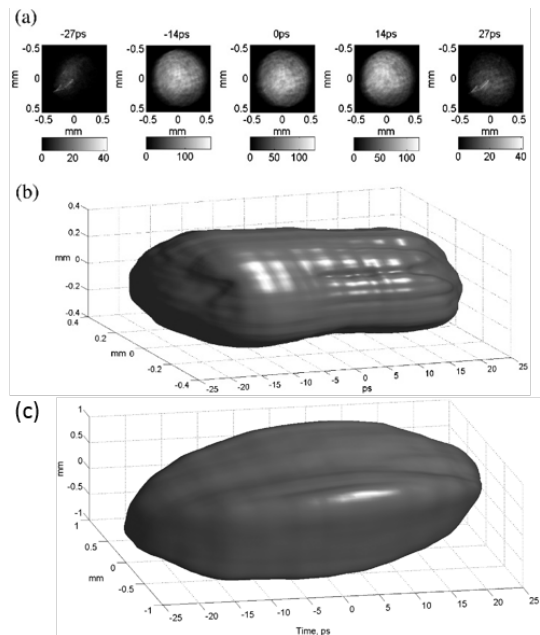

Figure 14 Demonstrated 3D laser distributions based on SLMs. (a) measured transverse intensity distributions of quasi-cylindrical pulses; (b) reconstructed 3D distribution of quasi-cylindrical; (c) reconstructed 3D distribution of quasi-ellipsoidal. From (Mironov *et al.*, 2016).

IV. BEAM CONTROL WITHIN ONE DEGREE OF FREEDOM USING EXTERNAL FIELDS

This section discusses the use of externally applied fields to shape the beam distribution. We introduce the general formalism and discuss various methods commonly employed or recently proposed to control the beam distribution in longitudinal or transverse directions and some of the associated projections.

A. General considerations

We categorize beam-shaping methods using external fields into three main approaches. The first approach consists of finding a phase-space transformation that maps a given initial distribution to a final target distribution. A second approach introduces a local coupling between two variables so that the overall transformation is still within one degree of freedom (DOF) but the local coupling enables access to another variable. A common example is the use of dispersive collimation where particles with given energies are removed via the introduction of a local correlation between energy and position. Finally, another technique employs an interceptive mask that modifies the momentum or affects the particles' transmission according to their transverse position.

The use of an external electromagnetic field to impart correlation within one DOF is commonly employed to control the beam properties along any of the DOFs.

For instance time-dependent fields are commonly used to introduce, e.g., a linear correlation in the longitudinal phase space (LPS) such as $\delta = hz$, that can later be exploited using a longitudinally dispersive beamline to alter the final bunch length via $\sigma_{z,f} = C^{-1}\sigma_{z,0}$, where the compression factor is defined in Eq. (90) and h and R_{56} are, respectively, the chirp in the incoming (z, δ) LPS and the longitudinal dispersion associated with the dispersive section. It should be pointed out that the same mechanism applies to the other DOFs, e.g., in the transverse planes to focus the beam.

Generally, this type of manipulation can be split into two stages referred to as a modulator and a converter. The modulator introduces a position-dependent momentum whereas the converter consists of a beamline providing a momentum-dependent change in the position. For the sake of simplicity, we consider the case of uncoupled motion between the three DOFs, and focus on one of the DOFs, with particle conjugated variables (ζ_i, ζ_{i+1}) [where $i \in [1, 3, 5]$]. We consider the effect of the modulator to impact a position-dependent external force along $\hat{\zeta}_{i+1}$ of the form $F_{i+1}(\zeta_i)$, resulting in the coordinate transformation

$$\zeta_{i,0} \to \zeta_{i,m} = \zeta_{i,0},$$
$$\zeta_{i+1,0} \to \zeta_{i+1,m} = \zeta_{i+1,0} + F_{i+1}(\zeta_{i,0}), \qquad (174)$$

under an impulse approximation (where we ignore change in the particle position). In the latter equation, the subscripts $_0$ and $_m$ respectively refer to the values before and after the modulator section. The downstream converter beamline introduces a momentum-dependent change in position so that the final position $\zeta_{i,f}$ is related to the upstream coordinate as

$$\zeta_{i,m} \to \zeta_{i,f} = \zeta_{i,m} + G_i(\zeta_{i+1,m}),$$
$$\zeta_{i+1,m} \to \zeta_{i+1,f} = \zeta_{i+1,m}. \qquad (175)$$

Overall, the transformation associated with the beamline takes the coordinate $(\zeta_{i,0}, \zeta_{i+1,0})$ and transforms it to the final coordinate

$$\begin{pmatrix} \zeta_{i,f} \\ \zeta_{i+1,f} \end{pmatrix} = \begin{pmatrix} \zeta_{i,0} + G_i[\zeta_{i+1,0} + F_{i+1}(\zeta_{i,0})] \\ \zeta_{i+1,0} + F_{i+1}(\zeta_{i,0}) \end{pmatrix}, \quad (176)$$

and a transfer map defined such that $(\zeta_{i,f}, \zeta_{i+1,f}) = \mathcal{M}(\zeta_{i,0}, \zeta_{i+1,0})$ can be formally associated with the transformation.

Considering an initial phase-space-density distribution $\Phi_0(\zeta_{i,0}, \zeta_{i+1,0})$ and invoking Liouville's theorem II.C.1 $\Phi_f(\zeta_{i,f}, \zeta_{i+1,f})d\zeta_{i,f}d\zeta_{i+1,f} = \Phi_0(\zeta_{i,0}, \zeta_{i+1,0})d\zeta_{i,0}d\zeta_{i+1,0}$, we can write the final phase-space distribution as

$$\Phi_f(\zeta_{i,f}, \zeta_{i+1,f}) = \Phi_0[\mathcal{M}^{-1}(\zeta_{i,0}, \zeta_{i+1,0})], \quad (177)$$

since the Jacobian of the transformation is unity. We now consider the projection along the position direction $\hat{\zeta}_i$ defined as

$$P_i(\zeta_i) = \int d\zeta_{i+1} \Phi(\zeta_i, \zeta_{i+1}). \qquad (178)$$

By virtue of the charge-conservation we now have

$$P_f(\zeta_{i,f}) = P_0[\zeta_{i,0}(\zeta_{i,f})]\frac{\partial \zeta_{i,0}}{\partial \zeta_{i,f}}, \qquad (179)$$

where the RHS can be written solely in term of $\zeta_{i,f}$ via inversion of the map described in Eq. (176). Therefore, by properly tailoring the transformation $\mathcal{M}$, one can modify the shape of the projection along any direction. To illustrate the set of derived equations we first consider the simple example of the bunch compression discussed in Section II.C.1. Given the incoming-LPS coordinate (z_0, δ_0), the correlation introduced by the linear accelerator and the energy-dependent path length from the compression can be described, respectively, by $F(z) = hz$ and $G(\delta) = R_{56}\delta$. Here for sake of simplicity we take both F and G to be linear functions of s such that the overall transformation is described by

$$\begin{pmatrix} z_f \\ \delta_f \end{pmatrix} = \begin{pmatrix} z_0 + R_{56}[\delta_0 + h(z_0)] \\ \delta_0 + h(z_0) \end{pmatrix}, \qquad (180)$$

so that the final longitudinal charge distribution is

$$\lambda_f(z_f) = C\lambda_0\left(\frac{z_f - R_{56}\delta_0}{C}\right), \qquad (181)$$

where C represents the compression factor defined in Eq. (90). Taking $\lambda_0(z_0) = 1/\sqrt{2\pi\sigma_{z,0}^2}\exp[-z_0^2/(2\sigma_{z,0}^2)]$ and assuming that δ_0 represents a random uncorrelated fraction energy spread, we obtain

$$\lambda_f(z_f) \quad = \quad \frac{C}{\sqrt{2\pi}\sigma_{z,0}}e^{-\frac{C^2 z_f^2}{2\sigma_{z,0}^2}}e^{-\frac{C^2 R_{56}^2 \delta_0}{2\sigma_{z,0}^2}}, \quad (182)$$

which showcases the well-known results associated with linearized bunch compression previously mentioned in Section II. A simple extension is to consider the case where nonlinearities play a role in the compression. For instance, owing to RF curvature in the linac employed to impart the chirp, a quadratic dependence on $\zeta_{5,0}$ is also introduced when the bunch length does not strictly verify the condition $\sigma_{z,0} \ll \lambda$ (where λ is the wavelength of the accelerating mode in the linac) so that $F(z) = hz + h_2 z^2$. Likewise, standard four-bend bunch compressors are known to introduce a second-order longitudinal dispersion T_{566}, and consequently $G(\delta) = R_{56}\delta + T_{566}\delta^2$. Following the same approach as before yields a final distribution along the longitudinal axis to be of the form (Li, 2001)

$$\lambda_f(z_f) = \frac{1}{\sqrt{2\pi}\sigma_{z,f}}\frac{e^{z_f/(\sqrt{2}\sigma_{z,f})}}{[-z_f/(\sqrt{2}\sigma_{z,f})]^{1/2}}\Theta(-z_f), (183)$$

where $\Theta()$ is the Heavide function. Thus, the nonlinear transformation introduced by the functions F and G now results in a change of the bunch current profile commonly encountered in, e.g., magnetic compression, subjected to strong nonlinearities; see for instance (Dohlus et al., 2004).

The example considered so far can be generalized via the introduction of arbitrary nonlinear functions F and G to control the final phase space (ζ_i, ζ_{i+1}) correlations. By controlling the degree of the non-linearity introduced, one can, in principle, tailor the correlations within the phase space to produce the desired profile along each of the phase-space directions.

In beam physics, it is customary to expand the transport map via a truncated Taylor series and consider each nonlinear order separately; see Section II.C.5. This is a typical process in optics where, e.g., the energy-dependent path length introduced by a dispersive section, e.g., a bunch compressor as discussed above, is now written as $\zeta_{5,f} = \sum_{i=1}^{6} R_{5i}\zeta_{i,0} + \sum_{i=1}^{6} \sum_{j\geq1}^{6} T_{5ij}\zeta_{i,0}\zeta_{j,0} + \mathcal{O}(\zeta_0^3)$ and likewise for any beamline; see Eq. (120). For sake of simplicity we assume a pencil bunch so that high-order coupling between the longitudinal and transverse phase spaces can be neglected in the expansion of the transfer map.

Although we have taken the modulator function to be a continuous function of the coordinate over a finite interval one can also consider a periodic function of the form $F(u) = F_0 \cos(ku + \phi)$. In such a case, a series of concatenated modulators described with a function $F_i(u) = F_{0,i} \cos(k_i u + \phi_i)$ could be used to synthesize the desired final distribution by controlling the term of the Fourier series associated with the final distribution. Such a description can be employed to describe a chain of linacs operating at different frequencies. By shaping the energy spectrum and using a transformation in $z = \mathcal{M}(\delta)$ one could tailor the energy spread or current profile. In practice introducing an arbitrarily high harmonic may be challenging, especially for the time-dependent field given, e.g., the limited set of klystron frequencies.

1. Interceptive beam shaping

The shaping techniques described so far combine nonlinear external fields with the beamline providing nonlinear correlation between the position and momentum of particles. A straightforward shaping technique consists of intercepting the beam with a mask that has a given "transmission" function $T(x,y)$ so that the mask can only affect the distribution in the transverse spatial coordinates. The incoming phase-space distribution $\Phi_i(\zeta_1...\zeta_4)$ is then simply transformed as

$$\Phi_f(x,y) = \Phi_i(x,y)T(x,y), \qquad (184)$$

where we assume the mask to be thin and only affecting the beam according to the transverse coordinate (and not the momentum). Consequently, the transverse profile along, e.g., ζ_1 can be found from

$$P_f(x) = \int dy \Phi_i(x,y)T(x,y) = \int_{\Upsilon_-}^{\Upsilon_+} dy \Phi_i(x,y), (185)$$

where we have taken the mask to be a binary function with unity value in the domain $y \in [\Upsilon_-(x), \Upsilon_+(x)]$.

A drawback of such a masking technique is its intrusive nature, which may hinder its application to high-power or high-repetition-rate beams, as the beam loss associated with the shaping process could result in radiological activation or hardware damage.

2. Manipulation with local coupling

In such a transformation, an external field introduces the required correlations $\zeta_j = H(\zeta_i)$ and the coordinate ζ_j is manipulated (e.g., via a function similar to F or G). Finally, the inverse transformation H^{-1} to remove the correlation between ζ_j and ζ_i. In the process the shaping imparted to ζ_i via shaping of ζ_j is preserved. A simple example of implementation of a manipulation based on local coupling regards dispersive collimation where a local dispersion bump locally introduces a correlation between transverse position and energy where a collimator is used to tailor the energy distribution (e.g. remove energy tail). One advantage of local-coupling methods combined with a mask is its simple implementation while providing a high degree of control over the beam shape (via precise shaping of the intercepting mask (Majernik et al., 2021)). However, the mask can result in significant particle losses.

B. Generation of shaped current distributions

One important aspect of LPS control resides in the ability to control the beam current profiles for application in beam-driven wakefield accelerators and light sources. Historically, current beam shaping has been an integral part of electron injectors based on continuous wave (CW) electron sources where a combination of masks and RF cavities − often dubbed "chopping" systems − are commonly employed to form bunches for injection in the subsequent linear accelerators (Smith, 1986; Tiefenback and Krafft, 1993).

1. Local coupling combined with transverse masking

Current shaping techniques were initially discussed as a means to prebunch the beam for free-electron laser (FEL) (Nguyen and Carlsten, 1996) applications. The method was eventually demonstrated at the ATF facility (Muggli et al., 2008) where it was also extended to shaping beyond microbunch generation (Shchegolkov et al., 2015). In brief, the method combined local coupling between the transverse (horizontal) and longitudinal phase spaces with masking. To explain the technique we consider an incoming bunch with an LPS chirp h_0 sent to a dispersive section with a transfer matrix producing

the horizontal and longitudinal dispersions η and ξ_-, respectively. Downstream of such a dispersive section the final horizontal and longitudinal positions of an electron is

$$x_m = x_\beta + \eta \delta_0,$$
$$z_m = z_0 + \xi_- \delta_0, \qquad (186)$$

where x_β is the geometric contribution to the beam spot. An intercepting mask in (x_m, y_m) with transmission function dependent on the horizontal coordinate $f(x_m)$ will tailor the transverse profile. For a beam with correlated momentum spread $\delta_0 = h_0 z_0$ the mask will alter the shape of the longitudinal distribution following $f[x_m/(\eta h_0)]$. A downstream dispersive section with longitudinal dispersion ξ_+ and designed to suppress the dispersion introduced by the section upstream of the mask results in the overall shaping function $f[x_m/(\eta h_0)]$. This technique bears similarities with the frequency-domain temporal shaping method commonly encountered in ultrafast laser shaping, where a frequency-chirped laser is dispersed, its spectrum modified with a mask, and subsequently recombined (Kuzmin *et al.*, 2018; Weiner *et al.*, 1988); see Section III.D.3.a.

The method was implemented in (Muggli *et al.*, 2008) to form train of sub-ps bunch. In such a case the mask consists of a set of N slits so that the transmission can be approximated as $f(x_m) = \sum_{\ell=1}^{N} \delta(x_m - X_\ell)$ where X_ℓ are the horizontal positions of the vertical slits. Considering $X_\ell = \ell D$, with D being the inter-slit spacing on the mask, the final modulation period is shown to be (Hyun *et al.*, 2019; Muggli *et al.*, 2008)

$$\Delta z \simeq D \frac{1 + \xi_+ h_0}{\eta h_0}. \qquad (187)$$

A practical implementation of the method used for a proof-of-principle experiment at the Accelerator Test Facility (ATF) appears in Fig. 15(a). The beam was locally dispersed in a dogleg beamline with vanishing net dispersion. The dogleg included an optical lattice providing an antisymmetric dispersion function with maximum value attaining $\eta \simeq 1.5$ m in close proximity to the mask location. The mask consisted of sub-mm tungsten wire; see inset photograph in Fig. 15(a). The produced bunch was modulated in energy and time (z) owing to the final LPS chip. Given the fixed dispersion, the mask sets the energy modulation period and ultimately the final current modulation, which can be controlled via the incoming chirp and ξ_+ function of the beamline. It was demonstrated that the technique enables some control over the final temporal period downstream of the dogleg beamline: for ~ 10 pC modulations with sub-mm periods were produced, consistent with the resonant excitation of wakefield in plasmas (Muggli *et al.*, 2010a). Likewise, this method was extended to produce bunches with triangular beam distribution at the ATF (Shchegolkov *et al.*,
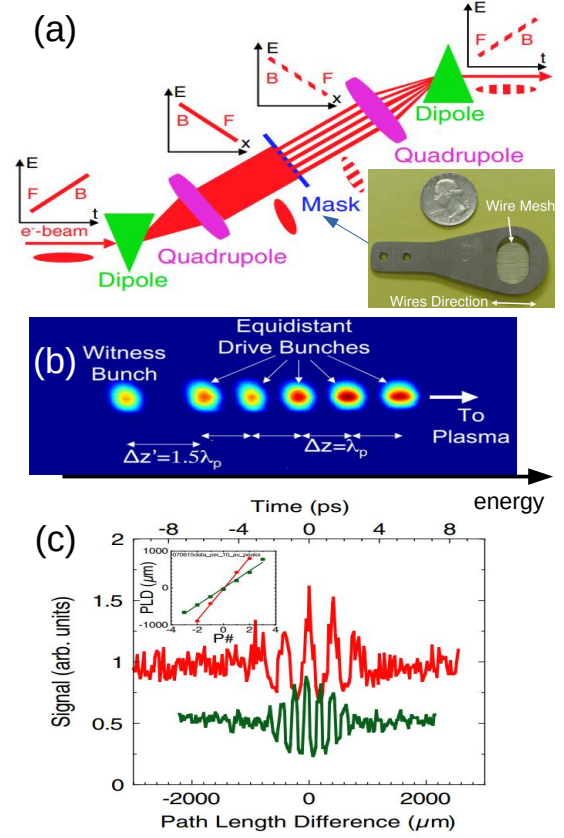

Figure 15 Experimental generation of a sub-picosecond bunch train at the ATF using a dogleg beamline (a) with final beam distribution in a downstream spectrometer (b) and measured temporal distribution via interferometry of coherent transition radiation (c). Figure adapted from (Muggli *et al.*, 2008) and (Muggli *et al.*, 2010b).

2015) for wakefield excitation in dielectric-lined waveguides (Antipov *et al.*, 2012).

More recently, a current-shaping method based on directly introducing a spatio-temporal coupling was proposed using a pair of transverse-deflecting cavities (Kur *et al.*, 2009; Zholents *et al.*, 1999). Such a concept is similar to the one commonly employed in DC electron injectors as part of the low-energy "chopping" system required to form the bunched beam before injection in an RF linac (Wilson *et al.*, 1985).

The main advantage of this shaping method compared to the one based on dispersive coupling stems from its ability to directly couple z to one of the transverse coordinates. In the method does not suffer from CSR-induced phase-space degradation. It can also provides control over the LPS chirp. In its simplest implementation, the beamline consists of two transverse deflecting cavities (TDCs) separated by a beamline described by a transfer matrix M. The overall matrix of the system in

(x, x', z, δ) is

$$S = R_{TDC}(\kappa_d) M R_{TDC}(\kappa_u) \equiv \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \quad (188)$$

where $R_{TDC}(\kappa)$ represents the matrix of a TDC with strength κ, see Section (II.B.3). In order for the beamline to be globally uncoupled, the 2×2 block matrix should vanish: $B = C = 0$. Solving for $B = 0$ yields the condition on the elements of M and cavity strength to

$$m_{12} = -\frac{L\left(\kappa_d + \kappa_u\left(Lm_{21} + m_{11} + 2m_{22}\right)\right)}{2\kappa_u},$$
$$\kappa_d = -\frac{\kappa_u\left(Lm_{21} + 2m_{22}\right)}{2}, \text{and} \quad (189)$$
$$m_{11} = -\frac{Lm_{21}}{2} - \frac{\kappa_u}{\kappa_d}.$$

These matching conditions ensure S becomes block diagonal. The (x, x') block matrix is independent of the TDC strengths, and the overall effect of the combined TDCs is to modify the LPS chirp $h \equiv \langle \zeta\delta \rangle / \langle \zeta^2 \rangle^{1/2}$ as $h \to h = h_0 - S_{65}$ with

$$S_{65} = \frac{L\kappa_u^2\left[(Lm_{21} + 2m_{22})^2 + 4\right]}{16}. \quad (190)$$

We have $S_{65} > 0$ for a cavity length $L \neq 0$. It is interesting to note that the chirp vanishes in the thin-lens model of the TDC ($L = 0$). Likewise the block matrix becomes

$$M_\perp = \begin{pmatrix} m_{22}^{-1} & 0 \\ m_{21} & m_{22} \end{pmatrix}, \quad (191)$$

consistent with Ref. (Kur *et al.*, 2009) where the choice $m_{11} = m_{22} = -1$ was made so that $\kappa_u = \kappa_d$ (given that decoupling forces $m_{12} = 0$). The shaping mask is located between the two TDCs and with sufficient phase advance from the upstream TDC to ensure the beam has a significant correlation $\langle xz \rangle$ at its location. The profile of the mask $T(x, y)$ is then tailored to provide the desired temporal shape. The application of this method to produce a modulated relativistic electron bunch was first discussed in (Du *et al.*, 2012) in the context of superradiant THz radiation, and further investigated numerically in (Ha *et al.*, 2020) to support the formation of various current shapes to support a beam-driven wakefield accelerator with enhanced transformer ratio. Figure 16 presents numerical simulation results for a minimal system composed of two TDCs and a quadrupole magnet with a different type of mask to produce doorstep distributions [Fig. 16(a,b)] for beam-driven acceleration and sub-ps bunch trains with variable microbunch spacing [Fig. 16(c,d)].

The masks described so far stop a fraction of the incident beam. An alternative solution is to use a spoiling mask made of a thin foil to alter the beam properties associated with a fraction of the beam. An example of
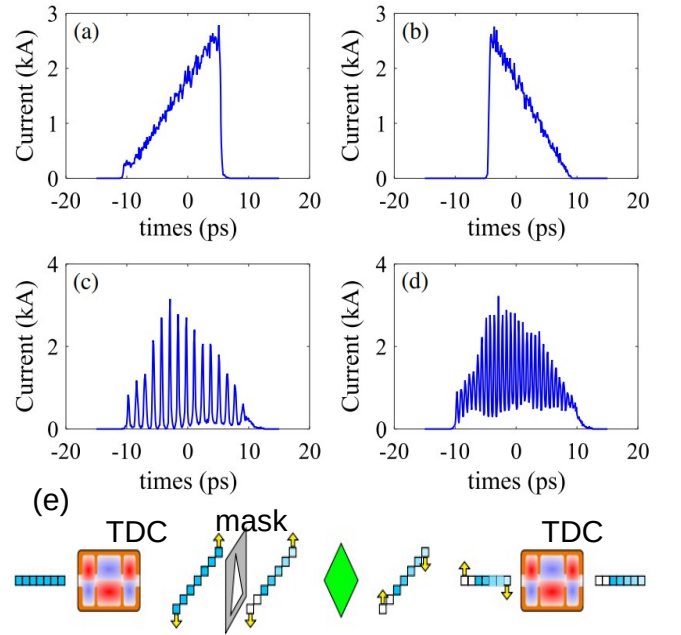

Figure 16 Numerical modeling of current-profile shaping showcasing the generation of ramped (a), reverse-triangle (b) and modulated (c,d) bunch distributions using local correlations imparted by a pair of transverse-deflecting cavities (TDC) combined with a mask (e) (adapted from (Ha *et al.*, 2020)).

implementation uses a slotted foil at the Linac Coherent Light Source (LCLS) (Emma *et al.*, 2004) to selectively spoil the transverse emittance of an electron beam (the fraction of the beam intercepted by the foil undergoes multiple scattering and suffer an emittance growth). The method was employed to ultimately control the duration of X-ray pulses in FELs, as only the unscattered beam population contributed to the lasing mechanism and supported the generation of isolated femtosecond X-ray pulses along with twin variable-delay X-ray pulses using a dual-slot foil (Ding *et al.*, 2015).

2. Modulators combined with longitudinally dispersive sections

A simple implementation of a current-shaping technique consists of introducing an energy modulation using RF accelerating cavities to properly control the longitudinal dispersion in the downstream beamline. The energy modulation is introduced as a time-varying field with accelerating voltage of the form $V(z) = \widehat{V}\cos(kz + \psi)$, where $\widehat{V}$, ψ, and k are, respectively, the peak accelerating voltage and phase, and the accelerating mode wavevector. The LPS transformation associated with the downstream dispersive sections is described via the Taylor expansion of the longitudinal map with, e.g., first- and second-order the coefficients $R_{56} \equiv \frac{\partial z}{\partial \delta}$, $T_{566} \equiv \frac{\partial^2 z}{\partial \delta^2}$ as de-
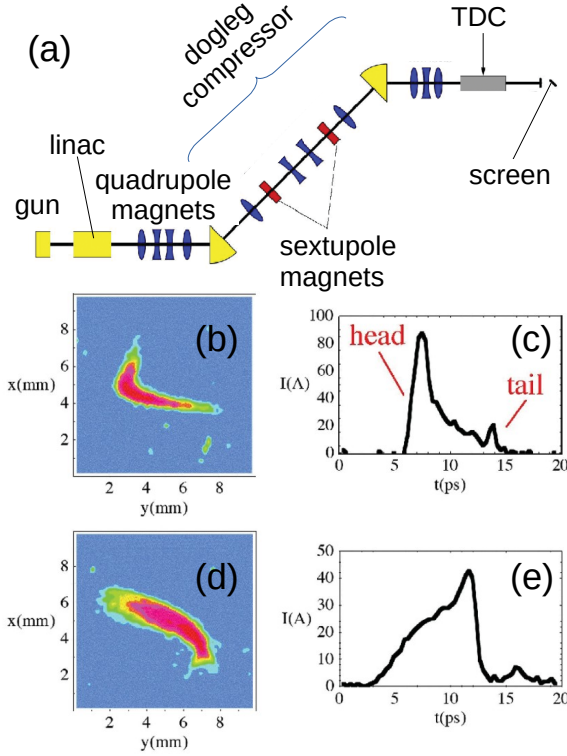
Figure 17 Experimental setup for phase-space shaping via introduction of nonlinear longitudinal dispersion at the NEP-TUNE facility (a) and resulting LPS (b,d) with associated current distribution (c,e) experimentally generated. Figure adapted from (England *et al.*, 2008).

tailed in Section II.C.5. The two widely-used approaches are to either (*i*) introduce a linear energy modulation and controls the nonlinear dispersion in the dispersive section, or to (*ii*) control the nonlinear modulation using accelerating sections operating at multiple frequencies.

A technique to control the nonlinear longitudinal dispersion to shape the beam current distribution was introduced in (England *et al.*, 2005). It was observed that compression of a "long" incoming bunch accelerated in an S-band linac such that $kz \ll 1$ is not strictly valid. As a result, the final fractional energy offset is

$$\delta \simeq \frac{e\widehat{V}}{\bar{E}} \left[-kz \sin\psi + \frac{(kz)^2}{2} \cos\psi \right], \qquad (192)$$

(where $\bar{E}$ is the final beam energy) resulting in significant quadratic correlation in the LPS with associated current profile described by an equation similar to Eq. (183). It was recognized that the addition of sextupole magnets in the dispersive section (arranged as a dogleg beamline) provided a sufficient handle over the T_{566} coefficient to ultimately control the current distribution. The experimental implementation of this technique at the NEPTUNE facility demonstrated the

generation of $\sim$ 12-MeV bunches with linearly ramped current profiles (England *et al.*, 2008); see Fig. 17. The technique was also shown to provide some control over the ramp shape and direction.

The introduction of higher-order longitudinal dispersion (up to third order) was also proposed to suppress current spikes arising from collective effects in the LCLS accelerator (Charles *et al.*, 2017). At LCLS, extreme current values at the head and tail of the electron bunch are responsible for substantial CSR-induced projected emittance growth. The nominal operating mode of LCLS consits of truncating the head and tail to suppress the current spikes to improve the FEL lasing performances. It was shown that current spikes can be strongly suppressed via control of the $U_{5666} \equiv \frac{1}{6}\frac{\partial^3 z}{\partial \delta^3}$ longitudinal-dispersion term using an octupole magnet without a significant increase of the horizontal slice emittance.

Further improving the precision of shaping methods discussed in this section can be accomplished using multiple energy modulators operating at different frequencies. Such an approach allows for the incoming LPS to be an arbitrary polynomial function $\delta(z)$ with its coefficients controlled by the settings of the modulator accelerating-voltage amplitude and phase. The solution was investigated using a dual frequency modulator operating at the frequencies f_1 and $f_n \equiv nf_1$ with total accelerating voltage $V(z) = V_1 \cos(k_1 z + \psi_1) + V_n \cos(k_n z + \psi_n)$, where $\widehat{V_{1,n}}$ and $\psi_{1,n}$ are, respectively, the accelerating voltages and operating phases of the two linac sections, and $k_{1,n} \equiv 2\pi f_{1,n}/c$. Assuming $k_{1,n}z_0 \ll 1$, the electron's LPS coordinates downstream of the linac are ($z_l = z_0, \delta_l = a_l z_0 + b_l z_0^2$), where $a_l \equiv a_0 - e(k_1 V_1 \sin\psi_1 + k_n V_n \sin\psi_n)/\bar{E}_l$, $b_l \equiv b_0 - e(k_1^2 V_1 \cos\psi_1 + k_n^2 V_n \cos\psi_n)/(2\bar{E}_l)$ with e being the elementary charge and $\bar{E}$ the beam's average energy downstream of the linac. The passage of the bunch through a longitudinally dispersive section results in an electron final coordinate to be given as a function of the initial coordinates following $z_f = a_f z_0 + b_f z_0^2$ with $a_f \equiv 1 + a_l R_{56}$ and $b_f \equiv b_l R_{56} + a_l^2 T_{566}$. Taking the initial current to follow the Gaussian distribution $I_0(z_0) = \hat{I}_0 \exp[-z_0^2/(2\sigma_{z,0}^2)]$ (where $\hat{I}_0$ is the initial peak current), and invoking the charge conservation gives the final current distribution

$$I_f^u(z_f) = \frac{\hat{I}_0}{\Delta^{1/2}(z_f)} \exp\left[-\frac{(a_f + \Delta^{1/2}(z_f))^2}{8b_f^2 \sigma_{z,0}^2} \right],$$
$$\times \Theta[\Delta(z_f)], \qquad (193)$$

where $\Delta(z_f) \equiv a_f^2 + 4b_f z_f$ and $\Theta()$ is the Heaviside function.

An example of the method discussed above was experimentally implemented in the FLASH facility at DESY (Piot *et al.*, 2012), where a linac composed of 1.3-
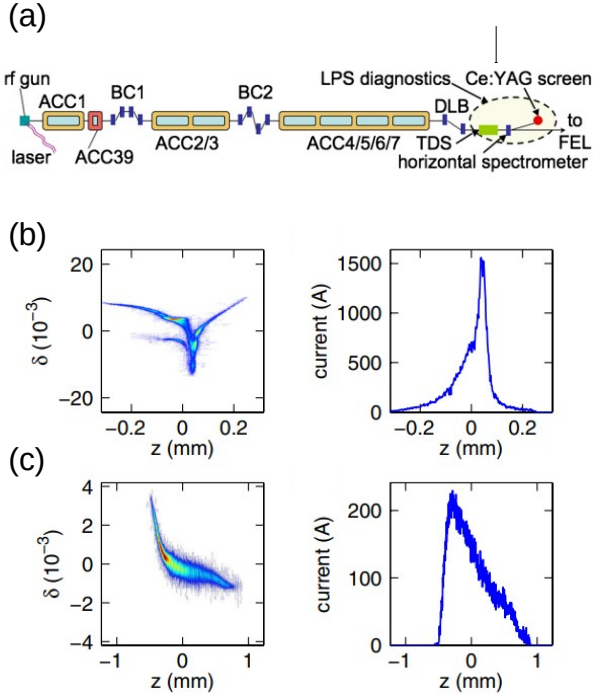
Figure 18 Experimental results on phase-space shaping using a dual-frequency linac at the FLASH facility at DESY (a) with corresponding nominally compressed beam (b) and linearly-ramped bunch generation (c) obtained via proper control of the accelerating voltage phase and amplitude of the linac sections ACC1 (1.3 GHz) and ACC39 (3.9 GHz) prior to the bunch compressor 1 (BC1) (adapted from Ref. (Piot *et al.*, 2012)).

Figure 19 Examples of LPS (a,c) and associated current distribution(b,d) obtained via a fully integrated combining the photocathode-laser shaping method discussed in Section III with multi-frequency linacs and nonlinear dispersive sections. Plots (a,b) are reproduced from (Cornacchia *et al.*, 2006), which simulated a possible operation mode of the FERMI@ELETTRA FEL. Plots (c,d) are extracted from (Tan *et al.*, 2021), which investigated the generation of shaped bunches for efficient acceleration in a corrugated-waveguide wakefield accelerator. In plot (c), the superimposed trace corresponds to the slide RMS energy spread (right label).

GHz and 3.9-GHz superconducting-linac modules and two magnetic bunch compressors was used; see Fig. 18(a). Specifically, the accelerator settings were optimized to form a linear-ramped current profile; see Fig. 18(b,c). The technique provided ample control over the shape owing to the number of variables available and produced a shaped beam at 300 MeV for future injection in the FLASH-Forward facility (Aschikhin *et al.*, 2016).

This method could be generalized in principle by introducing an arbitrary number of accelerating sections operating at different frequencies to ultimately synthesize any correlation in the LPS via a Fourier series. The generalized scheme is challenging in practice due to the limited number of RF sources and associated prohibitive cost. However, it could be implemented passively as discussed in Section V.

As a final note, we should point out that most of the time, current shaping relies on an integration of different shaping techniques as explored in (Cornacchia
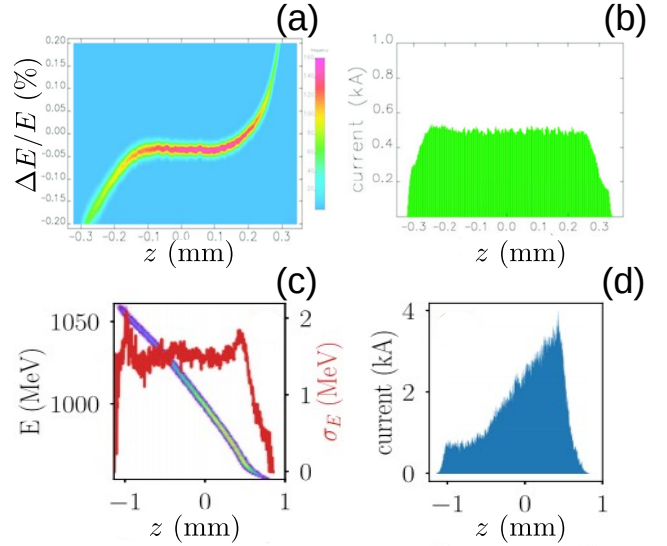
et al., 2006; Lemery and Piot, 2015). For instance, the generation of uniform beams required for cascaded harmonic lasing in X-ray FELs combines photocathode laser shaping to precisely precompensate for nonlinearities introduced in the compression process or arising from collective effects (Cornacchia *et al.*, 2006). This work especially demonstrated the generation of uniform current distribution based on numerical simulations; see Fig. 19(a,b). Likewise, numerical simulations indicate that a properly shaped photocathode laser pulse combined with a multi-frequency linac and a two-stage nonlinear compression process could provide precise control over the LPS and current distribution (Tan *et al.*, 2021), ultimately realizing the sought-after door-step distribution without the requirement for any collimation; see Fig. 19(c,d).

C. Realizing ultra-low energy spread

A recurrent topic associated with the use of bright electron beams regards the generation of ultra-low longitudinal emittance. One application of such a capability is the production of a low-energy-spread beam, ultra-short electron bunch for use in ultra-fast electron scattering experiments. There has been a renewed interest
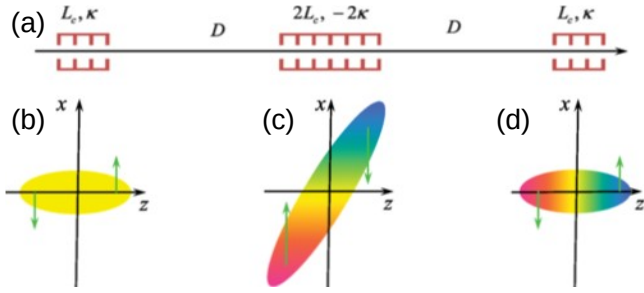
Figure 20 A TDC-based beamline for the generation of strong LPS chirps (a). The diagrams (b-d) illustrate the $x-y$ electron-bunch distribution inside each TDC with the arrows gives the direction of the transverse kick x'. The color map indicates the energy distribution. Adapted from (Yampolsky et al., 2020).

in meV energy-spread, sub-MeV electron bunches for applications in electron energy-loss spectroscopy (Egerton, 2008).

In principle, the method described above and detailed in (Zeitler et al., 2015) could be tailored to produce low energy spread. However, most of the low-energy-spread instruments have been based on dispersive collimation implementing an 'omega filter' (Tsuno and Munro, 1997). One issue associated with this class of monochromator commonly used in electron-scattering instrument is the limited beam transmission. To palliate this limitation, a lossless monochromator was proposed (Duncan et al., 2020) and demonstrated via numerical simulation the ability to produce 200-keV electron bunch with meV-level energy spread. The setup combines a photoemission electron source with a pair of accelerating cavities. The bunch-length lengthening ultimately set the final fraction energy spread $\sigma_{\delta,f}$ under Liouville's theorem as $\sigma_{\delta,f} = \sigma_{\delta,0}\frac{\sigma_{z,0}}{\sigma_{z,f}} \geq \frac{\hbar}{2}$ where $_0$ refers to the initial bunch length and fractional energy spread. (Duncan et al., 2020) note that the final kinetic energy E_K distribution along the bunch is quadratically dependent on the particles' time of arrival and radial position within the bunch $E_K(t,r) = at^2 + br^2$ such a dependence can be compensated by the pair of accelerating cavities.

D. Controlling longitudinal-phase-space (LPS) correlations

As discussed so far, the shaping of the bunch's current profile is often implemented by controlling correlation in the LPS. In this section, we focus on methods that have been implemented with the primary purpose of controlling correlation in the LPS. An example is the removal of correlated energy spread downstream of an accelerator to produce, e.g., a more monoenergetic bunch.

Although linear energy-spread control is conventionally achieved using an accelerating cavity operated off the crest, the method introduced in (Yampolsky et al.,

2020) combines several deflecting cavities to provide a power-efficient control of the LPS chirp. Specifically, the technique considers three horizontally deflecting cavities separated by a distance D and with respective deflection strengths κ, -2κ, and κ to produce an uncoupled transport matrix in (x,x',z,δ) with longitudinal 2×2 block given by

$$R_{z|\delta} = \begin{pmatrix} 1 & 0 \\ -\frac{2}{3}\kappa^2(3D+2L_c) & 1 \end{pmatrix}, \qquad (194)$$

where L_c is the transverse-deflecting cavity length; see Fig. 20. The latter equation reveals the main advantage associated with the proposed TDC-based scheme over conventional off-crest acceleration. The introduced chirp R_{65} scales quadratically with the cavity's strength κ while the off-crest acceleration scales linearly with the accelerating field. Although this simple model is not capable of altering the polarity of the chirp, introducing focusing elements between the TDC does enable such control. Additionally, the focusing elements can also be optimized to mitigate transverse-emittance degradation throughout the beamline (Yampolsky et al., 2020).
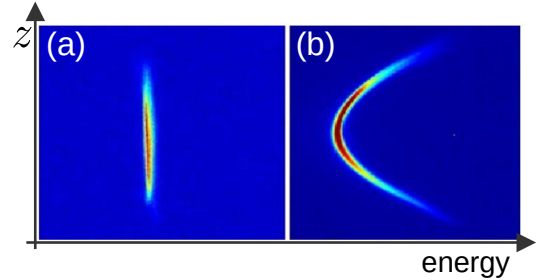
Figure 21 Example of longitudinal phases-space linearization at the LCLS facility (adapted from (Akre et al., 2008). Shown is the LPS measurement without (a) and with (b) operation of the fourth-harmonic linearizing cavity (operating at 11.424 GHz).

Accessing higher-order correlation in the LPS is more challenging especially when dealing with short (picosecond-scale) bunches, as the required RF wavelength needs to be comparable to the bunch length. However, some correction can be achieved with proper choice and interplay between RF systems operating at different frequencies or through non-relativistic effects. For example, correcting or imposing quadratic nonlinearities in the LPS can be achieved by combining a harmonic RF field to the fundamental mode, as introduced in (Smith, 1986) and later adapted for the LCLS and FLASH FELs (Flöttmann et al., 2001). The method relies on introducing a harmonic frequency so that the final fractional energy offset associated with an electron

located at z is approximately

$$\delta(z) \simeq \delta_0 \frac{E_0}{E_f} + \left(\frac{eV_1}{E_f} k_1 \sin\psi_1 + \frac{eV_n}{E_f} k_n \sin\psi_n \right) z$$
$$- \left(\frac{eV_1}{2E_f} k_1^2 \cos\psi_1 + \frac{eV_n}{2E_f} k_n^2 \cos\psi_n \right) z^2, \quad (195)$$

where δ_0 is the initial relative energy offset, ψ_1 and V_1 (resp. ψ_n and V_n) are, respectively, the phase and amplitude of the fundamental (resp. harmonic) section of the linac, and E_f the final energy. Considering the simple case of on-crest operation $\psi_1 = \psi_n = 0$, we note that cancellation of the second-order term can be accomplished when the harmonic section is operating on the deceleration phase with voltage amplitude $V_n = V_1 \left(\frac{k_1}{k_n} \right)^2$. Such an approach was first demonstrated at the Free Electron LASer in Hamburg (FLASH) facility in DESY (Harms *et al.*, 2011) by adding a third-harmonic linac ($f_n = 3.9$-GHz) to the fundamental-frequency linac ($f_1 = 1.3$ GHz). The same scheme was implemented at the European X-ray FEL (Decking *et al.*, 2020). Figure 21 showcases a measurement of the LPS-linearization process at the LCLS where a fourth-harmonic cavity ($f_n = 11.424$ GHz $= 4 \times 2.856$ GHz) is employed (Akre *et al.*, 2008). Similar, methods were investigated in the context of ultrafast electron-diffraction setups (Flöttmann, 2014).

It has also been recognized that this type of LPS linearization can be implemented without the need for a harmonic field when dealing with non-ultrarelativistic bunches. Such a scheme was introduced in (Krafft, 1996) were it was shown via numerical simulations and experimentally confirmed that a proper control of the phase and amplitude of a buncher and capture cavities could control the R_{56} and T_{566} in the CEBAF injector (Wang *et al.*, 1998). A similar technique was explored theoretically and via simulation for an ultra-fast electron source using an RF photoinjector; see (Zeitler *et al.*, 2015). Specifically, the authors show that expanding the beam after the electron source enabled a higher-order correction of the longitudinal focus by a subsequent accelerating cavity that is operated at the same frequency as the electron gun. Although the method was implemented as part of a ballistic-compression scheme to demonstrate the generation of sub-fs bunches at low energy, it could in principle be extended to higher energy when combined with standard compression beamlines.

Additionally, introducing high-order longitudinal dispersion, as discussed previously in Section IV.B can also be used to control LPS nonlinearity. Such a method was successfully implemented in a high-power energy-recovery linac (Piot *et al.*, 2003) and supported the first demonstration of same-cell energy recovery of a high-power electron beam with increased energy-spread after its interaction in an FEL oscillator (Neil *et al.*, 2000); see Fig. 22. Higher-order correction using octupole mag-

nets was also implemented in the JLab 10-kW FEL facility (Neil *et al.*, 2006).
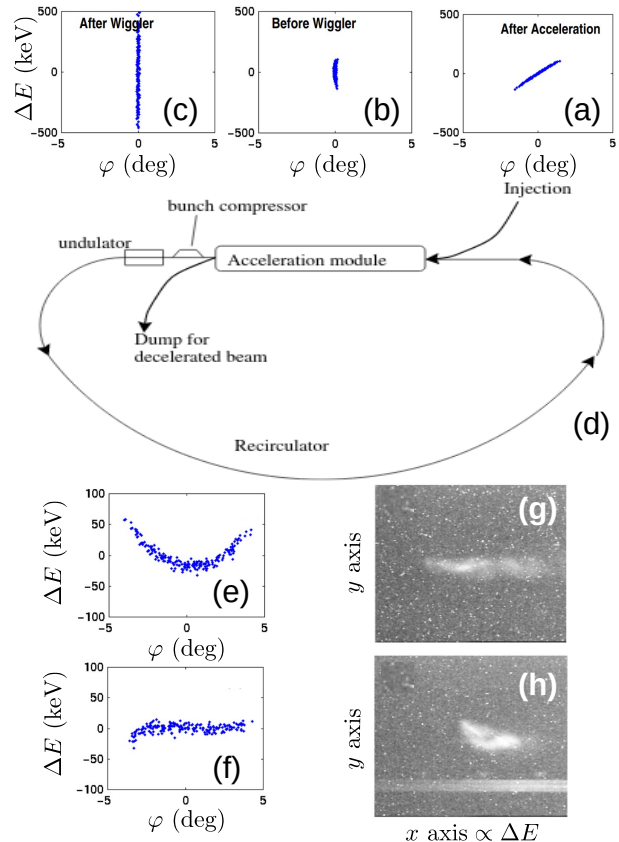
Figure 22 Nonlinear control of LPS with high-order longitudinal dispersion in JLab IR-FEL Demo for two settings of the longitudinal lattice along with resulting measured energy spread measured after energy recovery [images (g) and (h)]. The simulated LPS are displayed at acceleration-module exit (a), before(b) and after (c) the FEL iteration, and after energy recovery for the cases of linear (e) and nonlinear (f) correction (via independent control of the R_{56} and T_{566} of the recirculator) of the LPS. . The measured $(\delta E, y)$ distributions shown in images (g) and (h) correspond respectively to simulated LPS presented in (e) and (f). Adapted from Ref. (Piot *et al.*, 2003).

Finally, a laser-pulse shaping similar to the one discussed in Sections II and IV.B initially proposed in (Cornacchia *et al.*, 2006) can also be combined with a beamline with nonlinear longitudinal dispersion to control the final LPS of a relativistic bunch. An experimental demonstration of the concept was performed at the FERMI@ELETTRA accelerator using a 1.4-GeV beam (Penco *et al.*, 2014). The photocathode laser was temporally shaped to precompensate nonlinear correlations nominally accumulated during the acceleration and two-stage bunch compression processes. The measured LPS upstream of the FEL beamline is uncorrelated and linearized; see Fig. 23. In the process, the authors demon-

strated that the beam dynamics of the ramped bunch did not significantly affect the beam transverse emittance compared to their nominal operating point.

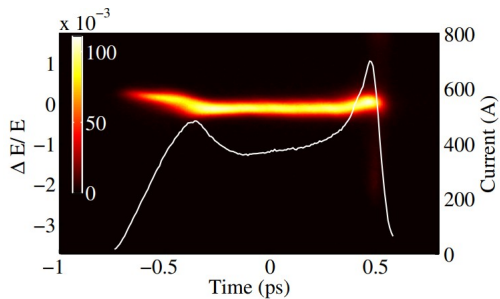Figure 23 Experimental demonstration of control of LPS correlation at at the FERMI@ELETTRA accelerator using a 1.4-GeV beam. Electron-bunch LPS (image) with superimposed current profile (white line, right axis). These data were obtained using a shaped photocathode laser with ramped temporal profile configuration. The 300-fs bunch core has a nearly constant incoherent energy spread of about $\sigma_E \simeq 150$ keV. From (Penco *et al.*, 2014).

E. Transverse phase-space control

We now examine techniques to transversely shape the beam expanding on the methods described in Sec. III.D. This type of shaping is ultimately impacted by collective effects so that devising shaping methods implementable at high energy is beneficial. For instance, producing uniform electron beams with uniform transverse distribution is critical to a broad range of applications including beam irradiation of a target, e.g., to produce X rays, as it provides a uniform dose and mitigates thermo-mechanical stress on the target (Pasquali *et al.*, 2019). Likewise, using bright electron beams with uniform transverse distribution offers an effective way of improving the performance and output power of tapered X-ray FELs (Emma *et al.*, 2014a). Finally, transversely shaped beams combined with the mask-based shaping techniques discussed earlier can lower beam losses on the mask. These transverse-shaping methods have gained considerable interest, as their possible combination with the phase-space-exchanging techniques discussed in Section VI could enable control of the current distribution with unprecedented precision and versatility.

1. Nonlinear transformations

The idea of using nonlinear transformations parallels the early discussion related to LPS manipulation. Such an approach was first explored in (Merminga *et al.*, 1991) to remove the beam tail via nonlinear focusing (implemented with sextupole magnets) as part of a "nonlin-

ear collimation" scheme proposed for future linear colliders. Likewise, the use of nonlinear focusing to form uniform beam distributions was discussed in (Meads, 1983). Numerical simulations (Kashy and Sherrill, 1987) demonstrated the formation of uniform distribution using an octupole magnet implemented in practical beamline. Likewise, (Batygin, 1993) derives the nonlinear force required to redistribute an incoming beam nonuniform distribution into a uniform distribution using a beamline composed of a multipole lens followed by a drift space. Further work discussed in (Meot and Aniel, 1996) analytically shows that the odd-order multipole fields, such octupole and dodecapole components, are required for transverse uniformization. Efforts to produce transversely uniform distributions over appreciable distances along an accelerator beamline has been investigated for possible use in combination with undulator tapering in X-ray FELs (Jiao and Cui, 2015). Over the years several experiments have been conducted. An early demonstration experiment using a 200-MeV H^- beam was performed at BNL (Tsoupas *et al.*, 1991), and an example of distribution measured at the NASA Space Radiation Laboratory (NSRL), using an electron beam at BNL (Tsoupas *et al.*, 2007), appears in Fig. 24. Similarly, uniform beam distributions are sometime used in medical accelerators.
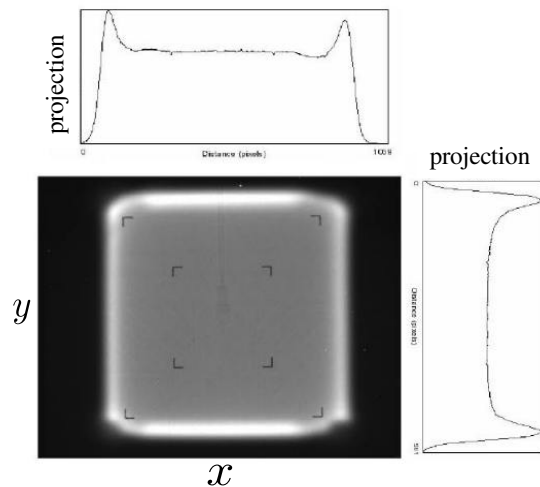
Figure 24 Measured beam distribution obtained downstream of a uniformizing beamline composed of octupole magnets. Adapted from (Tsoupas *et al.*, 2007).

Reference (Yuri *et al.*, 2007) develops a theoretical framework to understand the formation of shaped beam profiles using a nonlinear focusing beamline. Following this work, the phase-space coordinates downstream of the beamline under consideration can be expanded as a truncated power series of the initial phase-space coordinate upstream of the beamline (x_0, x_0') via a transformation
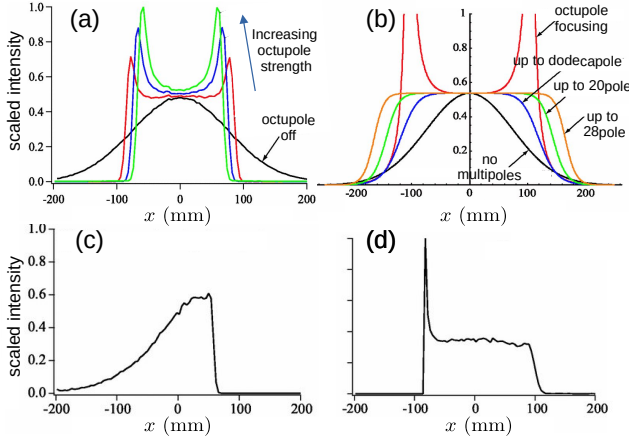
Figure 25 Example of computed transverse profiles downstream of a nonlinear-transport beamline composed of multipole field. The upper row demonstrates (a) the impact of an octupole magnet with increasing strength, and (b) the effects associated with various multipole orders on the horizontal beam distribution. The lower row showcases the use of two sextupole magnets to control the beam distribution [plots (c) and (d) correspond to the distributions recorded at two locations downstream of the sextupole magnets]. Adapted from (Yuri *et al.*, 2007).

of the form

$$\begin{pmatrix} x_t \\ x'_t \end{pmatrix} = \begin{pmatrix} (M_{11} - \frac{\alpha_0}{\beta_0} M_{12})x_0 - M_{12} \sum_{n=3}^{\infty} \frac{K_{2n}}{(n-1)!} x_0^{n-1} \\ (M_{21} - \frac{\alpha_0}{\beta_0} M_{22})x_0 - M_{22} \sum_{n=3}^{\infty} \frac{K_{2n}}{(n-1)!} x_0^{n-1} \end{pmatrix}.$$

Imposing Liouville's theorem as discussed in Section II.C.1 indicates that the final distribution along one of the position coordinates can be written as

$$\rho_t = \frac{\rho_0}{M_{11} - \frac{\alpha_0}{\beta_0} M_{12} - \sum_{n=3}^{\infty} \frac{K_{2n}}{(n-2)!} x_0^{n-1}}. \quad (196)$$

This equation provides the values of the magnetic multipole strength K_{2n} required to transform the initial distribution into the desired one. The work also extends previous methods by investigating the effect of the even-order multipole fields on beam uniformization and demonstrating the feasibility of uniformization of a Gaussian beam for even-order fields, or considering incoming transversely asymmetric beams. The authors conclude that using a combination of even-order multipoles may ultimately provide uniform beam over a larger transverse section; see Fig. 25(b).

Finally, although the emphasis has often been on the generation of uniform beams, the technique can also produce beams with transversely shaped distribution. For instance, Fig. 25 demonstrates the generation of a ramped horizontal distribution obtained as an intermediary step to generate a uniform distribution. Furth development demonstrated the generation of patterned hallow beams (Yuri *et al.*, 2019). This latter capabilities could be taken advantage of and combined with phase-space-exchange methods discussed in Section VI.
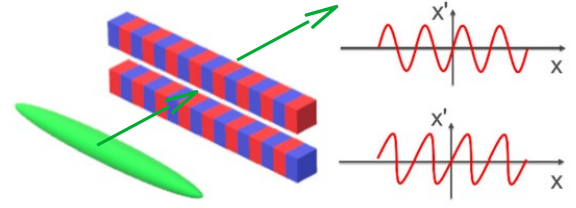

Figure 26 Configuration for transverse beam shaping using a "transverse"-wiggler magnet (left) and horizontal phase-space (right) immediately downstream (upper plot) the wiggler and after a subsequent drift (lower plot). Adapted from (Ha *et al.*, 2019a).

Another method toward shaping the transverse phase space is to use a transversely periodic magnetic field to modulate the electron divergence across the transverse beam distribution as proposed in (Ha *et al.*, 2019b) and illustrated in Fig. 26. Taking the example of the horizontal phase space (x, x') passing through a set of transverse wigglers − a wiggler oriented transverse to the beam direction − with wiggler parameter $K_{w,i}$, the divergence of an electron at initial position x_0 will be given by

$$x' = x'_0 + \sum_{i=1}^{n} K_{w,i}(\sin k_i x), \quad (197)$$

where $k_i \equiv 2\pi/\lambda_i$ is the transverse wavevector with λ_i being the ith wiggler period. A proper choice of k_i and $K_{w,i}$ allows the synthesis of any correlation in the (x, x') phase space that could produce a tunable profile along the horizontal spatial or divergence direction.

2. Interceptive beam shaping: beyond binary masks

As discussed in Section IV.B, an interceptive "binary" mask with optimized contour provides a versatile and simple tool to shape the beam transverse distribution but is rarely used owing to the accompanying losses. These losses are especially problematic when the shaping technique is implemented in a high-current accelerator as they can lead to, e.g., radiological activation and damage of beamline hardware.

For ultra-low emittance beams where quantum coherence is achieved and the beam can be described by its wavefunction, optical (photon) techniques have been adapted to shape the beam with a high degree of control, e.g., in electron microscopes (Nagayama, 2011; Shiloh *et al.*, 2019). An example of such a manipulation is the generation of "vortex" electron beams carrying orbital angular momentum (Uchida and Tonomura, 2010) using a nano-engineered spiral-like phase plate made of stacked graphite thin film.
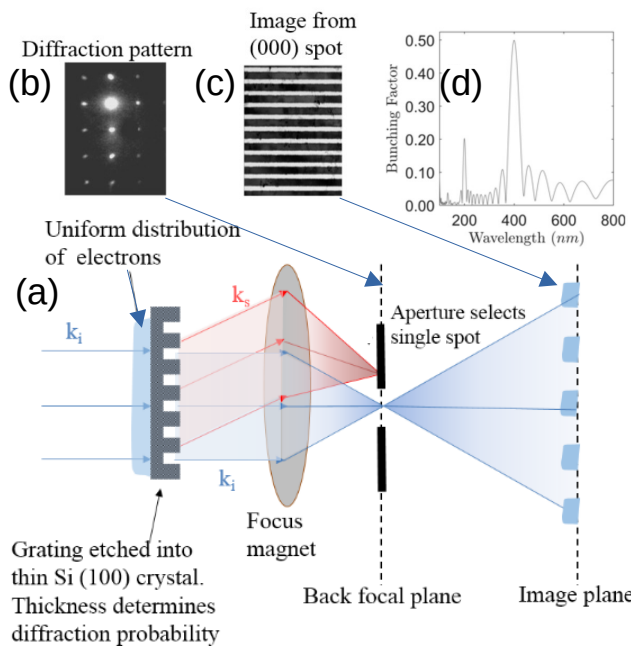
Figure 27 Principle of diffraction modulation to produce a nano-patterned beam (a) with diffraction pattern downstream of the grating (b) and final transverse distribution of the transmitted beam (c) with associated spatial spectrum along the vertical coordinate (d). Figure adapted from (Graves *et al.*, 2019).

Although electron beams produced in conventional accelerators are incoherent, ultralow-emittance relativistic electron beams can be manipulated with variable-transmission masks, as discussed in (Nanni *et al.*, 2018), provided their transverse coherent $L_\perp = \frac{\hbar}{mc}\frac{\sigma_\perp}{\varepsilon_\perp}$ (van Oudheusden *et al.*, 2007) is comparable or larger than the intercepting-mask crystalline structure (where $\sigma_\perp$ and $\varepsilon_\perp$ respectively refer to the transverse beam size and emittance). In this reference, a diffraction contrast modulation technique was proposed to realize nano-patterned electron beams. The method relies on electron diffraction in a transmission grating with periodically variable thickness. The grating structure spatially modulates the fraction of electrons diffracted into a particular Bragg peak. An aperture selecting the transmitted beam or one of the Bragg peaks then results in a final modulated beam. The method was demonstrated at SLAC using a 2.3-MeV 1-pC electron beam from an ultrafast electron-diffraction setup (Graves *et al.*, 2019). The beam was diffracted through a thin lithographically patterned Si membrane [Fig. 27(a)] to produce a vertically modulated beam distribution [Fig. 27(d)] with a spatial period of ~ 400 nm and an associated bunching factor of 0.5 [see spatial spectrum in Fig. 27(b)]. In this experiment, the unwanted electrons were scattered into the (220) Bragg peak, and the transmitted electron associated with the (000) peak was used to produce the final 300-fC modulated beam

[Fig. 27(c,d)]. This masking technique has some limited tunability accomplished by varying the tilt angle of the grating and selecting the Bragg peak to diffract the beam where desired. This method ultimately suffers from the mask's limited lifetime due to atomic displacement over long periods of time. The achieved sub-micrometric spatial period combined with demagnifying optics could produce beams with nanometric modulations.

V. LONGITUDINAL SHAPING WITH BEAM SELF-GENERATED FIELD

Self-generated fields are usually considered to be an obstacle to improving the beam quality. While fully decoupled Gaussian beam or beams having linear correlations are desirable, most self-generated fields introduce non-linear correlations on the beam. The non-linearity on the beam's phase space not only increases the emittance, but it introduce intrinsic limitations on the manipulation of the beam. Overcoming the degradation of the beam quality caused by the self-fields is one of the major research topics in accelerator physics. In the best case, self-fields might enable manipulation of the beam and avoid its degradation simultaneously.

This section focuses on shaping mechanisms that are based on beam-generated fields. Most of the methods introduced in this section are for longitudinal shaping because the use of self-generated fields for transverse shaping is scarce. We first discuss longitudinal profile shaping using the space-charge field and CSR in Sec. V.A and Sec. V.B. Then we describe the mechanism using wakefields in Sec. V.C. Each subsection provides a short description of the results and their applications.

A. Shaping profiles using space-charge field

The space-charge field has two characteristics. The first is that it depends on the beam's spatial distribution. If the beam has a symmetric profile, the field strength is symmetric to the beam center. The second is that it acts in a way to defocus or lengthen the bunch when the beam is not density modulated because electrons repel each other. These two characteristics give us several methods for using the space-charge field for shaping. The following subsections describe these shaping methods and give relevant results.

1. Space-charge field with a single bunch

Due to the field's dependence on the beam's spatial distribution, the usefulness of the space-charge field from a single bunch is limited. However, we can imagine two different uses of the space-charge field from a single bunch.
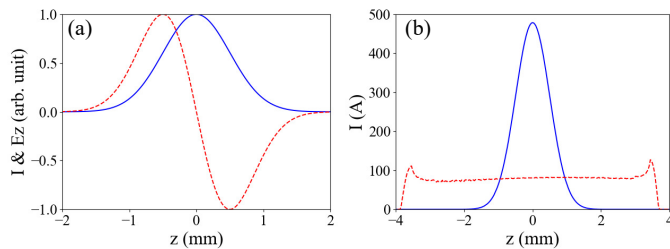
Figure 28 Current profile shaping using space-charge force. (a) Gaussian longitudinal profile (solid) and corresponding longitudinal space-charge field from Eq. (198). (b) Longitudinal profile of the initial particle distribution (solid), and the longitudinal profile after a 3-m-long drift (dash). (b) is a simulation result of a beam with 1-nC charge and 3-MeV energy.

First, the space-charge force applied to the bunch is not linear except for a few profiles. Thus, similar to the transverse shaping using an octupole magnet (Yuri *et al.*, 2007) (see Sec. IV.E.1 for further details), a symmetric-nonlinear space-charge field can induce a symmetric change in the profile. For example, in a cylindrical beam with a uniform transverse distribution and a longitudinal distribution of $\lambda(z)$, the space-charge field can be expressed as in Eq. (136). If the initial longitudinal distribution is Gaussian, then the axial space-charge field will be

$$E_z(z) = \frac{e}{4\pi\epsilon_0\gamma^2}\frac{N_b z}{\sigma_z^3}\left(1 + 2\ln\frac{b}{a}\right)\exp\left[-\frac{z^2}{2\sigma_z^2}\right]. \quad (198)$$

The Gaussian profile and corresponding space-charge field in Eq. (198) are visualized in Fig. 28(a). In the case of the example shown in Fig. 28, all particles move outward due to the space-charge field as the beam drifts, and the bunch length will increase. The space-charge field in this example is nonlinear and has its maximum strength at around ± 0.5 mm. Thus, particles initially located at ± 0.5 mm gain more momentum than particles in $|z| > 0.5$ mm, and eventually overtake these particles in $|z| > 0.5$ mm. This situation is very similar to octupole-based transverse shaping, see Fig. 24. Thus, if the beam traverses a long enough distance (i.e., large enough R_{56}), then this Gaussian distribution would evolve into a uniform distribution. Figure 28(b) shows a particle tracking simulation result, which demonstrates the idea. As we can see from Eq. (198), the space-charge field gets weaker as the beam energy increases. Thus, the change arising from the space-charge field requires a large R_{56} to convert momentum change to position change. This requirement limits the use of the space-charge field to near the electron gun; thus, so far this type of shaping is not demonstrated experimentally.

We can consider a similar type of shaping for generating asymmetric bunches, but it is necessary to introduce an originally asymmetric profile. Thus, using the space-charge field for asymmetric control is not an attractive option, but the understanding of this mechanism may help to improve the shaping quality of emission-based shaping methods introduced in Section III.

The other usage is exploiting the symmetric lengthening feature of the space-charge field. In (Luiten *et al.*, 2004) they suggested a new concept, the so-called blow-out regime, to generate a 3D ellipsoidal distribution with uniform density. In this concept, an ultra-short laser pulse with an appropriate radial distribution ($\lambda_\perp(r) \sim \sqrt{1 - (r/A)^2}$ where A is the maximum radius) shines on a photocathode. This generates an ultra-thin sheet of electrons that evolves into a uniform ellipsoidal shape. Here, the ultra-short pulse length plays a key role because it allows that all particles in the bunch to experience a similar space-charge field strength. This allows the beam to evolve into a uniform ellipsoid regardless of the original pulse shape. The imperfection here will appear as a soft edge of the ellipsoid. In (Luiten *et al.*, 2004) they ran simulations using a 100-pC beam with 1-mm radius and 0.4-eV energy at the cathode surface. The electric field at the cathode surface is assumed to be 100 MV/m, and the incident laser pulse is assumed to have a Gaussian profile with 30-fs FWHM. The original disk-shaped distribution has evolved into a nearly elliptical one that is slightly less than ideal due to the imperfection.

In (Rosenzweig *et al.*, 2006), further theoretical work was performed to combine the blow-out regime and emittance compensation. Later, this blow-out regime was experimentally demonstrated by (Musumeci *et al.*, 2008; O'Shea *et al.*, 2011; Piot *et al.*, 2013). The first experiment in (Musumeci *et al.*, 2008) demonstrated the blow-out regime using 15-pC beam; see Fig. 29. Later, the experiment in (Piot *et al.*, 2013) demonstrated the blow-out using a higher charge level (0.5-nC) with a Cs_2Te cathode, and it showed well-shaped bunches. However, the method required a disk-shaped beam, making it difficult to reduce the laser spot size at the cathode to obtain target charge. This limits the intrinsic emittance, which scales with laser spot size, Eq. 168. The emittance demonstrated so far does not show a clear advantage of this method. There is another emission scheme, based on a cigar-beam, introduced in (Rosenzweig *et al.*, 2019) that achieved lower emittance. While the blow-out regime uses a short but large radius laser profile, which produces a longitudinal expansion of the beam, the cigar-beam regime uses a long but small radius laser profile which produces radial expansion of the beam. Simulations showed a noticeable improvement in beam brightness of the cigar compared to the blow-out regime. Due to these limitations, research efforts were moved to using ellipsoidal laser profiles instead of the blow-out regime to generate an ellipsoidal beam; see (Khojoyan *et al.*, 2014). See Section III for further details of laser shaping.
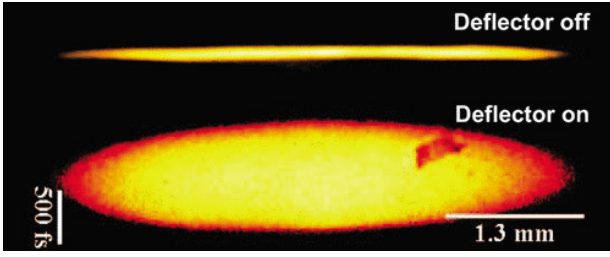
Figure 29 Experimental demonstration of ellipsoidal beam generation using blow-out method. The beam images were measured after a deflecting cavity. The beam was focused vertically, and the deflecting cavity kicked the beam vertically. x and y axes correspond to the beam's horizontal and longitudinal distributions, respectively. The beam's energy was 3.75 MeV and the charge was 15 pC. From (Musumeci et al., 2008).

2. Space-charge field with a few bunches

When more than a single bunch exists, the bunches mutually repel each other via the space-charge field. If there are extra bunches in front of and behind the target bunch, the electromagnetic field felt by the target bunch will depend on these extra bunches and the target bunch's self-field. Here, the charge ratio of each bunch and the spacing between the bunches can be used as knobs to control the space-charge field applied to the target bunch.

In (Lu et al., 2018), this concept was used to compress an electron bunch. During the experiment, they generated a low-charge main bunch and two extra bunches for space-charge field shaping using a laser beam splitter and α-BBO crystals. The main bunch was located between two field-shaping bunches. These two field-shaping bunches push the target bunch, which generates a negative longitudinal chirp of the target bunch's longitudinal phase space (LPS). Note that the negative chirp means the head of the bunch has lower energy. Because the space-charge field outside of the bunch shows the usual $1/r^2$ tendency, they located the target bunch at the middle so that the nonlinearity of the space-charge field coming from the field-shaping bunches almost cancels out. This provided a nearly linear longitudinal chirp to the target bunch; the relative separation was adjusted by using a laser delay line. The laser was injected into the RF photocathode gun that accelerated the beam to about 3.4 MeV. Due to the beam's low energy, a few meters of drift provided enough R_{56} for ballistic bunching. Note that ballistic bunching means that the bunch compression is accomplished by the velocity difference of the particles in a drift. Velocity bunching is a more general term, but the ballistic bunching is used when acceleration is not included in the compression.

During the experiment, they fixed the target bunch's charge level to 50 fC while varying the field-shaping
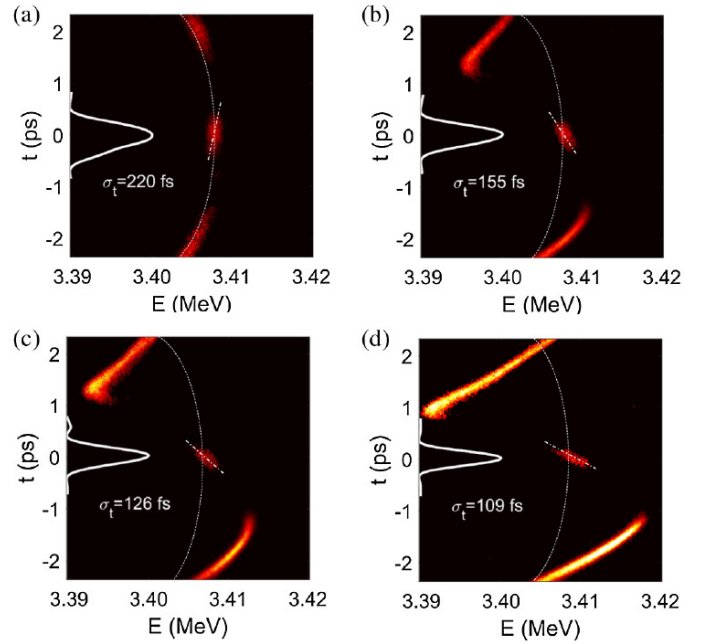
Figure 30 Space-charge field shaping for bunch compression. Measured longitudinal phase spaces are displayed. Each panel corresponds to different charge level of field-shaping bunches. (a-d) correspond to charges of 0.2, 3.2, 4.5, and 6.7 pC, respectively. $t > 0$ corresponds to the head. From (Lu et al., 2018).

bunches' charges from 0.2 to 6.7 pC. Here, the charges of the field-shaping bunches were used to control the field strength applied to the target bunch. They observed bunch compression with an LPS measurement as shown in Fig. 30. Also, the main bunch length was controlled from 220 fs to 109 fs by varying the charge. As we can see from Fig. 30(d), there is still room for further compression.

Another interesting point of this work is stability. In the case of chicane compression, the RF jitter of the cavity controlling the chirp affects the longitudinal chirp or the beam energy. This RF jitter, in turn, result in an increase of the bunch length jitter and arrival time jitter. On the other hand, all three bunches originated from a single laser pulse, and there are no RF cavities other than the gun in this scheme. Thus, it can provide better stability than other methods. This is an interesting method for applications requiring a short bunch with good stability such as ultra-fast electron diffraction (UED), ultra-fast electron microscopy (UEM), or FELs.

A similar method was also experimentally demonstrated for transverse phase space. In (Rihaoui et al., 2009), they introduced extra bunches surrounding a central main bunch in the transverse space; see Sec. III.D.2. Due to the interaction via the space-charge force and the beam's distribution change by solenoid focusing, these extra bunches either introduced extra focusing to the

main beam or changed the beam's distribution to a shape that was other than a Gaussian distribution. This could be an interesting option for a system having difficulty using magnets. For example, locating a solenoid near the cathode of superconducting guns is difficult, but transverse control is necessary to preserve the beam's quality. These extra bunches may provide all-beam-focusing to the target beam, so it may be able to locate a solenoid further downstream while preserving the beam quality.

Although the experimental demonstrations discussed in this section are limited to one-dimensional density control, we can imagine appling this method to more general bunch shaping applications. For example, nonlinearity is a key for asymmetric bunch shaping, as we saw in Section IV. We can imagine exploiting nonlinearities that space-charge force introduces. Bunch charges and spacing can be used to control the shape of LPS. Controlling the space-charge field of the main bunch by using additional bunches which may provide a new way to apply a laser-shaping technique to high-charge bunches.

We should note that beam shaping with extra bunches always has the problem of eliminating the extra bunches after they are no longer needed for shaping. The elimination process may introduce additional disadvantages that make the method less beneficial than described so far. A beam cutting process may happen using a mask and an additional beamline element such as chicane or RF deflecting cavities. However, these beamlines can easily add extra timing jitter to the beam, negating one of the benefits of this method.

3. Space-charge field with multiple bunches: space-charge oscillation

We now consider more than three bunches by providing a density modulation at the beginning. This initial density modulation introduces oscillatory behavior in the longitudinal phase space, which is similar to the plasma oscillation that ions and electrons form. In the electron bunch case, the initial density modulation and bunch-to-bunch interaction via the space-charge force introduces the same oscillation.

This behavior is shown in the left panel of Fig. 31. The initial density modulation lengthens each micro-bunch, due to the space-charge force, so that the modulation becomes weaker as the beam drifts until it finally disappears at a certain distance. Once the beam passes this point, the particles in each micro-bunch continue to move in the same direction, due to their momentum. These particles build another density modulation whose phase is 180° off from the original modulation, thus the direction of the space-charge force now is reversed. At a certain point, the momentum direction changes, due to the reversed space-charge force, and this density modulation starts to oscillate. The oscillation frequency of this oscillation is
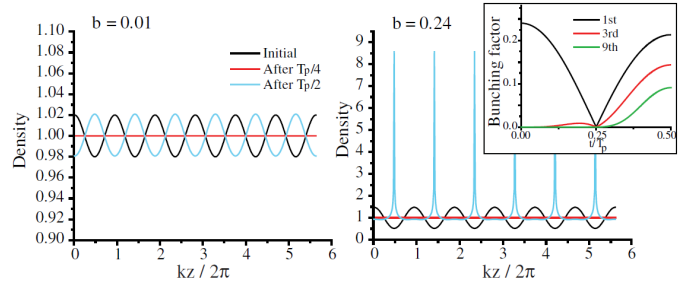

Figure 31 Nonlinear plasma oscillation of the beam to generate a bunch train with a high bunching factor. Beam's longitudinal density profiles and their evolution are displayed. The left panel started with a bunching factor of 0.01, and the right panel started with a bunching factor of 0.24. These are numerical solutions of the system of equations describing a coasting beam's density profile evolution. From (Musumeci et al., 2011).

equal to the plasma oscillation's frequency in Eq. (146). The change of the relative energy spread induced by the longitudinal space-charge force can be written as

$$\Delta\delta_{LSC} = 4\frac{\gamma^3\omega_p b}{ck}, \quad (199)$$

where k is wave number ($= \frac{2\pi}{\lambda}$), and b is bunching factor $= |\int \lambda(z,0)e^{ikz}dz|$; see (Musumeci et al., 2013).

In the linear oscillation situation described above, a sine-like shape on the phase space does not break and keeps oscillating because the energy modulation induced by the initial density modulation is small so that all particles cannot arrive at the density peak area before the momentum is reversed (i.e., $R_{56}\delta \ll \lambda/4$). However, when the modulation becomes big enough, the space-charge force can induce wave-breaking where the momentum modulation is large enough to cause all particles to reach the density peak area (i.e., $R_{56}\delta \approx \lambda/4$).

It is also possible to understand this process as the summation of harmonics. In the case of linear oscillation, only the fundamental mode governs the oscillation. However, as the initial density modulation gets larger, the oscillation process picks up more harmonic components of the fundamental wavelength. After half of the plasma period, these harmonic components sum in-phase and generate current spikes. This viewpoint can be seen from the equation below showing the electron density. More details are described in (Musumeci et al., 2013).

$$\lambda(z,t) = \left[1 + \cos\left(\omega_p t\right)\sum_{m=1}^{\infty} m c_m(t)\cos\left(mkz\right)\right], \quad (200)$$

where $c_m(t) = \frac{(-1)^{m+1}}{m}\frac{2}{\alpha(t)}J_m\left[m\alpha(t)b\right]$, and $\alpha(t) = 2\sin^2(\omega_p t/2)$.

Although this concept can be applied to any coasting beam, it is impractical for a high-energy beam because

of its long plasma wavelength. Thus, it was applied to an electron gun by several groups (Harris *et al.*, 2007; Neumann *et al.*, 2009, 2003). Here we describe more details of the experiment done by (Musumeci *et al.*, 2011), which successfully demonstrated the generation of current spikes.

They used three α-BBO crystals to generate eight equally spaced laser pulses to form a 1-THz ($\lambda = 300$ μm) pulse train. They tested the bunch charge up to 40 pC, and the gun solenoid was used to focus the beam. To demonstrate the method, they varied the laser intensity, which controls the charge, to change the phase advance of the oscillation. The phase advance determines the final modulation density. The modulation was measured with a deflecting cavity at the end of an experimental beamline. The solenoid field was fixed, and the gun phase was set to $35°$ to preserve the original modulation period. The result is summarized in Fig. 32.

According to their simulation, the phase advance of the lowest charge case (1.6 pC) was about 0.18π; see Fig. 32. Thus, it still showed the original modulation and all eight peaks appeared at the end of the beamline. When the charge was increased to 3.9 pC, the modulation was washed out; this charge level corresponds to the phase advance of 0.25π in their simulation. As the charge increased further, the phase advance became slightly lower or higher than 0.5π, so a density modulation appeared again, but now there were only seven peaks because of the π phase shift of density peaks in a given total bunch length.

A few years later, this method was revisited and experimented at a higher charge level to consider an application to terahertz (THz) radiation. In (Zhang *et al.*, 2016) they set up an experimental beamline similar to (Musumeci *et al.*, 2011) and added a chicane to broaden the frequency tuning range. They used a charge up to 1 nC, and results were very similar to the ones from (Musumeci *et al.*, 2011). They controlled both the charge level and the solenoid focusing strength to control the phase advance. The initial modulation was washed out with a low phase advance while the modulation was shifted by $180°$ when the phase advance was 0.25π.

In addition to confirming the oscillation at a higher charge level, they also tried to generate THz coherent transition radiation (CTR) using a foil. They adjusted two tuning parameters to control the modulation frequency and measured the spectrum of CTR using a Michelson interferometer equipped with a Golay cell (Golay, 1947). The first parameter was the launching phase of the gun. Depending on the launching phase, the bunch had different longitudinal chirps. This eventually introduced different compression ratios for ballistic bunching in the low energy area. They varied the phase from $25°$ to $50°$, and the frequency of CTR was varied from 1 THz to 0.7 THz, correspondingly. The second parameter was the chicane's bending angle, which controlled R_{56}. Com-
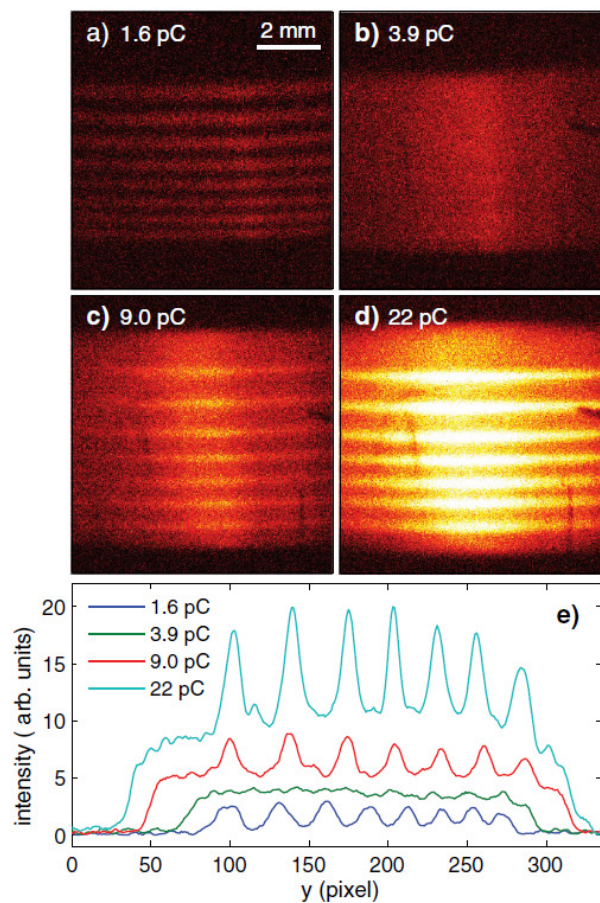
Figure 32 Experimental demonstration of bunch train generation using nonlinear plasma oscillation. (a-d) shows temporally streaked beam images, and corresponding longitudinal profiles are shown in (e). Streaking direction for beam images is vertical. From (Musumeci *et al.*, 2011).

pared to the launching phase, this second parameter provided a much wider tuning range. They varied R_{56} from 0 to 18 mm, and it changed the frequency from 0.8 THz to 1.6 THz.

Because of the high charge level, they achieved 2 μJ of THz energy from CTR. They expected that a 30-mm-long quartz tube with 0.3- and 0.4-mm inner and outer diameters, respectively, would provide 8-MW, 1.4-mJ THz radiation at 0.7-THz frequency. This intense and tunable THz radiation from a compact beamline can be useful for a typical spectroscopy-type equipment, pump-probe measurements in XFELs, or THz wakefield accelerators.

4. Space-charge field with multiple bunches: longitudinal cascade amplifier

When a momentum modulation from an initial density modulation is strong enough, the initial density modulation can be amplified with an appropriate R_{56}. This
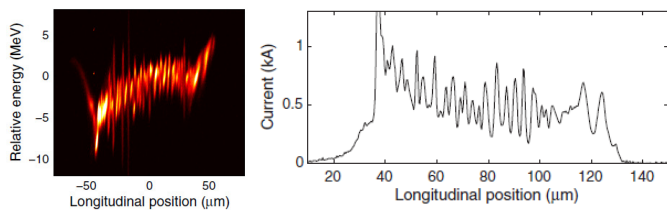
Figure 33 Measurement of microbunching instability. The left panel shows the measured longitudinal phase space without a laser heater. The strong modulation originated from the shot noise and its amplification. The right panel shows the corresponding current profile. Adapted From (Ratner *et al.*, 2015).

phenomenon was studied extensively in the early 2000s to understand the significant beam quality degradation and radiation generated from the beam. This is called the microbunching instability and many theories and experimental papers were published to explain and suppress this phenomenon; for example, (Heifets *et al.*, 2002; Huang *et al.*, 2010; Huang and Kim, 2002; Lumpkin *et al.*, 2009; Marinelli *et al.*, 2011; Prat *et al.*, 2017; Ratner *et al.*, 2015; Saldin *et al.*, 2002a, 2004, 2002b; Spampinati *et al.*, 2014; Wu *et al.*, 2008). One of them, (Ratner *et al.*, 2015), showed a picture of the longitudinal phase space with various conditions, and Fig. 33 clearly shows the modulation amplified by the gain process; see Sec.II.D.5.

As described in Sec. II.D.5, (Saldin *et al.*, 2002a) derived a simple equation to estimate the gain (G) of the density modulation, which is the ratio of final to initial modulation amplitude. Longitudinal space-charge fields introduced energy modulation to the beam, as appeared in Eq. (156) as an impedance term. Then, R_{56} of the given beamline converted an energy modulation to a density modulation. For the microbunching instability, the longitudinal space-charge force from the shot noise makes the gain higher than 1 and amplified this unwanted shot noise. Thus, an effort, such as using a laser heater, was made to reduce the gain [see the σ_δ term in the exponent of Eq. (156)].

On the other hand, (Schneidmiller and Yurkov, 2010) introduced a concept to use this amplification of shot noise as a way for generating a high-frequency bunch train; it was called the longitudinal space-charge cascade amplifier (LSCA). In this concept, they used a drift with a focusing channel to introduce a momentum modulation from the density modulation, then a chicane is used to introduce an R_{56} optimized for the modulation conversion with maximum gain. This focusing channel and chicane form a single cell of the amplifier, and each cell provides gain to the density modulation. Thus, they used a cascade of cells to saturate the density modulation. The numerical example they provided shows that the gain per cell reaches more than 40 for a beam with an energy of 3 GeV, a peak current of 2 kA, a normalized emittance of

2 μm, and an energy spread of 0.3 MeV, which is close to typical FEL parameters. Thus, two cascades would provide a $G > 1,000$. This is high enough to saturate the shot noise for this example case.

The LSCA method was experimentally demonstrated in 2013 by (Marinelli *et al.*, 2013). The experiment amplified shot noise using three chicanes with drifts between them. The amplified modulation was used to generate radiation by using an undulator, with an undulator parameter of 0.58, a period of 1.9 cm, and a total of 11 periods. Each chicane provided R_{56} of 4 mm, 2.5 mm, and 1.5 mm, in order, and the chicanes were 2 m apart. The initial momentum modulation was accumulated for a 10-m-long drift before the first chicane, and the beam energy was 72 MeV. The spectrum of undulator radiation was measured. They used a photodiode detector to measure the integrated intensity gain. With a 12-pC bunch, the average gain in the integrated intensity over the incoherent background was 600. When they measured the gain of on-axis radiation and only considered the bunch charge contributing to the coherent radiation, the gain went up to 1.5×10^4, which was in good agreement with their estimated gain of 2.5×10^4 from the linear theory.

This method is considered for FEL facilities to generate high-power coherent radiation, large radiation bandwidth, or attosecond pulses as described in (Schneidmiller and Yurkov, 2010) and (Dohlus *et al.*, 2011). On the other hand, another use of this cascade method was proposed by (Ratner, 2013) wherein they suggested using the cascade method to achieve a high cooling rate for coherent electron cooling. Here, the modulation imprinted by an hadron beam onto the electron beam is amplified with the cascade amplifier, and the field from the amplified modulation of electron beam is then used to cool the hadron beam.

5. Space-charge field with multiple bunches: plasma cascade amplifier

While the previous LSCA did not consider transverse behavior much, (Litvinenko *et al.*, 2018) introduced a new method called the plasma-cascade amplifier (PCA), which generates modulation at the plasma frequency using solenoid focusing. As previously mentioned, $G > 1$ occurs when the space-charge induced momentum is strong enough, which is converted to the density by an appropriate R_{56}. It is hard to expect a high gain in a drift because the R_{56} of a drift (L_d/γ_0^2) is much lower than the one from a chicane ($2L_{dogleg}\theta^2$, $\theta \ll 1$). However, the final density modulation is governed by how many particles gather together at one spot, so $R_{56}\delta$ is the critical term. Thus, even though a drift provides a low R_{56}, providing a stronger momentum modulation from the space-charge force can make a similar amplification. One of the clear ways to achieve it is to increase

the density by focusing the beam, as we can infer from Eq. (199).

The plasma cascaded instability can also be understood as a parametric resonance of harmonic oscillators as described by (Litvinenko *et al.*, 2018). In the harmonic oscillator, the oscillation can increase exponentially when the oscillation frequency is modulated at a certain frequency (i.e., $\omega_0(t) = \omega_0 \left[1 + A \cos(\omega t)\right]$). This resonance happens when $\omega \simeq 2\omega_0$. In the case of the PCA, ω_0 corresponds to the plasma oscillation frequency. Thus, if the plasma oscillation frequency can be modulated, then a parametric resonance will occur in the beam's longitudinal density modulation. Here, a series of focusing solenoids are used to modulate the plasma oscillation frequency, as shown in Fig. 34. Due to the required relationship between the plasma oscillation frequency and its modulation frequency, it is clear that system parameters such as the beam envelope, solenoid-to-solenoid distance, and charge level should be carefully designed. Thorough theoretical work has been done by (Litvinenko *et al.*, 2018), and further work is on-going [see (Blaskiewicz, 2019)].

The method was experimentally demonstrated by (Litvinenko *et al.*, 2019; Petrushina *et al.*, 2019). A long bunch (400 ps) with low energy (1.76 MeV) was used for the experiment, and the charge was varied. They used 5 cells to amplify the shot noise, and the solenoids were not placed periodically, which works equally well as a periodic setup. From the simulation, they expected a gain of 400-500 for 0.4 THz, and a gain around 200 for 0.6 THz. Figure 34 shows measured longitudinal beam profiles. The density modulation is very clear, and it shows a charge dependency, which is expected because $G > 1$ occurs with a strong enough momentum modulation from the space-charge force. The spectrum of these modulated bunches is broadband with a peak around 0.4 THz as (Litvinenko *et al.*, 2019; Petrushina *et al.*, 2019) expected from their simulation. See (Litvinenko *et al.*, 2021) for further details.

As LSCA had applications to radiation and coherent electron cooling, the PCA can be applied for the same purposes. While the LSCA requires a dispersive beamline such as a chicane, the PCA does not introduce any dispersion to the beam. This may provide a relative benefit in terms of beam quality preservation due to the lack of CSR and timing issues for the electron cooling.

B. Shaping profiles using coherent synchrotron radiation

CSR is another beam-generated field that changes a longitudinal beam momentum distribution along time. Similar to the space-charge field, a CSR field is determined by the bunch's longitudinal profile, and the field strength is the only controllable knob, as we can see from Eq. (142). CSR cannot be generated without a disper-
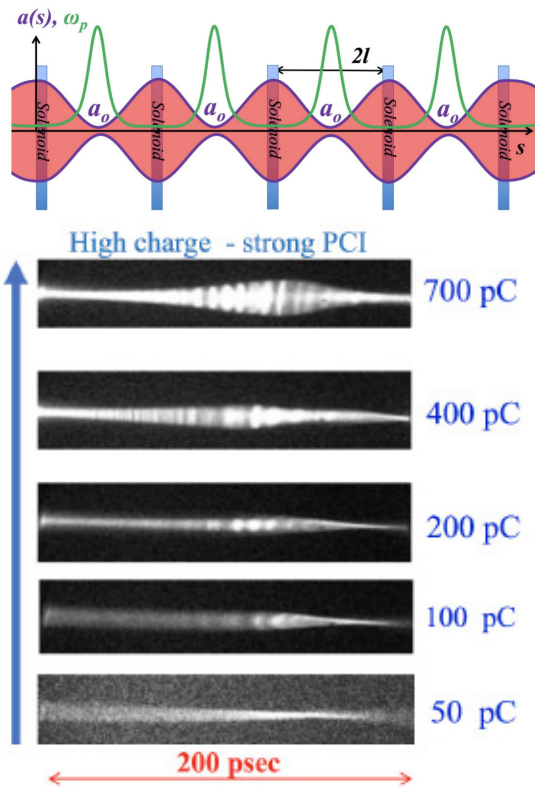
Figure 34 Experimental demonstration of plasma cascade amplification. The conceptual figure on top shows the configuration of a PCA with corresponding plasma wavelength modulation (top), and the measured longitudinal density distribution (bottom). Adapted from (Petrushina *et al.*, 2019).

sive element such as a dipole magnet. This means that the transverse and longitudinal phase space will be correlated along the path, and it makes the analysis and control more difficult than the space-charge field of a single bunch. Thus, using CSR directly for shaping is challenging, and no direct use of it has been proposed so far.

However, CSR is indeed a field that can change the longitudinal profile. It is normally considered an obstacle for longitudinal beam manipulation. There have been several research efforts that tried to preserve the shape through a beamline with CSR. In the rest of this section, we describe efforts to suppress CSR's impact on the profile. Studying these methods will tell us how to use CSR as a shaping tool.

(Mitchell *et al.*, 2013) showed theoretically that a certain beam shape can flatten the CSR wake, and it can minimize the CSR's impact on transverse emittance growth through the chicane. The beam profile is given by

$$\lambda(z) = \frac{4}{3} \frac{(z-a)^{1/3}}{(b-a)^{4/3}}, \tag{201}$$

where a and b are the limits of the longitudinal bunch
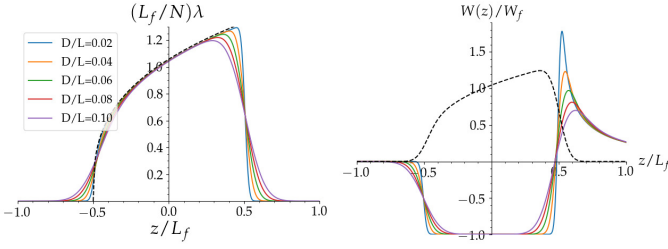
Figure 35 Theoretical estimation of aberrations of ideal current profile flattening CSR wake and aberration's impact on the CSR wake. Current profile with aberrations are displayed in the left panel. The right panel shows corresponding CSR wakes. The dashed curve in the left panel shows the ideal profile from Eq. (201). Solid curves show profiles aberrated by the chicane. Each curve corresponds to a different $\mathcal{D}/L$ where L is the initial bunch length. L_f ($= L/C$) is the final bunch length after compression. Adapted from (Mitchell et al., 2013).

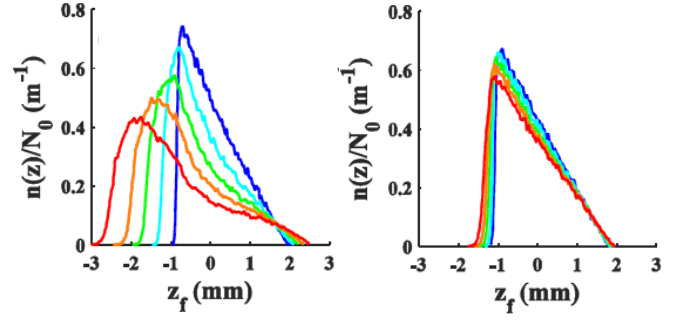Figure 36 Correction of CSR-induced aberration in start-to-end simulation. The x-axis is the beam's longitudinal position after a EEX beamline, and the y-axis is the normalized longitudinal density. The colors represent charges of 1.1, 3.3, 5.5, 7.7, and 11 nC. The left panel was simulated with a $20°$ bend for the dogleg, and the right panel used a $12°$ bend. Adapted from (Ha et al., 2016).

domain ($a \leq z \leq b$). This shape is given on the left of Fig. 35, and the corresponding CSR field is on the right.

Here, preserving the initial shape is critical to minimize CSR's impact on the emittance, so (Mitchell et al., 2013) derived an equation for the longitudinal profile starting from a decoupled 6D distribution. The profile at distance s can be written as

$$\lambda(z; s) = \frac{C}{\sqrt{2\pi}\mathcal{D}} \int_{-\infty}^{\infty} \lambda(\zeta; 0) \exp\left(-\frac{(\zeta - zC)^2}{2\mathcal{D}^2}\right) d\zeta, \tag{202}$$

where C is the compression factor, $\mathcal{D}$ is given by

$$\mathcal{D}^2(s) = C^2 \left(R_{56}^2(s)\sigma_\delta^2 + \mathcal{H}(s)\right), \tag{203}$$

and $\mathcal{H}$ is given by

$$\mathcal{H}(s) = \frac{\left[\sigma_x^2 R_{51}(s) - \sigma_{xx'} R_{52}(s)\right]^2 + \varepsilon_x^2 R_{52}^2(s)}{\sigma_x^2}. \tag{204}$$

The change of the profile depends on C and $\mathcal{D}$. When $\mathcal{D}$ is close to zero, $\lambda(z; s) = C\lambda(zC; 0)$, which means the final profile has the same shape as the initial profile but is compressed. Thus, the profile would not change significantly when $\mathcal{D}$ is small. Figure 35, which shows the longitudinal profile for different values of $\mathcal{D}$ and their corresponding CSR wakes, proves that CSR has a small impact on the profile when $\mathcal{D}$ is small. This tells us that a small R_{56}, initially low energy spread, and a small transverse emittance are preferred. Also, beam matching into the beamline is another important factor.

(Ha et al., 2016) tracked CSR effects on the beam in an EEX beamline to find a way to preserve the shaping quality of EEX-based shaping (see Sec. II.C.3 and Sec. VI.C.1). They found that R_{56} of the second dogleg is the carrier of CSR-induced momentum change. Thus, reducing R_{56} was simulated to confirm its impact. Figure 36 shows current profiles at the exit of an EEX beamline

for several different charge levels. The left figure corresponds to an EEX beamline with a $20°$ bend, and the right figure corresponds to a bending angle of $12°$. The R_{56} of the doglegs were 0.29 m and 0.18 m, respectively. The profile of 11 nC case has several clear differences from the 1.1 nC case when R_{56} is high. For example, the tail is lengthened, the peak is rounded, the linear ramp is changed to a concave curve, and the bunch length is significantly lengthened. On the other hand, the low R_{56} case preserved all features (tail, peak, ramp, and bunch length) even for 11 nC.

(Tan et al., 2021) performed an optimization by using a reverse tracking simulation to find a beamline setting and an initial beam condition for providing the desired final profile. By optimizing the entire beamline, nonlinearities introduced by each part of the beamline could cancel each other out. The beamline they used in the simulation consisted of an SRF gun, 650-MHz linac cavities, a 3.9-GHz linac cavity, and two chicane compressors. Optimization variables included accelerating gradient, phase, frequencies of 650-MHz linac cavities, R_{56} and T_{566} of each chicane, and parameters defining the initial profile.

Even though the charge of the beam they used was 10 nC, they found that a reasonable beamline setting provided the desired final longitudinal profile. Even this setting included quite a high R_{56} for both chicanes, 0.129 m and 0.131 m, which can introduce a strong CSR effect on the profile. However, the work was done with a simulation code with several simplified physics (e.g., beam propagation using R_{56} and T_{566}, steady-state CSR only, etc.). Thus, more work is required to reach a definite conclusion. However, it is promising that control of nonlinearities of the beamline may compensate a strong CSR's impact on the final profile.

From the study of this research on CSR suppression, we can infer how to make CSR's impact on the final profile stronger. Strong CSR, due to high charge or short bunch,

coupled with high R_{56} may provide the CSR dominant control over the final profile. This would induce a profile change along the beamline; thus the final profile could be made to have completely different features than the initial one. However, the final profile that the CSR can generate would be limited because the CSR field is determined by the beam's profile along the beamline. This means that we need to either shape the initial profile or create other knobs to control the beam's longitudinal profile inside the beamline. R_{51}, R_{52}, and transverse phase space may be good candidates for the additional knob, and we may be able to use nonlinearities from other parts of the beamline to manipulate the final profile further.

C. Shaping profiles using wakefields

The longitudinal wakefield inside a given structure can be expressed as the convolution of current distribution and wake function, which is the wakefield from a single particle; see Eq. (122). Here, the longitudinal wake function can be written as the summation of each mode excited in the structure:

$$w_z(z) = \sum_i 2\mathcal{K}_i cos(k_i z), \qquad (205)$$

where k_i is the wavenumber of the ith mode, and $\mathcal{K}_i$ is the loss factor of ith mode which is defined as the energy lost by the particle exciting the mode per unit charge squared. While the space-charge field and CSR did not provide an additional knob other than its strength, in the case of the wakefield, the material or geometry of the vacuum pipe determines the wake function (i.e., loss factor and fundamental frequency). Thus, manipulation using the wakefield provides higher degrees of freedom than other beam-generated fields.

In this section, we discuss the use of the wakefield in shaping the longitudinal density profile. Similar to methods described in Section IV, wakefields control time-energy correlation and the subsequent beamline (e.g. drift and anisochronous beamline) make the correlation that determines the longitudinal density distribution. Thus, we can easily imagine applications such as bunch compression, bunch train generation, and single-bunch shaping. These applications will follow the same principle as the ones in Section IV, but the wakefield-based techniques could be more compact and cost effective than external devices.

In addition to profile shaping, the degrees of freedom that wakefields provide enable further control over the longitudinal phase space. Therefore, we also discuss manipulations of the energy distribution and longitudinal phase space using wakefields.
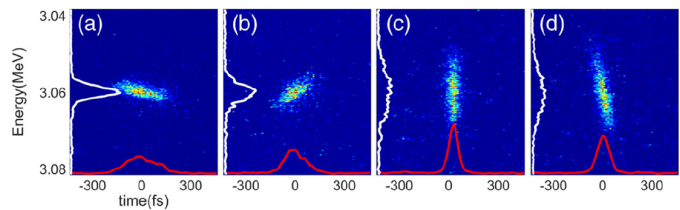

Figure 37 Experimental demonstration of wakefield-based bunch compression using THz structure with drive bunch. The figures show measured longitudinal phase space of the target bunch and corresponding projections. Each figures corresponds to different drive bunch charges. The charges were (a) 0, (b) 0.6, (c) 0.9, and (d) 1.3 pC. Adapted from (Zhao et al., 2018a).

1. Bunch compression

A wakefield can work with two compression mechanisms; ballistic bunching and magnetic bunching. Both mechanisms require control over the longitudinal chirp, which can be provided by the wakefield. The chirp can be produced in two different ways. If there is a wakefield-driving bunch, the target bunch can be placed at the zero-crossing phase of the wakefield, to provide either a negative or a positive chirp. If there is no driving bunch, then the self-wakefield inside the target bunch always makes the tail lose more energy than the head of the bunch (with an appropriate choice of frequency). Thus, the target bunch always has a positive chirp. The positive chirp can work with magnetic bunching, but it is not appropriate for ballistic bunching because a drift has $R_{56} > 0$.

Ballistic bunching using a wakefield was experimentally demonstrated by (Zhao et al., 2018a). They split a Ti:sapphire laser at 800 nm into three pulses. Two pulses were used to generate the drive and target bunch while the third pulse was used for diagnostics purposes by converting it to THz radiation through optical rectification. A 5-cm-long quartz tube with an inner diameter of 400 μm was used to generate a wakefield, and a drive bunch charge ranging from 0.6 to 1.3 pC was used to generate different longitudinal chirps. Thus, for a given $R_{56} \approx 3.07$ cm, they generated different compression ratios. The results are displayed in Fig. 37. An uncompressed bunch length of 150 fs rms was compressed to 2.8 fs rms in the case of (c).

While a similar compression can be easily accomplished with conventional RF cavities, the wakefield-based ballistic bunching offers a few advantages. First, this compression requires a driving bunch that can easily be generated by splitting UV laser pulses. Therefore, it does not require extra cavities for chirping and corresponding RF power sources. Second, RF sources always introduce jitter on the beam energy. Because the wakefield fully relies on the given structure and the beam, it can provide better energy stability. (Zhao et al., 2018a)

measured energy stability of 2.4×10^{-4} rms for wakefield-based bunching, while it was 1.5×10^{-3} rms for bunching with a buncher cavity. Third, this bunching also reduces arrival time jitter. In (Zhao *et al.*, 2018a) they measured arrival timing jitter using a THz streaking method proposed by (Zhao *et al.*, 2018b). The measured arrival time jitter for wakefield-based bunching was about 60 fs rms, while the buncher provided a jitter of 170 fs rms.

Similarly, magnetic bunching with a wakefield can also have the benefit of reduced timing jitter. Small timing jitter is highly demanded in many modern accelerators. Especially, it is important for wakefield accelerators. Wakefield accelerators usually operate at very high frequency, so placing a short target bunch at the right time is critical because it determines the beam's energy spread and energy gain. However, controlling the separated bunch's timing in a few ps or lower is not straightforward.

(Zhao *et al.*, 2018a) proposed a scheme to use the wakefield to reduce the timing jitter between an externally injected drive bunch and main bunch. The proposed concept uses the wakefield from a drive bunch to produce a negative chirp on the main bunch's longitudinal phase space. This chirping process is followed by a chicane to compress the main bunch. If the main bunch is a little late compared to the ideal time, it will be accelerated by the driving bunch's wakefield. In the following chicane, the main bunch will take a shorter path than the path it would take when it arrives at the ideal time (i.e., zero-crossing of the wakefield). This path length difference compensates the original timing error after the compression. Similarly, if the main bunch arrives earlier than the ideal time, the main bunch will be decelerated by the wakefield and this bunch will take a longer path than the bunch at the right timing. Thus, the timing errors can be compensated.

2. Bunch train generation

The bunch compression discussed in the previous section requires a condition $\frac{\sigma_z}{\lambda_{wake}} \ll 1$. If we go to the other regime ($\frac{\sigma_z}{\lambda_{wake}} \gg 1$), the wakefield can imprint a sinusoidal energy modulation on the beam, which can be used for bunch train generation. Once again, we can use either a drift or other anisochronous beamline as a momentum-to-density modulation converter. A drift requires a low-energy beam to provide sufficient R_{56}, and the overall beamline setup can be simple and compact. On the other hand, an anisochronous beamline, such as a chicane, can generate a high R_{56} that is compatible with any beam energy. This was experimentally demonstrated by (Lemery *et al.*, 2019) using a low-energy beam with drift and by (Antipov *et al.*, 2013) using a compact chicane.

In (Lemery *et al.*, 2019), they used a beam with charge of 1 nC and an energy of 6.2 MeV. Two structures were prepared (Inner diameter: 450, 750 μm, 5 and 8 cm long), followed by an accelerating cavity providing 14 MeV of energy gain. For both structures, downstream diagnostics measured the phase space, and the modulation was successfully imprinted in the LPS. The measured wavelength of the final modulation was 1.01, and 1.81 mm for each structure. They showed good agreement with the expected fundamental wavelengths of each structure.

On the other hand, (Antipov *et al.*, 2013) used a chicane consisting of four permanent magnets to provide a high R_{56} ($\approx$ 4.9 cm). Here the beam's charge and energy were 0.5 nC and 57 MeV, respectively. A 5-cm-long Kapton capillary with an inner diameter of 300 μm was used. The corresponding fundamental frequency was around 0.8 THz. During the experiment, they controlled the incident longitudinal chirp by controlling the phase of the linac before the compressor. This macro-chirp on top of sinusoidal modulation, can provide either a frequency upshift or downshift. The modulation frequency was measured from the autocorrelation of the coherent transition radiation, and the result is shown in Fig. 38. When there was no chicane, the modulation conversion to the density modulation was not observed. When a chicane was introduced, a density modulation appeared, and its frequency was close to the expected 0.8 THz (see the 189 keV/mm case). Also, depending on the linac phase, the chirp was changed from -567 keV/mm to +567 keV/mm. Although the initial modulation frequency was about 0.8 THz, the final density modulation frequency could be tuned from 0.68 to 0.9 THz.

Equation (154) is applicable to both (Lemery *et al.*, 2019) and (Antipov *et al.*, 2013) methods. In the case of the wakefield, the modulation amplitude induced by the initial energy modulation and R_{56} was $A_{ind} = Ck|R_{56}|\Delta\delta \exp\left(-\frac{1}{2}C^2k^2R_{56}^2\sigma_\delta^2\right)$, where $\Delta\delta$ is the amplitude of the initial energy modulation [see (Saldin *et al.*, 2002a)]. Note that the required $R_{56}\Delta\delta$ for the same final modulation amplitude decreases as modulation frequency increases. $R_{56}\Delta\delta$ is the particle's travel distance. To build a density spike from the sinusoidal energy modulation, particles need to travel a little less than $\lambda/4$. Thus, high-frequency modulation naturally requires either a small R_{56} or a small initial modulation amplitude. Although the beam energy was 57 MeV for (Antipov *et al.*, 2013), the total length of the chicane was less than 50 cm, and R_{56} was only around 0.049 m. Their high modulation frequency (0.8 THz) enabled the use of a compact chicane. Also, as the modulation frequency increases, the initial uncorrelated energy spread becomes more important. The exponent of A_{ind} includes a $k\sigma_\delta$ term. As the frequency increases, the modulation amplitude decays exponentially. Thus, a higher frequency requires a lower initial spread to keep the exponential decay at the same level.

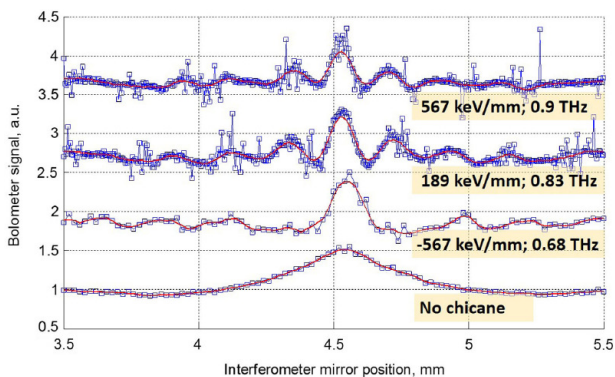Finally, we note that modulation does not occur if the

Figure 38 Experimental demonstration of bunch train generation using a dielectric structure followed by a chicane. The figure shows measured CTR signals, which are autocorrelation results. Each curve corresponds to a different initial energy chirp as described in the label next to the curve. The squares are the data points, and the curves are the smoothed fits. From (Antipov *et al.*, 2013).

bunch's longitudinal profile's frequency spectrum does not overlap with the structure's frequency spectrum. Thus, imprinting multi-period modulation is not possible for a Gaussian profile. Both experiments described above used flat-top profiles.

3. Single-bunch profile shaping

The current profile of a single bunch can be shaped by introducing an appropriate nonlinearity or by masking. As we saw from transverse manipulation by (Yuri *et al.*, 2007), introducing an appropriate nonlinearity on the phase space can change the profile. Figure 39 shows an example for a triangular profile. A wakefield introduces a nonlinear curvature in the linear phase space. In this example, particles in the tail are accelerated due to the wakefield, while particles near the center lose their energy. Thus, when the beam enters a chicane, particles in the tail overtake the leading particles and build up a density spike. At the same time, due to the nonlinearity in the energy gain along the longitudinal position, the tail becomes sharp. The head of the current profile smears out for the same reason. Note that this happens when the wavelength of the wakefield is comparable to the bunch length so that the appropriate nonlinearity is introduced to the phase space.

In (Andonian *et al.*, 2017) this method was experimentally demonstrated using an 80-pC, 50-MeV beam and a compact chicane with $R_{56} = 9.2$ mm. They used a dielectric wakefield structure having an inner diameter of 200 μm and a length of 5 cm. The fundamental mode frequency was 0.39 THz, which is about a quarter of $1/\sigma_z$. The beam's profile was measured with and without the structure to compare the effect of the wakefield. The reconstructed profile from CTR interferometry is given in
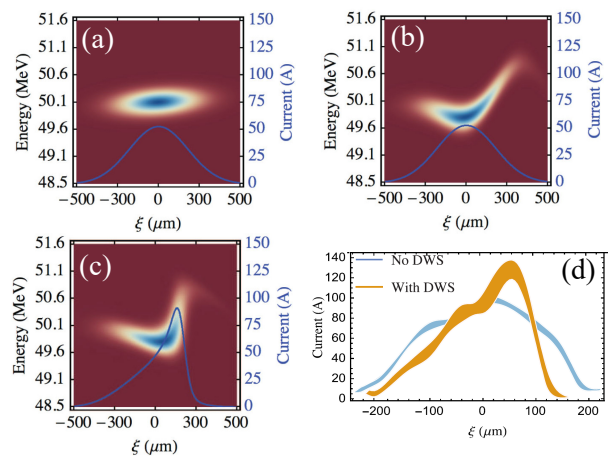

Figure 39 Experimental demonstration of triangular current profile generation using a dielectric structure followed by a chicane.(a-c) are simulation results, and (d) is reconstructed profile from interferometer measurements with and without the structure. (a-c) corresponds to longitudinal phase space before the structure, after the structure, and after the chicane, respectively. The initial phase space (a) is generated artificially using beam parameters used for the experiment. The modulation applied to this phase space is calculated from a formula using structure parameters used for the experiment. Adapted from (Andonian *et al.*, 2017).

Fig. 39(d).

The method can be extended to generate arbitrary current profiles when we use a series of structures with appropriate frequencies. Similar to (Piot *et al.*, 2012) and (Ha *et al.*, 2019a), each structure represents a Fourier component of the target correlation. With an appropriate R_{56}, the correlation can change the initial profile to the desired profile. A scheme using transverse wigglers was proposed by (Ha *et al.*, 2019a) and can be used for the longitudinal phase space directly by simply replacing transverse wigglers with wakefield structures. Simulations in (Mayet *et al.*, 2020) showed the feasibility of using several structures with different frequencies. They also showed the feasibility of structure fabrication.

A masking-based technique has not been proposed yet. However, combining wakefield-based shaping with methods introduced this paper could provide advantages. For example, one can adopt a wakefield-based deflector (see (Novokhatski, 2015)) instead of RF deflecting cavities for the method in (Ha *et al.*, 2020); see Sec. IV.B.1. The wakefield deflector uses a transverse wakefield that kicks the trailing particle transversely. Thus, it generates time and transverse correlation just like RF deflecting cavities. While this passive device may provide a big advantage on timing jitter, the nonlinearity of the transverse wake will require more extensive analysis to design the mask and eliminate the correlation after chopping (Bettoni *et al.*, 2016; Craievich and Lutman, 2017; Seok *et al.*, 2018).
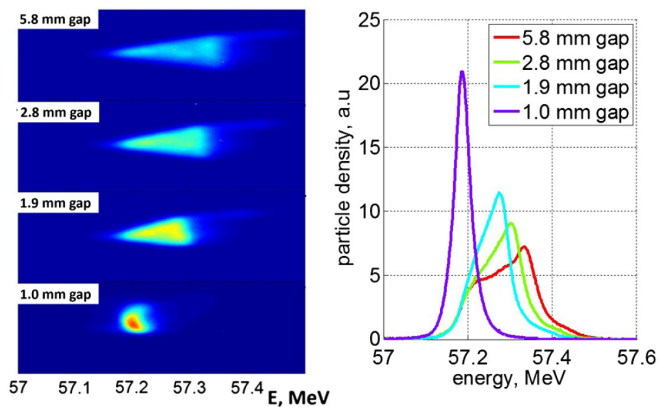
Figure 40 Experimental demonstration of energy spread reduction by wakefield from dielectric structure. The left panel shows beam images on the spectrometer with different structure gap sizes. Corresponding profiles are shown in the right panel. From (Antipov *et al.*, 2014).

4. Control over the energy distribution

Both (Antipov *et al.*, 2014) and (Emma *et al.*, 2014b) experimentally demonstrated the effectiveness of the wakefield dechirper. Both experiments used a slab structure, which consists of two jaws. Thus, the wakefield strength was adjusted by the gap size. When the operating conditions, such as a beam's profile, are known, the structure can be optimized for the operating condition. However, these proof-of-principle experiments used the slab structure with an adjustable gap to provide flexibility to the experiment. Figure 40 shows measured beam images after the spectrometer and corresponding spectrum from (Antipov *et al.*, 2014). As the gap size decreased, a quasi-linear wakefield amplitude increased, and it successfully reduced the energy spread.

The low correlated energy spread from a dechirper can benefit all modern accelerator applications. For example, the X-ray free-electron oscillator (XFELO) requires energy spread of $< 1 \times 10^{-4}$ (Kim *et al.*, 2008), which is challenging to achieve with existing methods. However, a recent study by (Qin *et al.*, 2016) showed its feasibility. Optimization of most of the linac parameters (phase, gradient, R_{56}, etc.) was performed, and the beamline had a dechirper at the end. This optimization in the simulation generated the core of the beam ($\sim$400 fs) having a relative energy spread of $\sim 2.6 \times 10^{-5}$ (105 keV rms).

Although the methods just introduced demonstrate the compensation of the chirp by experiments or simulations, there was a common limitation. The methods introduce a nonlinearity in phase space that increases the correlated energy spread. The linearity in the core of the phase space can be preserved by designing a structure having a low-frequency single-mode, whose wavelength is much longer than the bunch length. However, preserving the linearity in the periphery is challenging. When

$k\sigma_z \ll 1$, Eq. (122) and Eq. (205) can be simplified to

$$W_z(z) \approx 2\mathcal{K}Q_b \int_{-\infty}^{z} \lambda(\zeta)d\zeta, \qquad (206)$$

where Q_b is the total charge, and $\lambda(z)$ is the normalized density profile. Thus, only a uniform profile provides $W_z(z) \sim z$, which preserves both the core and the periphery's linearity. All other profiles would would have a nonlinear periphery in their phase space. .

A simulation in (Antipov *et al.*, 2014) showed the feasibility of extending linear region of the wakefield by using a multi-mode structure. However, this approach does not provide any tunability after the fabrication of a structure. (Antipov *et al.*, 2014) also remarked about the use of two or more dechirpers to control nonlinearity issues. This multi-structure approach was recently simulated by (Mayet *et al.*, 2020). It provided high-quality dechirping, but the dechirping for the periphery was still limited due to the limitation of Fourier synthesis known as the Gibbs phenomenon.

5. Control over longitudinal phase space

As we notice with the previous single-bunch shaping and energy distribution control, a wakefield applied to the beam changes the longitudinal phase space, and it can be used to control the correlation of the phase space. The most demanding control over the correlation is linearization, and it is experimentally demonstrated by (Deng *et al.*, 2014) and (Fu *et al.*, 2015). For a simple description, if we assume that only the RF curvature from the linac generates nonlinearities and the beam's energy gain from the linac can be written as $E_{RF}\cos(k_{RF}z)$, this gain can be approximated to $E_{RF}\left[1 - \frac{(k_{RF}z)^2}{2}\right]$ for $k_{RF}z \ll 1$. Similarly, the wakefield from a structure having only a single mode and a uniform profile can be written as $\frac{n_0 W_0}{k_w}\sin\left(k_w(z + \Delta z/2)\right)$, where Δz is the width of the uniform profile, and n_0 is the density. Because $k_w t$ should be small while the wavelength of the wakefield should be shorter than the wavelength of the RF field in the linac, we can approximate it as $n_0 W_0 \left[-\frac{\Delta z k_w^2}{4}z^2 + \left(1 - \frac{k_w^2 \Delta z^2}{8}\right)z + \frac{\Delta z}{2}\right]$. Thus, it is possible to eliminate the quadratic term that generates the curvature in the phase space by choosing the proper strength of the wakefield. This can be simply done by adjusting the gap size in the case of slab structures.

An experiment by (Deng *et al.*, 2014) demonstrated the linearization, and they generated undulator radiation to estimate the impact of the linearization on the radiation quality. They expected to see a significant reduction of the radiation bandwidth due to the linearization and a shift of the central wavelength. As the wakefield strength increased, the bandwidth became narrower. The bandwidth from the 6-mm gap was measured to 7.8 nm. On
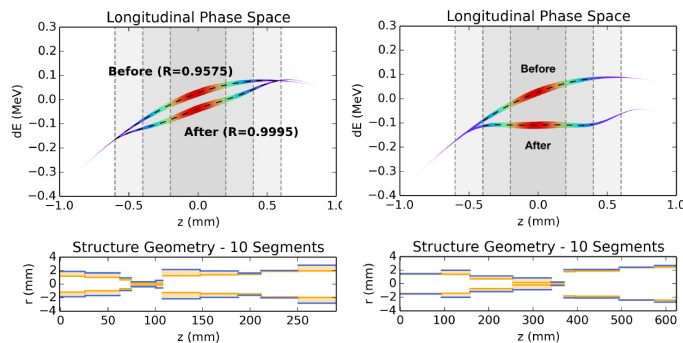
Figure 41 Longitudinal correlation control via a series of dielectric structures. The top panels show simulated longitudinal phase spaces before and after the correction using 10 dielectric structures, and the bottom panels show the corresponding structure geometries for simulations. The left column shows an example of linearization while the right column shows dechirping. Adapted from (Mayet *et al.*, 2020).

the other hand, the 2-mm gap provided 3.7-nm bandwidth.

This experiment was designed to observe the effect of linearization on the radiation. However, we can imagine that the bunch compression will benefit from this linearization too. Currently, most of the linearization is done with harmonic cavities, which require additional RF power sources. However, the beam-generated field may be able to manage the nonlinearity with lower cost. (Penco *et al.*, 2017) used wakefields from a high-impedance linac and a dielectric waveguide structure to replace high-harmonic cavities. The result was similar to the linearization result by high-harmonic cavities.

A series of structures with different frequencies may enable further control of the phase space such as correlation control of the phase space for an arbitrary current profile or control of longitudinal chirp with the linearization. The wakefield from each structure will work as a Fourier component and can generate arbitrary correlation in the phase space; see (Ha *et al.*, 2019a). This concept was demonstrated in a simulation (Mayet *et al.*, 2020) using a series of dielectric structures. Because the choice of parameter set becomes too complex, they used an optimization algorithm to find the best result. Figure 41 shows two of their simulation results. The top panel shows the longitudinal phase space with and without linearization by wakefields, and the bottom panel shows the corresponding structure geometries. They used a total of 10 structures to correct the shape and control the chirp. Although this approach requires experimental demonstration and further study on beam transport, instability, fabrication error, etc., this is indeed a powerful method that would enable us to optimize the longitudinal phase space for each application.

VI. COUPLING BETWEEN DEGREES OF FREEDOM FOR PHASE-SPACE TAILORING

A. Introduction

This section discusses shaping methods that rely on the coupling between phase spaces associated with different degrees of freedom (DOF). In Section IV, the introduction of *local* coupling was shown to provide some control over a coordinate not usually accessible in an uncoupled beamline. In that case, shaping is performed within "coupling bumps" and the correlation is removed downstream. Ultimately, the fact that the coupling is partial limits the precision of that shaping method. This section focuses on beamlines that have strong coupling or can swap phase-space coordinates. This class of phase-space manipulation enables precise shaping of the beam phase space and opens a path to emittance repartition among the different DOFs.

B. Coupling between the two transverse degrees of freedom

1. Producing beams with canonical angular momentum

Canonical angular momentum (CAM)-dominated, or "magnetized", beams have important applications in electron cooling of heavy-ion beams (Budker *et al.*, 1975; Derbenev and Skrinsky, 1978; Parkhomchun and Skrinskii, 2000). In such a scheme, a cold electron beam co-propagates with the ion beam at the same average velocity. Collisions between ions and electrons transfer thermal motion away from the ion to the electron beam. The cooling efficiency can be greatly improved by using a magnetized beam. More recently, the use of a CAM-dominated beam was also considered for mitigating a resonance-driven instability in long periodic focusing channels (Cheon *et al.*, 2020). A simple technique for forming CAM-dominated beam consists of immersing a cathode in an axial magnetic field B_z, i.e., inside a solenoid magnet. In such a case, the canonical momentum $L = \frac{eB_z}{2}r^2$ is proportional to B_z. In a rotationally invariant system, the conservation of angular momentum implies that the beam acquires a mechanical angular momentum (MAM) equal to the CAM once it exits the solenoidal field as B_z vanishes; henceforth the beam's motion in the two transverse planes becomes coupled.

As discussed in Section- II, beam dynamics is determined by several factors including space charge, thermal emittance, angular momentum, and external fields. However, beam dynamics differs drastically when one factor dominates over the others; see Fig. 42. In Fig. 42(a) the dynamics is dominated by the thermal emittance, and the electrons have random momentum direction; in (b) the beam is dominated by its angular momentum, and the electrons shear in a vortex pattern; and in (c) the beam

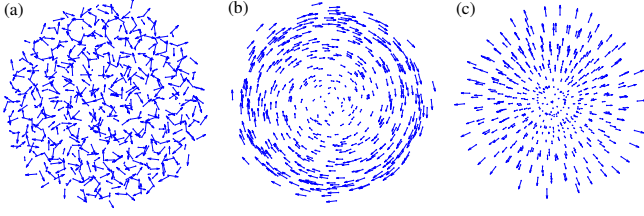is dominated by space charge, and the electrons repel each other by the Coulomb force and move outwards.

Figure 42 Motions of the electrons when the beam is dominated by (a) emittance, (b) angular momentum or (c) space charge. Each dot represents an electron in (x, y) space, and the arrows show the magnitude and direction of the electrons' velocities. From (Sun, 2005).

Taking into account all if these contributions, the rms transverse envelope equation for an electron bunch propagating in a drift space is described by (Reiser, 1994) as

$$\sigma'' - \frac{K_p}{4\sigma} - \frac{\varepsilon_u^2}{\sigma^3} - \frac{\mathcal{L}^2}{\sigma^3} = 0, \tag{207}$$

where σ is the transverse rms size, $K_p = \frac{2I}{I_0 \beta^3 \gamma^3}$ is the generalized perveance, I is the absolute value of the peak beam current, $I_0 = 4\pi\epsilon_0 mc^3/e \approx 17\,\text{kA}$ is the Alfvén current, ε_u is the uncorrelated transverse rms emittance, and $\mathcal{L}$ is related to the average canonical angular momentum $\langle L \rangle$ (see Section II.B.4) and the longitudinal momentum p_s of the beam via the magnetization

$$\mathcal{L} = \frac{\langle L \rangle}{2p_s}. \tag{208}$$

The second, third, and fourth terms of Eq. (207) represent the effects due to space charge, emittance, and angular momentum, respectively. When the fourth term is much greater than the second and the third terms, the beam is *angular-momentum-dominated*.

If there is external electromagnetic linear focusing, an extra term in the form of $k_0\sigma$ can be added to the envelope equation, where k_0 is related to the strength of the external focusing force.

The magnetic field on the photocathode is normally zeroed to minimize the projected emittances. This can be seen from Eq. (207), where the canonical angular momentum term $\mathcal{L}$ plays the same role as the emittance term in the beam-envelope equation so that it can be introduced to an effective emittance $\varepsilon_{eff} = \sqrt{\varepsilon_u^2 + \mathcal{L}^2}$, as noted in Section II. However, a large magnetic field is required to produce an angular-momentum-dominated beam in order that the correlation between the two transverse degrees of freedom dominates, i.e., $\mathcal{L} \gg \varepsilon_u$.

Recall the conservation of canonical angular momentum, Eq. (79). At the photocathode, the average of the first term is zero since $\langle \dot{\phi} \rangle = 0$. The second term must

not vanish in order to allow the beam to acquire angular momentum. Therefore, an axial magnetic field on the cathode is required to generate an angular-momentum-dominated electron beam. The first photoinjector-based generation of a CAM-dominated beam was demonstrated at Fermilab's A0 photoinjector facility (Sun *et al.*, 2004) using an L-band RF gun. The solenoidal lenses surrounding the RF gun were tuned to provide a variable magnetic field on the photocathode, and the beam's MAM was measured; see Fig. 43. In the latter work, the beam was further accelerated using a superconducing cavity, and the MAM was measured at ~ 15 MeV. Similar investigations were most recently conducted at Jefferson Laboratory on a 300-keV electron beam produced from a DC gun (Mamun *et al.*, 2018) and at the Fermilab FAST facility in a high-charge regime (3.2 nC) at 40 MeV (Fetterman *et al.*, 2022).
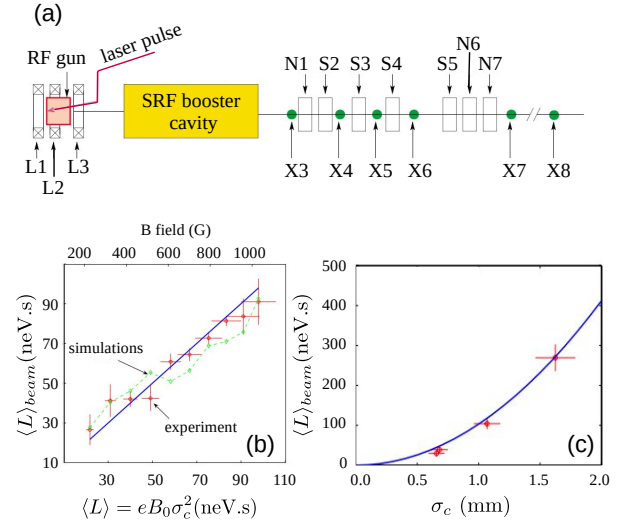
Figure 43 Experimental setup used to produce a CAM-dominated beam at the Fermilab A0 facility (a), example measured conversion between CAM and MAM (b), and demonstration of the quadratic dependence on laser spot size on the cathode (c). In diagram (a), the letters represent solenoidal magnetic lenses (L), skew quadrupoles (Q), and diagnostic stations (X, which means "cross"). Dimensions are in mm. In (b) and (c), the circles are measurements and the lines are fits. In plot (b), the dashed-curve shows the corresponding simulations. Adapted from (Sun *et al.*, 2004).

2. Decoupling of CAM-dominated beams and transverse emittance partitioning

A set of quadrupoles, with properly selected strength and separation, can apply a net torque to the CAM-dominated beam and remove its angular momentum. The result is an asymmetric beam with its two transverse degrees of freedom no longer coupled, i.e., a flat

beam. A flat electron beam, i.e., a beam with high transverse emittance ratio, can be produced from an angular-momentum-dominated beam (Brinkmann *et al.*, 2001). The technique consists of manipulating an angular-momentum-dominated beam produced by a photoinjector using the linear transformation described in (Derbenev and Skrinsky, 1978). A round-to-flat beam transformer, consisting of three skew quadrupoles and drift spaces, is discussed in (Burov and V.Danilov, 1998). This technique was proposed as a way of producing a high-aspect-ratio beam to mitigate beamsstrahlung in future linear colliders while also circumventing the use of an electron damping ring conventionally used to reduce the vertical emittance (Brinkmann *et al.*, 2001). Likewise, the technique was also adapted for applications in of microwave- and THz-radiation generation (Carlsten and Bishofberger, 2006; Kim and Kumar, 2007) using, e.g., the concept of Smith-Purcell backward oscillator (Andrews *et al.*, 2005).

Finally, flat beams were also proposed as an intermediary way of transporting and accelerating magnetized beams by transforming them from magnetized beams into flat beams and back again (Benson *et al.*, 2018; Piot and Sun, 2014).

The theory of generating a flat beam from an incoming angular-momentum-dominated beam is treated in several papers (Brinkmann *et al.*, 2001; Burov *et al.*, 2002; Derbenev, 1998; Kim, 2003). In this section, we follow the theoretical treatment based on the 4D beam matrix presented in (Kim, 2003), in which the round-to-flat beam transformation analysis was performed assuming that the beam and the transport channel upstream of the transformer are cylindrically symmetric and that the particle dynamics is Hamiltonian.

We will specify the coordinates of a particle in transverse trace space by two vectors:

$$\mathbf{X} = \begin{pmatrix} x \\ x' \end{pmatrix} \text{ and } \mathbf{Y} = \begin{pmatrix} y \\ y' \end{pmatrix}. \tag{209}$$

The corresponding 4×4 beam matrix is

$$\Sigma = \begin{pmatrix} \langle \mathbf{XX}^T \rangle & \langle \mathbf{XY}^T \rangle \\ \langle \mathbf{YX}^T \rangle & \langle \mathbf{YY}^T \rangle \end{pmatrix}. \tag{210}$$

Let $R(\theta)$ be the 4×4 rotation matrix of angle θ:

$$R(\theta) = \begin{pmatrix} I \cdot \cos\theta & I \cdot \sin\theta \\ -I \cdot \sin\theta & I \cdot \cos\theta \end{pmatrix}, \tag{211}$$

where I stands for the 2×2 identity matrix. The beam matrix is rotationally invariant if

$$\Sigma = R(\theta) \cdot \Sigma \cdot R(\theta)^{-1}. \tag{212}$$

From Eq. (212), we obtain

$$\langle \mathbf{XX}^T \rangle \cos^2\theta + \langle \mathbf{YY}^T \rangle \sin^2\theta$$
$$+ (\langle \mathbf{XY}^T \rangle + \langle \mathbf{YX}^T \rangle) \sin\theta \cos\theta = \langle \mathbf{XX}^T \rangle.$$

Since the rotation angle θ is arbitrary, Eq. (213) leads to

$$\langle \mathbf{XX}^T \rangle = \langle \mathbf{YY}^T \rangle, \tag{213}$$
$$\langle \mathbf{XY}^T \rangle = -\langle \mathbf{YX}^T \rangle. \tag{214}$$

Taking the transpose of both sides of Eq. (214)

$$\langle \mathbf{XY}^T \rangle^T = -\langle \mathbf{YX}^T \rangle^T = -\langle \mathbf{XY}^T \rangle, \tag{215}$$

we find that $\langle \mathbf{XY}^T \rangle$ is antisymmetric and can be written as

$$\langle \mathbf{XY}^T \rangle = \mathcal{L} J_{2D}, \tag{216}$$

where $\mathcal{L}$ is a constant related to the angular momentum L and longitudinal momentum p_s by

$$\mathcal{L} = \langle xy' \rangle = -\langle x'y \rangle = \frac{L}{2p_s}, \tag{217}$$

and J_{2D} is the 2×2 unit symplectic matrix given by Eq. (18).

By expressing the beam matrix in terms of *Courant-Snyder parameters* (also known as *Twiss parameters*) α, β (see, for example, (Wiedemann, 1999)), the general form of a round beam matrix in (x, x') or (y, y') subspaces can be written as

$$\langle \mathbf{XX}^T \rangle = \langle \mathbf{YY}^T \rangle = \varepsilon T_0,$$
$$\text{with } T_0 = \begin{pmatrix} \beta & -\alpha \\ -\alpha & \frac{1+\alpha^2}{\beta} \end{pmatrix}, \tag{218}$$

where ε is the rms transverse emittance, and $|T_0| = 1$.

Gathering Eq. (216) and Eq. (218), we may write the general form of a cylindrically symmetric 4×4 beam matrix in the following convenient form:

$$\Sigma_0 = \begin{pmatrix} \varepsilon T_0 & \mathcal{L} J_{2D} \\ -\mathcal{L} J_{2D} & \varepsilon T_0 \end{pmatrix}. \tag{219}$$

Let M be the transfer matrix of the transformer, which is symplectic. The beam matrix at the exit of the transformer is

$$\Sigma = M \Sigma_0 M^T. \tag{220}$$

Kim noticed two invariants associated with the symplectic transformation given by Eq. (220) (Kim, 2003):

$$I_1 = \varepsilon_{4D} = \sqrt{\det(\Sigma)}, \tag{221}$$
$$I_2(\Sigma) = -\frac{1}{2}\text{Tr}(J_{4D} \Sigma J_{4D} \Sigma). \tag{222}$$

where "Tr" is the trace operator, and J_{4D} is the 4×4 unit symplectic matrix .

Suppose a proper transfer matrix M exists such that the beam matrix at the exit of the transformer is block diagonalized

$$\Sigma = \begin{pmatrix} \varepsilon_- T_- & 0 \\ 0 & \varepsilon_+ T_+ \end{pmatrix}, \text{ with } T_\pm = \begin{pmatrix} \beta_\pm & -\alpha_\pm \\ -\alpha_\pm & \frac{1+\alpha_\pm^2}{\beta_\pm} \end{pmatrix}. \tag{223}$$

Applying Eq. (221) to the symplectic transformation given by Eq. (220), we have

$$\sqrt{\det(\Sigma)} = \sqrt{\det(\Sigma_0)} \Rightarrow \varepsilon_+ \varepsilon_- = \varepsilon^2 - \mathcal{L}^2. \qquad (224)$$

It is easy to calculate the second invariant once we verify that

$$J_{2D} T_0 J_{2D} T_0 = -I, \qquad (225)$$

which leads to

$$J_{4D} \Sigma_0 J_{4D} \Sigma_0 = \begin{pmatrix} -(\varepsilon^2 + \mathcal{L}^2)I & 0 \\ 0 & -(\varepsilon^2 + \mathcal{L}^2)I \end{pmatrix}, \quad \text{and}$$

$$J_{4D} \Sigma J_{4D} \Sigma = \begin{pmatrix} -\varepsilon_-{}^2 I & 0 \\ 0 & -\varepsilon_+{}^2 I \end{pmatrix}.$$

So, from Eq. (222) we have

$$I_2(\Sigma) = I_2(\Sigma_0) \Rightarrow \varepsilon_+{}^2 + \varepsilon_-{}^2 = 2(\varepsilon^2 + \mathcal{L}^2). \qquad (226)$$

Finally, the two transverse emittances can be derived from Eq. (224) and Eq. (226)

$$\varepsilon_\pm = \varepsilon \pm \mathcal{L}. \qquad (227)$$

Equation (227) gives the two transverse emittances of a completely decoupled asymmetric beam. One emittance (ε_+) can be orders of magnitude larger than the other (ε_-) given properly chosen initial conditions such as ε and $\mathcal{L}$, which are related to the beam matrix at the cathode surface:

$$\Sigma_c = \begin{pmatrix} \sigma_c^2 & 0 & 0 & \mathcal{L} \\ 0 & \sigma_c'^2 & -\mathcal{L} & 0 \\ 0 & -\mathcal{L} & \sigma_c^2 & 0 \\ \mathcal{L} & 0 & 0 & \sigma_c'^2 \end{pmatrix}, \qquad (228)$$

where σ_c and σ'_c are the initial beam size and divergence spread at the cathode, respectively. The intrinsic (or thermal) rms normalized emittance on the cathode (see Eq. (167)) is $\varepsilon_c^n = \sigma_c \sigma_{pc} = \beta \gamma \sigma_c \sigma'_c = \beta \gamma \varepsilon$.

The experimental generation of a flat beam from a CAM-dominated beam was first demonstrated at the Fermilab A0 photoinjector (Edwards *et al.*, 2000, 2001). The experimental setup was identical to the one displayed in Fig. 43 and three skew-quadrupole magnets located at ~ 15 MeV were employed to remove the angular momentum, as demonstrated via numerical simulation in Fig. 44(a-c). Ultimately, the experiment demonstrated the generation of a flat beam with a transverse emittance ratio of $\varepsilon_y/\varepsilon_x \simeq 100$ (Piot *et al.*, 2006); see Fig. 44. The experiment confirmed the underlying physics and was validated against numerical simulation; see Fig. 44(d,e). Most notably, the smaller measured normalized emittance was $\varepsilon_x^n \simeq 0.4$ μm. This number, although limited by the diagnostics resolution, was a factor ~ 2 smaller than the thermal emittance estimated from the laser spot size on the cathode which was ~ 1 μm.
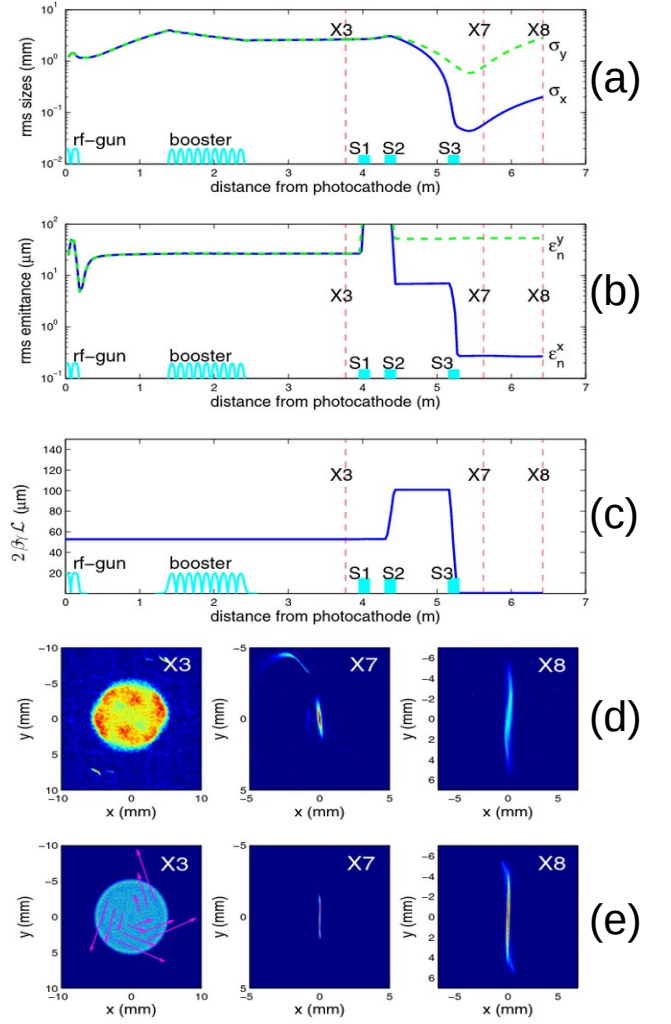
Figure 44 Simulations (a-c,e) and measurements of flat-beam generation for the A0 photoinjector. Plots (a), (b), and (c) are ASTRA simulation of respectively the beam size, emittance and magnetization evolution along the beamline showcasing the removal of angular momentum using three skew-quadrupole magnets S1, S2 and S3 to transform the incoming magnetized beam into a flat beam. The images displayed in (d) are measured beam distribution at X3, X7, and X8 screens while the distribution shown in (e) are the corresponding numerical simulations. The labels "X*i*" refer to the diagram appearing in Fig. 43(a). Adapted from (Piot *et al.*, 2006)

Further experiments, carried out at the AWA facility at Argonne, demonstrated the generation of a flat beam using a similar principle but with high charge up to 2 nC and an emittance ratio close to 200. The beamline also included a high-resolution phase-space measurement (using scanning slits) that permitted the reconstruction of the flat-beam phase space (Xu *et al.*, 2019).

3. Phase-space exchange between the two transverse planes

In this section, we consider the phase-space coordinate exchange in the **X**-**Y** 4D phase space. Such a phase-space swapping was first proposed for spectrometer applications in 1972 (Kowalski and Enge, 1972) and then for a Möbius accelerator (Talman, 1995). If introduced in an electron storage ring for a light source, the beam emittance in the x- and y-directions become equal and one half of the natural emittance in the normal ring, thus mitigating the lifetime limitation for intra-beam scattering (Aiba *et al.*, 2015). The XY exchange is also interesting for the next generation synchrotron radiation rings since it can lead to better horizontal injection efficiency (Kuske and Kramer, 2016). Likewise, such an exchange could be critical in lowering collective instabilities such as regenerative beam-break-up (BBU) instabilities in superconducting linacs (Rand and Smith, 1980; Tennant *et al.*, 2005) and in beam-driven wakefield accelerators (Gai *et al.*, 1997).

The transformation matrix of the form

$$R_{XY} = \begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix}; \quad \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \to R_{XY} \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \begin{pmatrix} A\mathbf{Y} \\ B\mathbf{X} \end{pmatrix} \tag{229}$$

can be constructed from an symmetric arrangement of five skew quadrupoles as follows:

$$R_{XY} = Q_{S1} L_1 Q_{S2} L_2 Q_{S3} L_2 Q_{S2} L_1 Q_{S1}. \tag{230}$$

Here Q_{Si} and L_i are the matrices for skew quadrupoles and drifts:

$$Q_{Si} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & q_i & 0 \\ 0 & 0 & 1 & 0 \\ q_i & 0 & 0 & 1 \end{pmatrix}, \quad L_i = \begin{pmatrix} 1 & d_i & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & d_i \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{231}$$

To simplify calculations, note that a skew quadrupole is obtained by a 45° rotation of a normal quadrupole:

$$Q_{Si} = R_{\pi/4}{}^{-1} Q_i R_{\pi/4}; \quad Q_i = \begin{pmatrix} 1 & 0 & 0 & 0 \\ q_i & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -q_i & 1 \end{pmatrix},$$

$$\text{and } R_{\pi/4} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix}. \tag{232}$$

Here we are using the 2×2 block matrix notation and since drift matrices are rotationally invariant, we can write

$$R_{XY} = R_{\pi/4}{}^{-1} Q_1 L_1 Q_2 L_2 Q_3 L_2 Q_2 L_1 Q_1 R_{\pi/4}. \tag{233}$$

Since the product of drifts and quadrupoles is a 2×2 block

diagonal, we have

$$R_{XY} = R_{\pi/4}{}^{-1} \begin{pmatrix} C & 0 \\ 0 & \bar{C} \end{pmatrix} R_{\pi/4}$$

$$= \frac{1}{2} \begin{pmatrix} C + \bar{C} & C - \bar{C} \\ C - \bar{C} & C + \bar{C} \end{pmatrix}. \tag{234}$$

The 2×2 matrix C is easy to calculate and $\bar{C}$ is obtained from C by replacing $q_i \to -q_i$. The condition that R_{XY} gives an emittance exchange is

$$\bar{C} = -C. \tag{235}$$

This gives

$$q_1 = \frac{d_2 q_2}{-d_1 - d_2 + d_1{}^2 d_2 q_2{}^2}, \tag{236}$$

$$q_3 = -\frac{d_1 + d_2 + d_1{}^2 d_2 q_2{}^2}{d_1 d_2 q_2 (d_1 + d_2)}. \tag{237}$$

The condition that C is a pure drift gives

$$d_2 = \frac{-1 + \eta^2 - \sqrt{1 + 2\eta^2 - 4\eta^3 + \eta^4}}{2 q_2 \eta (-1 + \eta)}; \quad \eta = d_1 q_2. \tag{238}$$

Lastly, the drift length becomes equal to the total drifts in the RHS of Eq. (230), that is, $d_T = 2(d_1 + d_2)$ if

$$d_1 = \frac{1}{9} \left[1 - \frac{4 \times 2^{2/3}}{\left(-67 + 9\sqrt{57}\right)^{1/3}} + \right.$$
$$\left. 2^{1/3} \left(-67 + 9\sqrt{57}\right)^{1/3} \right] \frac{1}{q_2} \approx \frac{-0.469}{q_2}. \tag{239}$$

The system is therefore completely determined by specifying the value of q_2 (Kim, 2020). The final transform matrix is

$$R_{XY} = \begin{pmatrix} 0 & 0 & 1 & d_T \\ 0 & 0 & 0 & 1 \\ 1 & d_T & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}. \tag{240}$$

The exchange of transverse emittance via five skew quadrupoles in a transfer line prior to injection to the storage ring was numerically studied in (Kuske and Kramer, 2016) and (Armborst, 2016) for the BESSY II machine. The purpose of the manipulation was to exchange the smaller vertical emittance with the larger horizontal emittance prior to injection in order to stay within the acceptance of the ring. However, due the additional dispersion generated, the authors concluded that a different way of transverse emittance exchange, i.e., "resonance crossing" (Aiba and Kallestrup, 2020; Carli *et al.*, 2002), was more feasible. Indeed, the emittance exchange via resonance crossing in a booster synchrotron was demonstrated in (Kallestrup and Aiba, 2020). On the other hand, in the upgrade of the Advanced Photon

Source at Argonne National Lab., a horizontal on-axis injection scheme from the booster to the storage ring has been adopted. The lattice of the booster to storage ring transport beamline was successfully designed to achieve transverse emittance exchange using a set of skew quadrupoles based (Kuske and Kramer, 2016).

C. Transverse-to-longitudinal phase-space exchangers

1. Emittance exchange

Transverse-to-longitudinal emittance exchange was first proposed by (Cornacchia and Emma, 2002) as a means to mitigate the microbunching instability in bright electron beams. The rationale there was that, generally, the transverse emittance produced by a state-of-the-art photoinjector is larger than desired for FEL generation, while the beam energy spread is smaller than necessary in order to avoid gain reduction in an x-ray FEL. As a result, they proposed a transverse-to-longitudinal phase-space exchange (i) to increase the slice energy spread, thereby reducing their sensitivity to the microbunching instability, while simultaneouly, to reduced the transverse emittance to improve the FEL performance (e.g., reduce the gain length). Utilization of emittance exchange in a compact XFEL facility was proposed in (Graves *et al.*, 2018). It was also suggested as a way to mitigate BBU instabilities and to improve efficiency in energy-recovery linacs (Piot, 2009). Mathematically, we follow a similar approach to the one described in Section VI.B.3 and seek a transformation that provides a matrix of the form

$$R_{XZ} = \begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix};$$
$$\begin{pmatrix} \mathbf{X} \\ \mathbf{Z} \end{pmatrix} \mapsto R_{XZ} \begin{pmatrix} \mathbf{X} \\ \mathbf{Z} \end{pmatrix} = \begin{pmatrix} A\mathbf{Z} \\ B\mathbf{X} \end{pmatrix}. \quad (241)$$

A possible beamline capable of providing such a transformation was first discussed in (Cornacchia and Emma, 2002). The beamline consists of a horizontal deflecting cavity located at the symmetry point of a chicane beamline; see Fig. 45(a). Using the transfer matrix of a deflecting cavity introduced in Section II.B.3, the overall transfer matrix of the system depicted in Fig. 45(a) is

$$R_{EEX} = R_{-DL} R_{TDC} R_{DL}, \quad (242)$$

where R_{DL} is the matrix of a dogleg under the small bending angle approximation; see Eq. (43) and R_{-DL} is the matrix of a reversed dogleg (i.e., R_{-DL} is obtained from R_{DL} via the substitution $\eta \mapsto -\eta$).

In (Cornacchia and Emma, 2002) they show that the condition $\kappa = 1/\eta$ (here η is the horizontal dispersion generated by one dogleg) causes most of the elements associated with the 2×2 anti-diagonal blocks of R_{EEX} to vanish for a thin-lens TDC. From this, they show that
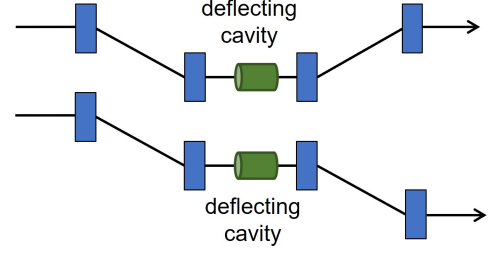

Figure 45 Top: Partial emittance exchange using a deflecting cavity and and chicanes. Bottom: Complete emittance exchange using a deflecting cavity and double doglegs.

the emittance is mapped as

$$(\varepsilon_{x,0}, \varepsilon_{z,0}) \mapsto (\varepsilon_x, \varepsilon_z)$$
$$= (\sqrt{\varepsilon_{z,0}^2 + \lambda^2 \varepsilon_{x,0}\varepsilon_{z,0}}, \sqrt{\varepsilon_{x,0}^2 + \lambda^2 \varepsilon_{x,0}\varepsilon_{z,0}}), \quad (243)$$

where λ is a coupling term that can be minimized via a proper choice of the incoming horizontal and longitudinal phase-space parameters. For $\lambda \ll 1$ we note that the beamline approximately exchanges the horizontal and longitudinal emittances $[(\varepsilon_{x,0}, \varepsilon_{z,0}) \xrightarrow{R_{EEX}} (\varepsilon_{z,0}, \varepsilon_{x,0})]$. It was subsequently recognized by (Kim and Sessler, 2006) that flipping the second half of the chicane so that the TDC is flanked by two horizontally dispersive sections arranged as doglegs [see Fig. 45(b)] produces an ideal exchange (taking $\ell_c = 0$). The emittance exchange condition is found to be

$$\kappa\eta + 1 = 0, \quad (244)$$

and the resulting transformation takes the form

$$R_{EEX} = \begin{pmatrix} 0 & \frac{\ell_c}{4} & -\frac{\ell+\frac{\ell_c}{4}}{\eta} & \frac{\eta^2 - \frac{\xi(4\ell+\ell_c)}{4}}{\eta} \\ 0 & 0 & -\frac{1}{\eta} & -\frac{\xi}{\eta} \\ -\frac{\xi}{\eta} & \frac{-\ell\xi - \frac{\ell_c\xi}{4} + \eta^2}{\eta} & \frac{\ell_c\xi}{4\eta^2} & \frac{\ell_c\xi^2}{4\eta^2} \\ -\frac{1}{\eta} & -\frac{\ell+\frac{\ell_c}{4}}{\eta} & \frac{\ell_c}{4\eta^2} & \frac{\ell_c\xi}{4\eta^2} \end{pmatrix}, \quad (245)$$

where ℓ_c and ℓ are, respectively, the TDC length and the distances between the dogleg dipole magnets. Under the thin lens approximation ($\ell_c = 0$), the latter matrix is 2×2-block anti-diagonal. In the most general case, it is possible to exchange the beam emittance by proper choice of the phase-space correlation on the incoming bunch. Specifically, the emittance mapping adopts the same form as Eq. (243) with the coupling term

$$\lambda^2 = \frac{\ell_c^2 (1 + \alpha_{x,0}^2)[\xi^2 + (\xi\alpha_{z,0} - \beta_{z,0})^2]}{16\eta^2 \beta_{x,0}\beta_{z,0}}. \quad (246)$$

Here $\alpha_{i,0}$ and $\beta_{i,0}$ are the initial Courant-Snyder parameters associated with the horizontal ($i = x$) and longitudinal ($i = z$) degrees of freedom. The quantity λ^2 can be

Table II Direct measurements of horizontal transverse (x) to longitudinal (z) emittance exchange compared to simulation. Emittance measurements are in units of µm and are normalized. From (Ruan *et al.*, 2011); experiment performed using the configuration displayed in Fig. 43(a).

	Simulated		Measured	
	In	Out	In	Out
ε_x^n	2.9	13.2	2.9 ± 0.1	11.3 ± 1.1
ε_y^n	2.4	2.4	2.4 ± 0.1	2.9 ± 0.5
ε_z^n	13.1	3.2	13.1 ± 1.3	3.1 ± 0.3

minimized by a proper choice of either the longitudinal or transverse C-S parameters. A possible solution consists of tuning the incoming LPS chirp to fullfill $\alpha_{z,0} = \frac{\beta_z}{\xi}$, corresponding to an incoming longitudinal phase space chirp $\mathcal{C} \equiv \frac{d\delta}{dz}\big|_0 = -\frac{1}{\xi}$ that produces a minimum bunch length at the cavity location (Sun *et al.*, 2007). Such a beamline was numerically investigated in (Emma *et al.*, 2006) to confirm its ability to mitigate the microbunching instability in an X-ray FEL.

This double-dogleg configuration was used to experimentally demonstrate a near-ideal horizontal-to-longitudinal emittance exchange at the Fermilab A0 photoinjector (Ruan *et al.*, 2011) with final results summarized in Table II.

Further work on improving the EEX beamline by (Zholents and Zolotorev, 2011) proposed combining the TDC with two accelerating-mode cavities for providing a simple way of cancelling the thick-lens effect of the TDC, which induces beam energy gain, as detailed in Section II.B.3. The design also developed a chicane-based exhanger with quadrupole magnets to control the dispersion and circumvent the limitation associated with the early design (Cornacchia and Emma, 2002). A similar design was later discussed in (Xiang and Chao, 2011). Ultimately, the capability to exchange emittance is limited by higher-order effects and requires the addition of higher-order multipole magnets, as discussed in (Nanni and Graves, 2015). Also, in the shaping process, collective effects such as CSR can significantly reduce the shaping quality (Carlsten *et al.*, 2011b; Ha *et al.*, 2016; Zholents and Zolotorev, 2011). Here, the exchange introduces extra difficulty than other well-studied beamlines such as chicane. Researches to control CSR's impact on the shaping are underway (Ha *et al.*, 2018a,b, 2017b; Malyzhenkov and Scheinker, 2018).

2. Current profile shaping

An important application of emittance exchange is its potential capability for shaping the beam's temporal distribution with unprecedented versatility and precision (Piot *et al.*, 2010). Considering Eq. (245) with

$\ell_c = 0$, we find that the initial phase space coordinates $(\mathbf{X}_0, \mathbf{Y}_0)^T$ of an electron will be mapped to final coordinates $(\mathbf{X}, \mathbf{Y})^T = R_{EEX}(\mathbf{X}_0, \mathbf{Y}_0)^T$. In particular, the electron's final longitudinal coordinates $\mathbf{Z} = (z, \delta)$ are solely functions of its initial transverse coordinates $\mathbf{X}_0 = (x_0, x_0')$:

$$\begin{cases} z = -\frac{\xi}{\eta}x_0 - \frac{\ell\xi - \eta^2}{\eta}x_0' \\ \delta = -\frac{1}{\eta}x_0 - \frac{\ell}{\eta}x_0' \end{cases}. \tag{247}$$

Exploiting the mapping described by these equations, one can produce arbitrarily shaped current or energy profiles by controlling the incoming transverse phase space. Specifically, the incoming phase-space distribution in $\mathbf{X}$ is mapped into the $\mathbf{Z}$ space via

$$\Phi_z(\mathbf{Z}) = \Phi_x(B^{-1}\mathbf{Z}), \tag{248}$$

where the subscript of the function Φ indicates to which of the two-dimensional phase spaces the function corresponds. Consequently, many of the techniques discussed in Section IV can be readily applied. We can, for instance, use a mask to shape the incoming distribution in the $\mathbf{X}$ space and map this distribution into the $\mathbf{Z}$ plane as originally proposed in (Piot *et al.*, 2010) and further discussed in (Jiang *et al.*, 2011). In principle, arbitrary bunch current distributions can be achieved using the EEX technique (Piot *et al.*, 2011b), but ultimately the incoming emittance partition limits the shaping resolution, as discussed in Section III; see Fig. 46.
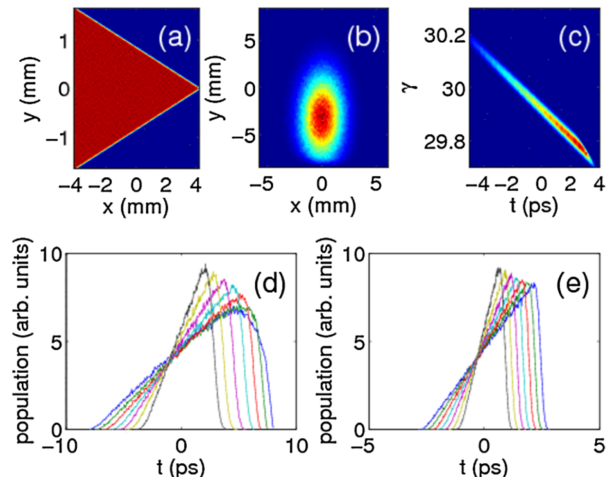
Figure 46 Example of simulated generation of a linearly ramped current profile from an initial uniform triangular distribution in (a) shown with corresponding final transverse distribution (b) and longitudinal phase space (c) downstream of the EEX beamline. The current profiles after EEX are shown in plots (d) and (e) for initial horizontal and longitudinal profiles equal to (10,1) µm and (1,10) µm, respectively. From (Piot *et al.*, 2011b).

An experimental demonstration of the shaping capability of an EEX beamline was demonstrated in (Sun *et al.*,

2010b), where the incoming beam was transversely sliced into multiple beamlets using a multislit plate and sent through an EEX beamline as diagrammed in Fig. 43(a). Doing so produced a train of subpicosecond bunches with tunable separation as depicted in Fig. 47, where an autocorrelation of coherent transition radiation emitted by the beam downstream of the EEX beamline confirmed the beam was temporally modulated when the incoming beam was intercepted by the multislit mask located upstream of the EEX beamline. In addition, the experi-
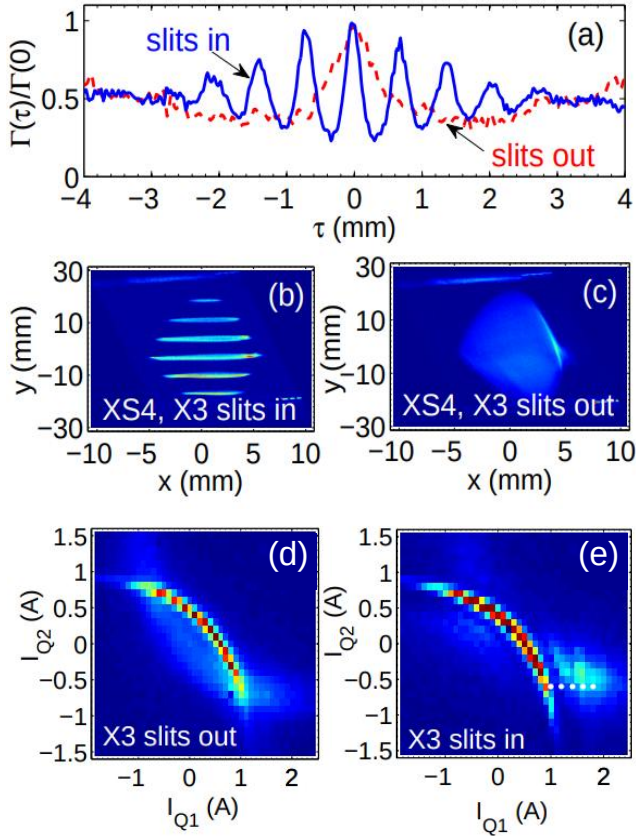
Figure 47 Experimental demonstration of sub-picosecond bunch train generation with an EEX beamline. (a) Normalized autocorrelation function $\Gamma(\tau)/\Gamma(0)$ of the CTR signal recorded with (solid) and without (dashed) insertion of a multi-slit mask upstream of the EEX beamline (τ is the optical path difference). The corresponding beam transverse densities downstream of the EEX appear in (b) and (c). The vertical axis on (b) and (c) is proportional to the beam's fractional momentum spread (δ). The nominal bunch charge is 550 ± 30 pC and reduces to $\sim 15 \pm 3$ pC when the slits are inserted. Total normalized CTR energy detected at X24 as a function of quadrupole magnet currents I_{Q_1} and I_{Q_2} with X3 slits out (d) and in (e) the beamline. The bolometer signal is representative of the inverse of the bunch duration σ_t. The intensity island appearing at $(I_{Q_1}, I_{Q_2}) \simeq (1.5, -0.5)$ in (e) is indicative of a density-modulated bunch. From (Sun *et al.*, 2010a).

ment demonstrated that tuning the upstream quadrupole magnets Q1 and Q2 (labeled as N6 and N7 in Fig. 43) provided control over the final LPS correlation and could transfer the incoming transverse density modulation into an energy or temporal modulation as discussed in (Sun *et al.*, 2010a); see also Fig. 47(b-e). This setup was also employed to investigate the generation of narrowband coherent transition radiation with tunable central wavelength (Piot *et al.*, 2011a).

LPS shaping was further investigated at the Argonne Wakefield Accelerator (AWA) (Ha *et al.*, 2017a) in the context of producing a beam suitable to improve the efficiency of a beam-driven wakefield accelerator. The EEX beamline used in the first generation of experiments at AWA adopted the double-dogleg configuration depicted in Fig. 48 using a 48-MeV high-charge bunch. Four quadrupole magnets downstream of the linac were used to manipulate the transverse beam phase space prior to the exchanger. A set of 100-μm-thick insertable tungsten masks of various shapes located ~0.2 m upstream of the exchanging beamline were used to demonstrate high-precision control over the final temporal shape of the bunch. Each dogleg provided $\eta \simeq 0.9$ m and the TDC was a 1/2+1+1/2-cell L-band cavity. The experimental setup incorporated full, single-shot LPS diagnostics that were used to directly measure the produced temporal distribution downstream of the exchanger. Figure 48 (g-j) showcases some examples of the final current distribution experimentally achieved with four different mask shapes. After these initial experimental demonstrations, AWA's EEX shaping beamline was then used set new records for transformer ratio. The beamline produced ramped current profiles for driving the wakefield in a dielectric structure (Gao *et al.*, 2018) and a plasma medium (Roussel *et al.*, 2020) with transformer ratios of $\mathcal{R} \simeq 4.5$ and $\mathcal{R} \simeq 7.8$ respectively.

The most recent developments on EEX have focused on using this class of beamline to form nanometer-scale modulations for coherent-radiation emission at X-ray wavelengths either by combing structured beams formed using a structure, e.g., field-emission-array, photocathode (Graves *et al.*, 2012), or by impressing beam modulation in the transverse phase space upstream of the exchanging beamlines which is then converted in a modulation on the LPS (either along the energy or temporal direction). In (Nanni *et al.*, 2018) a transmission mask such as described in Section IV.E.2 is employed to produce the initial transverse modulation while (Ha *et al.*, 2019b) explores the use of a transverse wiggler as shown in Fig. 26.

3. Bunch compression

The final bunch length downstream of the EEX is $\sigma_{z,f} = [R_{EEX,33}^2 \sigma_x^2 + R_{EEX,44}^2 \sigma_{x'}^2 + $
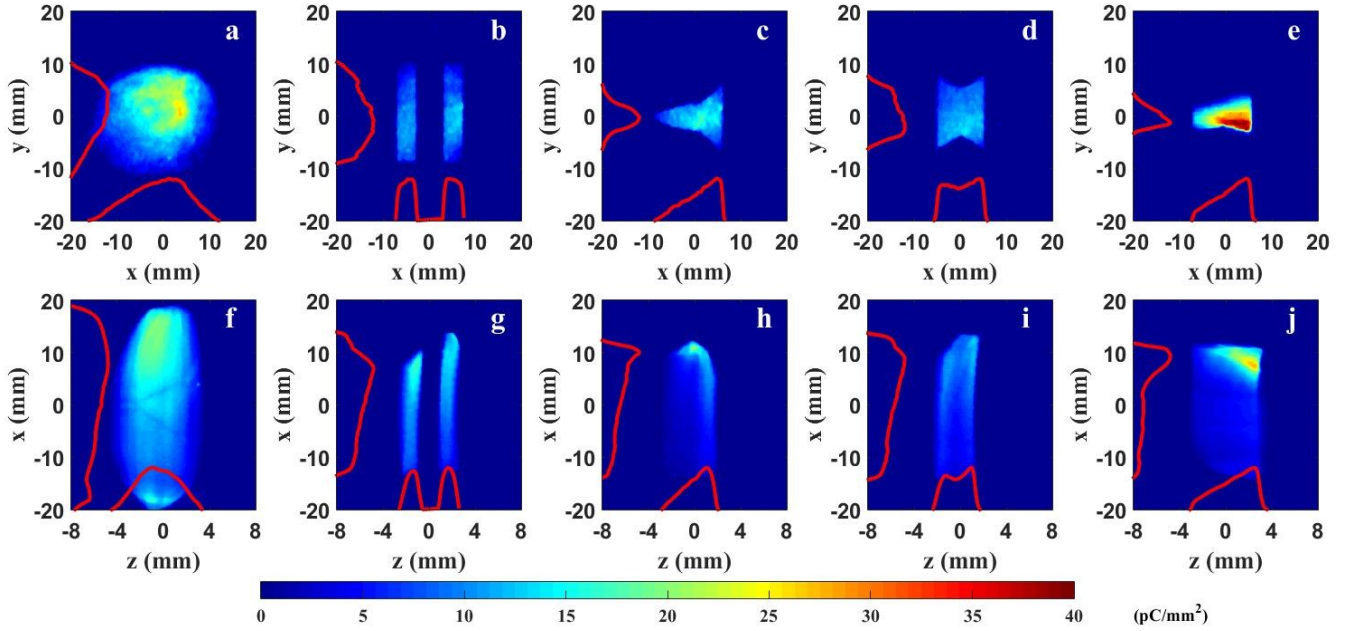
Figure 48 Experimental demonstration of versatile bunch shaping using an EEX beamline at AWA. (a-j) Beam shaping experiments using different transverse masks at the AWA EEX beamline. Images (a-e) are upstream of the EEX, and images (f-j) are downstream of the EEX. From (Ha *et al.*, 2016).

$2R_{EEX,33}R_{EEX,44}\langle xx'\rangle]$. Substituting Eq. (245), with the assumption $\ell_c = 0$ for simplicity, and expressing all beam quantities in term of ε_x and C-S parameters (α_x, β_x) associated with the horizontal phase space upstream of the EEX beamline, we obtain

$$\sigma_{z,f} = \sqrt{\frac{\varepsilon_x}{\beta_x}} \left[\left(\frac{\xi}{\eta}\beta_x - \left(\eta - \frac{\ell\xi}{\eta}\right)\alpha_x \right)^2 + \left(\eta - \frac{\ell\xi}{\eta}\right)^2 \right]^{1/2}. \quad (249)$$

The latter part of this equation indicates that upon choosing the proper incoming phase space correlation $\alpha_x/\beta_x = \xi/(\eta^2 - \ell\xi)$, the final bunch length becomes $\sigma_{z,f} = |\eta - \frac{\ell\xi}{\eta}|\sqrt{\frac{\varepsilon_x}{\beta_x}}$, which depends on the EEX design and can be made very small. Based on this observation, (Carlsten *et al.*, 2011b) investigated the performance of an EEX beamline for producing short bunches. The reference specifically discusses design choices that lead to extremely short final bunch duration $\sigma_{z,f}/c$ at the subfemtosecond time scale. The advantages of EEX beamlines employed for compression includes the reduced susceptibility to CSR-induced effects (including microbunching instability and bunch-length broadening), and elimination of the need for an initial longitudinal-phase-space chirp for compression using a chicane as well as any residual energy-phase correlation after compression. The drawback of an EEX-based bunch compressor is that the final horizontal phase space is determined by the incoming longitudinal phase space, and

the final transverse partition is generally not symmetric, which may be problematic for some applications, such as FELs (Emma *et al.*, 2006). In addition, any longitudinal phase space jitter (e.g., timing, energy and energy spread) will be manifested in the transverse phase space, resulting in beam position and spot size jitter (Ha *et al.*, 2017b).

4. Double phase-space exchangers

A limitation of the EEX technique for shaping the beam profile stems from the associated emittance exchange between, e.g., the horizontal and longitudinal phase spaces. To circumvent this limit, a possible configuration consists of a beamline composed of two concatenated EEX beamlines providing the mapping $(\mathbf{X}_0, \mathbf{Z}_0) \mapsto (\mathbf{Z}_1, \mathbf{X}_1) \mapsto (\mathbf{X}_2, \mathbf{Y}_2)$. Consequently, inserting a mask at location "1" produces a shaped beam downstream of the beamline while ideally leaving the incoming emittance unaltered. In practice such a method is not viable as collective effects such as space charge and CSR ultimately dilute the bending-plane emittance.

Despite tihs limitation, another advantage of the double phase-space exchanger configuration was recognized in (Zholents and Zolotorev, 2011) as providing a way to transparently tune the final LPS parameters without requirements on the incoming LPS. To understand such an application, we can write the matrix of the beamline as a block diagonal matrix flanked by two matrices rep-

resenting the EEX using Eq. (245):

$$R_{DEEX} = R_{EEX} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} R_{EEX}, \qquad (250)$$

assuming $L_c = 0$ so that $R_{EEX} \simeq \begin{pmatrix} 0 & M \\ N & 0 \end{pmatrix}$. Consequently, the total matrix of the double EEX (DEEX) simplifies as

$$R_{DEEX} = \begin{pmatrix} MBN & 0 \\ 0 & NAM \end{pmatrix} \overset{\gamma \gg 1}{\Rightarrow} \begin{pmatrix} MN & 0 \\ 0 & NAM \end{pmatrix}, \tag{251}$$

which is a 2×2 block diagonal matrix and confirmed that the DEEX beamline does not provide any global coupling. The simplification $B = I$ in the ultra-relativistic limit $\gamma \gg 1$ comes from the longitudinal dispersion associated with a longitudinal a drift and scales as $-\ell/\gamma^2$ where ℓ is the length of the telescope beamline. These results are generally available even when $\ell_c \neq 0$ as long as thick-lens effects are corrected via, e.g., the addition of accelerating-mode cavities (Zholents and Zolotorev, 2011). Equation (251) suggests a simple way of tuning the final LPS correlation of the beam by properly designing the insertion beamline optical lattice between the two EEXs. The second EEX simply converts the transversely fine-tuned beam via the insertion beamline to the longitudinal phase space. Focusing on the LPS mapping and treating the insertion beamline as a telescope, i.e. considering A to be diagonal with element $A_{11} = 1/A_{22} \equiv a$, we find

$$NAM = \begin{pmatrix} \frac{\ell\xi + a^2(\ell\xi - \eta^2)}{a\eta^2} & \frac{\xi(a^2+1)(\ell\xi - \eta^2)}{a\eta^2} \\ \frac{\ell(a^2+1)}{a\eta^2} & \frac{\ell a^2\xi + \ell\xi - \eta^2}{a\eta^2} \end{pmatrix}. \tag{252}$$

As an example, considering the case when the beamline is designed such that $\ell(a^2+1)/(a\eta^2) \ll 1$, the above matrix is approximately

$$NAM \simeq \begin{pmatrix} -a & \xi\left(a + \frac{1}{a}\right) \\ 0 & -\frac{1}{a} \end{pmatrix}; \tag{253}$$

therefore, the final rms bunch length is

$$\sigma_z = [\sigma_{z,0}^2 a^2 + \xi^2(a + 1/a)^2 \sigma_{\delta,0}^2$$
$$- 2a\xi(1 + a^2)\langle z_0\delta_0 \rangle]^{1/2}. \tag{254}$$

An important consequence of Eq. (254) is that bunch compression can be accomplished without any longitudinal-phase-space chirp (i.e. $\langle z_0\delta_0 \rangle = 0$) and designing the beamline such that $|a^2 + \xi^2(a+1/a)^2| < 1$. A numerical simulation of a double-EEX-based bunch compressor is shown in Figure 49.

Additionally, the configuration enables control of nonlinear correlation in the longitudinal-phase-space by using nonlinear magnets between the two EEXs. For instance, (Seok et al., 2019b) discusses various configurations to reduce the beam final energy spread via control of
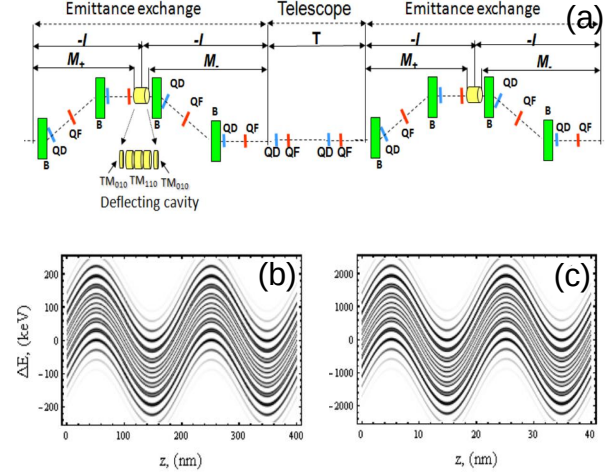
Figure 49 Schematic of a double-EEX-based bunch compressor (a) with a simulated example of an application to provide a 10-fold increase in modulation frequency of an incoming laser-modulated electron beam (b) and (c). From (Zholents and Zolotorev, 2011).

the longitudinal phase-space nonlinearities with a DEEX configuration. The method closely follows the technique described in Section IV.E.1. Here, the first EEX is applied to the phase space $(\mathbf{X}_0, \mathbf{Z}_0) \mapsto (\mathbf{X}_1 \simeq \mathbf{Z}_0, \mathbf{Z}_1 \simeq \mathbf{X}_0)$, a then a nonlinear transformation is applied to the horizontal phase space

$$x_1' = f(x_1, y_1), \tag{255}$$

and then the second EEX is applied to change back the longitudinal and transverse phase space $(\mathbf{X}_1, \mathbf{Z}_1) \mapsto (\mathbf{X}_2 \simeq \mathbf{Z}_1, \mathbf{Z}_2 \simeq \mathbf{X}_1) \simeq (\mathbf{X}_0, \mathbf{Z}_0)$. Figure 50 showcases a configuration capable of introducing a nonlinear phase-space correlation with a DEEX beamline. In this particular case, a sextupole magnet, judiciously located downstream of the first EEX, imposes a quadratic correlation. Extension of this technique to control higher-order distor-
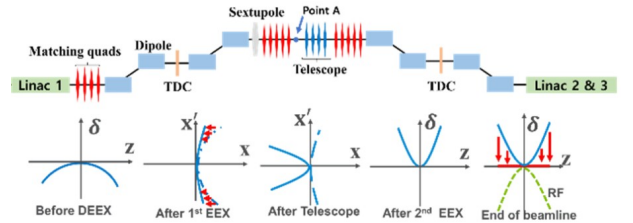
Figure 50 Overview of a DEEX-based energy-spread reduction via correction on quadratic nonlinearity with a sextupole magnet. From (Seok et al., 2019b).

tion is straightforward: correction of third order nonlinearities to, e.g., suppress current spike "horns" sometimes encountered in bright-beam injectors, was demonstrated via numerical simulations (Seok et al., 2019a) as an alternative to the technique described in Section IV.B.2;

Figure 51 Longitudinal phase spaces from numerical particle tracking. (a) shows the beam's initial phase space and (b) shows the phase space after a modulator. (c) and (d) show the final phase space after a double EEX beamline, and each of them corresponds to spectral and temporal bunching respectively. From (Seok *et al.*, 2021).

see also Ref. (Charles *et al.*, 2017).

Another recent development introduced in (Seok *et al.*, 2021) is using a DEEX beamline with a phase space modulator, such as wakefield str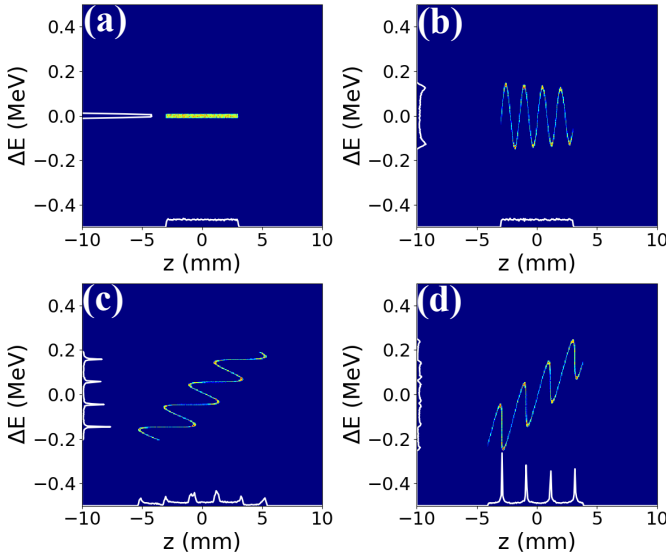ucture or transverse wiggler discussed in Sections IV and V. Similar to generating a bunch train from the energy modulation, this method can generate density spikes on both the time and energy axes. If we only consider a small fraction of the sinusoidal modulation that builds up the density spike, we can write this microbunch's longitudinal correlation as $\delta = hz$ (i.e., locally linear correlation with chirp h); see Fig. 51(b). In general, we can define the compression factor for this microbunch as $1/(R_{55} + hR_{56})$. Appropriate beamline parameters for a given local chirp can maximize the compression factor and build up density spikes. We can imagine the same situation for the energy distribution. The energy compression factor will be simply $1/(R_{65} + hR_{66})$. In the case of the chicane, R_{55} is 1 and R_{56} is non-zero, so there is a corresponding h. However, R_{65} is zero and R_{66} is 1, so no controls are allowed on the energy distribution. On the other hand, a double EEX beamline that provides non-zero R_{65} and R_{66} would be able to generate energy spikes. This beamline provides control over R_{65} and R_{66} using the quadrupole magnets between two EEX beamlines; see (Zholents and Zolotorev, 2011) and (Ha *et al.*, 2017c).

Figure 51 showcases numerical tracking results. Similar to the density modulation case, a dielectric structure in front of a double EEX beamline imprints a sinusoidal modulation on the longitudinal phase space. Then, this modulation is converted into energy spikes, controlled with quadrupole magnets located in the middle of the beamline. These energy spikes can be considered as several beams, each with small energy spread. Because the EEX beamline provides control of R_{55}, R_{56}, R_{65}, and R_{66}, it can be used to generate a multi-energy beam with controllable time separation from a single bunch. This may be a useful technique to generate multi-color radiation for various spectroscopy methods or pump-probe experiments.

D. Generalized phase-space repartitioning between the three degrees of freedom

1. Flat beam transformation combined with emittance exchange

Future electron-positron linear colliders (LCs) require unprecedented ultra-low transverse emittances. Additionally, the transverse emittance ratio should be high to mitigate beamsstrahlung effects (Yokoya and Chen, 1992). The present requirements from high-energy physics call for 80% spin-polarized electron beams. The electron bunch charge range from fC to nC depends on the LC technology choice. The vertical emittance required to mitigate beamsstrahlung effects is attained by injecting the beam into a damping ring to decrease its transverse emittance via radiative cooling, yielding a "flat" beam with a transverse emittance ratio ranging from 300 to 500, depending on the LC concept. Taking the example from the International Linear Collider (ILC) (Adolphsen *et al.*, 2013), the final emittance partition of $(\varepsilon_x^n, \varepsilon_y^n, \varepsilon_z^n) = (5.5, 20 \times 10^{-3}, 6.5 \times 10^4)$ µm corresponding to a six-dimensional emittance $\varepsilon_{6D}^n = \varepsilon_x^n \varepsilon_y^n \varepsilon_z^n \simeq 7150$ µm³ is generally larger than values typically achieved by state-of-the-art photoinjectors (e.g., for 3.2 nC bunch charges). Consequently, there has been some effort to develop a damping-ring-free electron injector for a future linear collider by combining the round-to-flat-beam transformer (RFTB) and EEX beamlines. Specifically, (Xu *et al.*, 2021) considers a 3.2-nC bunch produced in an L-band photoinjector and demonstrates via numerical simulation that a transverse intrinsic emittance partition of $(\varepsilon_+^n, \varepsilon_-^n, \varepsilon_z^n) \sim (500, 10 \times 10^{-3}, 10)$ µm assumes an ideal mapping of the intrinsic emittances $(\varepsilon_x^n, \varepsilon_y^n) = (\varepsilon_1^n, \varepsilon_2^n)$ using an RFBT beamline. The author suggests that the downstream EEX beamline would produce an emittance partition $(\varepsilon_x^n, \varepsilon_y^n, \varepsilon_z^n) \sim (10, 10 \times 10^{-3}, 500)$ µm.

2. Coupling between the longitudinal and transverse phase spaces

So far we have discussed beamlines capable of phase-space swapping between the two transverse phase space

planes $[(x,x') \leftrightarrow (y,y')]$ or between one of the transverse phase space planes and the longitudinal plane $[(x,x') \leftrightarrow (z,\delta)]$. We also presented a method to repartition the phase space between (x,x') and (y,y') using a magnetized beam and an RFBT beamline. We now examine whether a similar transformation is possible between, e.g., (x,x') and (z,δ) following (Carlsten *et al.*, 2011a). In this work, the authors noticed that the matrix of a deflecting cavity in (x,x',z,δ) under the thin lens approximation (i.e., described by Eq. (65) with $L=0$) is similar to the matrix of a skew-quadrupole magnet in (x,x',y,y'). Consequently, introducing an initial correlation similar to the one described for the case of a magnetized beam (i.e., by Eq. (228)) in the (x,x',z,δ) plane combined with several deflecting cavities separated by "drift" (with matrix described by Eq. (43) with $\eta=0$) could repartition the emittances between the (x,x') and (z,δ) planes. Specifically, (Carlsten *et al.*, 2011a) consider the practical case of a tilted-front laser pulse impinging on a photocathode so that the initial correlation is imparted in the (x,x',z,δ) phase space of the form

$$z_c = z + \tau x, \qquad (256)$$

where τ is a parameter representing the $x-z$ correlation introduced by the tilt. The corresponding beam matrix at the cathode location is written as

$$\Sigma_c = \begin{pmatrix} \sigma_x^2 & 0 & \tau\sigma_x^2 & 0 \\ 0 & \sigma_x'^2 & 0 & 0 \\ \tau\sigma_x^2 & 0 & \sigma_z^2+\tau^2\sigma_x^2 & 0 \\ 0 & 0 & 0 & \sigma_\delta^2 \end{pmatrix}. \qquad (257)$$

The corresponding intrinsic emittances are

$$\begin{aligned} \varepsilon_x &\simeq \sigma_x\sigma_\delta\tau, \\ \varepsilon_z &\simeq \frac{\sigma_x'\sigma_z}{\tau}, \end{aligned} \qquad (258)$$

where it is assumed that $\tau^2 \gg \frac{\sigma_x'\sigma_z}{\sigma_x\sigma_\delta}$. Then (Carlsten *et al.*, 2011a) shows that this initial coupling can be removed by locating four deflecting cavities downstream of the electron source to map the intrinsic emittance to conventional emittances. The cavities are separated with drifts with longitudinal dispersion (e.g. magnetic chicane described by Eq. (43) with $\eta=0$).

This approach can be further expanded by introducing both a transverse and longitudinal tilt to introduce arbitrary emittance repartitioning within the three degrees of freedom (Yampolsky *et al.*, 2010).

VII. FUTURE DIRECTIONS

The techniques described in this paper along with that discussed in (Hemsing *et al.*, 2014) have open the path to finer control over the beam's phase-space distribution beyond the ensemble-averaged techniques. Further development of phase-space tailoring methods will ultimately aim to provide full six-dimensional control of the phase-space distributions, possibly enabling the design of a tailored beam at the single-particle level. Examples of possible expansions based on recent developments are presented in this section for each of the beam-shaping classes considered. The list below is by no means exhaustive given the vigorous on-going research on the topic.

A. Ab-initio shaping

Controlling the final beam distribution by properly programming the initial conditions is widespread owing to its versatility and easy implementation, as discussed in Section III. Further expanding the resolution of this class of methods will benefit from new electron-source emission concepts, integrated photonic progress, and advances in ultrafast-laser systems.

Over the last decade, new ultra-cold electron sources based on trapped atoms have emerged (Claessens *et al.*, 2005; Zolotorev *et al.*, 2007). Although the primary motives of this work were related to the generation of electron beam close to quantum degeneracy, an experiment demonstrated a high degree of control over the beam distribution (McCulloch *et al.*, 2011) via shaping of the ionization laser using a spatial light modulator similar to the one discussed in Section III. These sources typically produce electron beams with keV energies, and their integration into relativistic electron sources remains a challenge (van der Geer *et al.*, 2014). Yet they will provide another toolkit for ab-initio shaping of the beam distribution.

Likewise, novel laser architecture based on cohrently combining laser pulses that have been individually controlled (Lemons *et al.*, 2021), could form laser pulses with finer structure to precisely program the photoemission process. Although it is ultimately limited by cathode response time and intrinsic emittance.

Finally, recent progress in integrated nanophotonic fabrication (Komljenovic *et al.*, 2016) has opened the path to the fabrication of pixelized cathodes (Blankemeier *et al.*, 2019). This type of cathode could ultimately support initial control over the emitted beam spatiotemporal distribution with unprecedented spatial and temporal scales.

B. Controlling the beam via external fields

Over the years the possibility of shaping the electromagnetic-field distribution using metamaterials has flowed in many applications, most notably optics, and has led to the development of "transformation electromagnetics" (Werner and Kwon, 2014), a branch of electromagnetism aimed at tailoring electromagnetic fields. These techniques have recently been applied to the de-

sign of electromagnets with unusual properties (e.g., producing a negative magnetic permeability) that could expand the methods described in Section IV. For instance, (Mach-Batlle *et al.*, 2020) experimentally demonstrated that an active magnetic metamaterial can emulate the field of a straight current wire at a distance. Such a demonstration opens the way to manipulating magnetic fields in inaccessible regions. It is expected that these emerging technologies will be critical to further the development of beam-control techniques based on external electromagnetic fields. Finally, the discovery of knotted solutions to Maxwell's equations (Kedia *et al.*, 2013; Ranada, 1990) could also have applications to beam shaping as they provide faster (e.g., localized) spatiotemporal variations of the field compared to conventional "plane-wave" synthesis solutions. Experimentally producing these knotted solutions is challenging as the solutions can be mathematically formulated as a summation of spherical-harmonic functions. A possible approximate experimental implementation discussed in (Irvine and Bouwmeester, 2008) uses tightly focused circularly-polarized laser pulses.

C. Shaping the beam using collective effects

Similar to the discussion in the previous section, we expect engineered new materials to play a critical role in fostering more precise control over the beam using the beam's self field. The main advantages of beam shaping techniques based on wakefield-driven structures reside in their ability to (*i*) support electromagnetic fields with wavelengths comparable to the bunch length (which is challenging in conventional RF cavities) and (*ii*) ensure these fields are carrier-envelope-phase (CEP) locked with the bunch, thereby alleviating the need for precise external synchronization and consistently mitigating possible issues associated with shot-to-shot jitter. Over the last two decades, wakefield structures based on metamaterial have been developed (Antipov *et al.*, 2007) and tested (Antipov *et al.*, 2008; Duan *et al.*, 2017; Lu *et al.*, 2019). Likewise, photonic band-gap (PBG) structures have been introduced (Smirnova, 2004). PBG structures can control the distribution of excited modes, e.g., to suppress noxious modes in a wakefield accelerators (Simakov *et al.*, 2016). It is conceivable that PBG structures designed to introduce a well-defined superposition of modes could find applications in bunch shaping (e.g., to synthesize the desired longitudinal or transverse force), expanding on techniques presented in Section V.

D. Redistributing phase space between planes

Exchanging phase space between two or three degrees of freedom has facilitated the generation of shaped beams

for, e.g., advanced acceleration concepts as showcased in Section VI. We note that so far the implementation of phase-space exchanging beamlines has often been based on a simple configuration and further optimization, e.g., to mitigate sources of 6D-emittance dilution would be critical to the deployment of this class of methods. Likewise, and similar to a laser-based method described in (Xiang, 2010), one could consider simpler versions of phase-space exchangers where the required time-dependent deflecting fields are introduced using transverse wakefields. Finally, extending the concept of phase-space exchange to control the beam at smaller time scales along the ideas discussed in (Graves *et al.*, 2019) could have ground-breaking consequences for the development of room-sized coherent X-ray sources.

LIST OF SYMBOLS

s, $\bar{s}$, s_2, τ	distance along the reference trajectory of the reference particle
x, y, z	coordinates relative to the reference particle
r, φ, s	cylindrical coordinates
x', y'	transverse angles
δ	relative momentum deviation
$\boldsymbol{\mathcal{Z}}$	6D canonical phase-space variables
ζ_i	ith component of $\boldsymbol{\mathcal{Z}}$
p_x, p_y, p_s	momenta
M	map corresponds to position s to $\bar{s}$
R	Jacobian or linear transformation matrix
e	electron charge
$\mathbf{E}$	electric field
$\mathbf{B}$	magnetic field
E_0	peak electric field strength
K_q	focusing strength of quadrupole magnet
η	dispersion
ξ	R_{56} of dogleg matrix
ω	angular frequency
k	wave number
λ	wavelength
Δ_p	fractional momentum increase
$\mathbf{A}$	vector potential
κ_L	(spatial) Larmor frequency
L	canonical angular momentum
$f(\boldsymbol{\mathcal{Z}}; s)$	normalized phase-space distribution of particles at s
ℓ	length of bunch or device
σ_i	root-mean square (rms) value of $i \in \mathcal{Z}$
$\lambda(z)$	normalized line charge density
$\lambda_\perp(x)$	normalized transverse charge density
C	compression factor
Σ	second-order-moment beam matrix
ε_i^{proj}	RMS emittance along $i \in [x, y, z]$
ε_i	intrinsic emittances where $i \in [1, 2, 3]$
ε_i^n	normalized emittance along $i \in [x, y, z]$
Q	bunch charge
$Z(k)$	impedance
Z_0	free-space impedance
$b(k; s)$	bunching factor
r_e	classical electron radius
λ_D	Debye length

ω_p	plasma frequency
h	longitudinal-phase-space chirp strength
K_{2n}	magnetic multipole strength
K_w	wiggler parameter
ϵ_0	permittivity in vacuum
N_b	total number of particles in bunch
I	current
I_A	Alfven current
$\mathcal{K}$	loss factor
$w(z)$	wake function
$W(z)$	wakefield of bunch

VIII. ACKNOWLEDGMENTS

REFERENCES

Adolphsen, Chris, Maura Barone, Barry Barrish, Karsten Buesser, and Philip Burrows (2013), "The International Linear Collider Technical Design Report - Volume 3.II: Accelerator Baseline Design," arXiv:1306.6328 [physics.acc-ph].

Aiba, Masamitsu, Michael Ehrlichman, and Andreas Streun (2015), "Round Beam Operation in Electron Storage Rings and Generalisation of Mobius Accelerator," in *6th International Particle Accelerator Conference*, p. TUPJE045.

Aiba, Masamitsu, and Jonas Kallestrup (2020), "Theory of emittance exchange through coupling resonance crossing," Phys. Rev. Accel. Beams **23**, 044003.

Akre, R, D. Dowell, P. Emma, J. Frisch, S. Gilevich, G. Hays, Ph. Hering, R. Iverson, C. Limborg-Deprey, H. Loos, A. Miahnahri, J. Schmerge, J. Turner, J. Welch, W. White, and J. Wu (2008), "Commissioning the linac coherent light source injector," Phys. Rev. ST Accel. Beams **11**, 030703.

Alvarez, L W (1946), "The design of a proton linear accelerator," Phys. Rev. **70**, 799.

Andonian, G, S. Barber, F. H. O'Shea, M. Fedurin, K. Kusche, C. Swinson, and J. B. Rosenzweig (2017), "Generation of ramped current profiles in relativistic electron beams using wakefields in dielectric structures," Phys. Rev. Lett. **118**, 054802.

Andrews, H, K. Nichols, D. Kim, E. I. Simakov, S. Antipov, G. Chen, M. Conde, D. Doran, G. Ha, W. Liu, J. Power, J. Shao, and E. Wisniewski (2020), "Shaped beams from diamond field-emitter array cathodes," IEEE Transactions on Plasma Science **48** (7), 2671–2675.

Andrews, H L, C. H. Boulware, C. A. Brau, and J. D. Jarvis (2005), "Dispersion and attenuation in a smith-purcell free electron laser," Phys. Rev. ST Accel. Beams **8**, 050703.

Antipov, S, M. Babzien, C. Jing, M. Fedurin, W. Gai, A. Kanareykin, K. Kusche, V. Yakimenko, and A. Zholents (2013), "Subpicosecond bunch train production for a tunable mj level thz source," Phys. Rev. Lett. **111**, 134802.

Antipov, S, S. Baturin, C. Jing, M. Fedurin, A. Kanareykin, C. Swinson, P. Schoessow, W. Gai, and A. Zholents (2014), "Experimental demonstration of energy-chirp compensation by a tunable dielectric-based structure," Phys. Rev. Lett. **112**, 114801.

Antipov, S, C. Jing, M. Fedurin, W. Gai, A. Kanareykin, K. Kusche, P. Schoessow, V. Yakimenko, and A. Zholents (2012), "Experimental observation of energy modulation in electron beams passing through terahertz dielectric wakefield structures," Phys. Rev. Lett. **108**, 144801.

Antipov, S, L. Spentzouris, W. Gai, M. Conde, F. Franchini, R. Konecny, W. Liu, J. G. Power, Z. Yusof, and C. Jing (2008), "Observation of wakefield generation in left-handed band of metamaterial-loaded waveguide," Journal of Applied Physics **104** (1), 014901.

Antipov, S, L. Spentzouris, W. Liu, W. Gai, and J. G. Power (2007), "Wakefield generation in metamaterial-loaded waveguides," Journal of Applied Physics **102** (3), 034906.

Armborst, Felix (2016), "Study and optimization of the transverse beam parameters before injection into the storage ring bessy ii," .

Arnold, VI (1978), *Mathematical Methods of Classical Mechanics* (Springer-Verlag New York).

Asaka, Takao, Hiroyasu Ego, Hirohumi Hanaki, Toru Hara, Taichi Hasegawa, Teruaki Hasegawa, Takahiro Inagaki, Toshiaki Kobayashi, Chikara Kondo, Hirokazu Maesaka, Shinichi Matsubara, Sakuo Matsui, Takashi Ohshima, Yuji Otake, Tatsuyuki Sakurai, Shinsuke Suzuki, Yasuyuki Tajiri, Shinichiro Tanaka, Kazuaki Togawa, and Hitoshi Tanaka (2017), "Low-emittance thermionic-gun-based injector for a compact free-electron laser," Phys. Rev. Accel. Beams **20**, 080702.

Aschikhin, A, C. Behrens, S. Bohlen, J. Dale, N. Delbos, L. di Lucchio, E. Elsen, J.-H. Erbe, M. Felber, B. Foster, L. Goldberg, J. Grebenyuk, J.-N. Gruse, B. Hidding, Zhanghu Hu, S. Karstensen, A. Knetsch, O. Kononenko, V. Libov, K. Ludwig, A.R. Maier, A. Martinez de la Ossa, T. Mehrling, C.A.J. Palmer, F. Pannek, L. Schaper, H. Schlarb, B. Schmidt, S. Schreiber, J.-P. Schwinkendorf, H. Steel, M. Streeter, G. Tauscher, V. Wacker, S. Weichert, S. Wunderlich, J. Zemella, and J. Osterhoff (2016), "The FLASHForward facility at desy," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **806**, 175 – 183.

Bane, K L F, P. Chen, and P. B. Wilson (1985), "On collinear wake field acceleration," IEEE Transactions on Nuclear Science **32** (5), 3524–3526.

Bane, Karl, and Gennady Stupakov (2016), "Dechirper wakefields for short bunches," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **820**, 156 – 163.

Bane, KLF, and G. Stupakov (2012), "Corrugated pipe as a beam dechirper," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **690**, 106 – 110.

Baryshev, Sergey V, Sergey Antipov, Jiahang Shao, Chunguang Jing, Kenneth J. Pérez Quintero, Jiaqi Qiu, Wanming Liu, Wei Gai, Alexei D. Kanareykin, and Anirudha V. Sumant (2014), "Planar ultrananocrystalline diamond field emitter in accelerator radio frequency electron injector: Performance metrics," Applied Physics Letters **105** (20), 203505.

Batygin, YK (1993), "Beam intensity redistribution in a nonlinear optics channel," Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms **79** (1), 770–772.

Benson, SV, *et al.* (2018), "Development of a Bunched Beam Electron Cooler Based on ERL and Circulator Ring Technology for the Jefferson Lab Electron-Ion Collider," in *Proc. of Workshop on Beam Cooling and Related Topics (COOL'17), Bonn, Germany, 18-22 September 2017*, Workshop on Beam Cooling and Related Topics No. 11 (JACoW, Geneva, Switzerland) pp. 72–76, https://doi.org/10.18429/JACoW-COOL2017-WEM12.

Bettoni, S, P. Craievich, A. A. Lutman, and M. Pedrozzi (2016), "Temporal profile measurements of relativistic electron bunch based on wakefield generation," Phys. Rev. Accel. Beams **19**, 021304.

Bich, A, J. Rieck, C. Dumouchel, S. Roth, K. J. Weible, M. Eisner, R. Voelkel, M. Zimmermann, M. Rank, M. Schmidt, R. Bitterli, N. Ramanan, P. Ruffieux, T. Scharf, W. Noell, H.-P. Herzig, and Nico De Rooij (2008), "Multifunctional micro-optical elements for laser beam homogenizing and beam shaping," in *Photon Processing in Microelectronics and Photonics VII*, Vol. 6879, edited by David B. Geohegan, Frank Träger, Jan J. Dubowski, Andrew S. Holmes, Michel Meunier, Craig B. Arnold, and Hiroyuki Niino, International Society for Optics and Photonics (SPIE) pp. 123 – 134.

Blankemeier, L, F. Rezaeifar, A. Garg, and R. Kapadia (2019), "Integrated photonics for low transverse emittance, ultrafast negative electron affinity gaas photoemitters," Journal of Applied Physics **126** (3), 033102.

Blaskiewicz, Michael (2014), "Cooling of high-energy hadron beams," Annual Review of Nuclear and Particle Science **64** (1), 299–317.

Blaskiewicz, Michael (2019), "3d theory of the plasma cascade instability," arXiv:1910.13300 [physics.acc-ph].

Brillouin, Leon (1945), "A theorem of Larmor and its importance for electrons in magnetic fields," Phys. Rev. **67**, 260–266.

Brinkmann, R, Y. Derbenev, and K. Flöttmann (2001), "A low emittance, flat-beam electron source for linear colliders," Phys. Rev. ST Accel. Beams **4**, 053501.

Brown, Karl L (1968), "A first and second order matrix theory for the design of beam transport systems and charged particle spectrometers," Adv. Part. Phys. **1**, 71–134.

Brown, Karl L (1999), "Single element optics," in *Handbook of Accelerator Physics and Engineering*, edited by Alexander Wu Chao and Maury Tigner, Chap. 2.2.1 (World Scientific, Singapore).

Budker, G I, Y. S. Derbenev, N. S. Dikansky, V. I. Kudelainen, I. N. Meshkov, V. V. Parkhomchuk, D. V. Pestrikov, B. N. Sukhina, and A. N. Skrinsky (1975), "Experiments on electron cooling," IEEE Transactions on Nuclear Science **22** (5), 2093–2097.

Burov, A, and V.Danilov (1998), *An Insertion to Eliminate Horizontal Temperature of High Energy Electron Beam*, Tech. Rep. Fermilab-TM-2043 (Fermilab).

Burov, Alexey, Sergei Nagaitsev, and Yaroslav Derbenev (2002), "Circular modes, beam adapters, and their applications in beam optics," Phys. Rev. E **66**, 016503.

Busch, H (1926), "Berechnung der bahn von kathodenstrahlen im axialsymmetrischen elektromagnetischen felde," Annalen der Physik **386** (25), 974–993.

Carey, DavidC (1987), *The Optics of Charged Particle Beams* (Hardwood Academic Publishers).

Carli, C, G. Cyvoct, M. Giovannozzi, E. Metral, G. Métral, and R. Steerenberg (2002), "Emittance exchange by crossing a coupling resonance," in *8th European Particle Accelerator Conference (EPAC 2002)*, pp. 1157–1159.

Carlsten, BE (1989), "New photoelectric injector design for the los alamos national laboratory xuv fel accelerator," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **285** (1), 313–319.

Carlsten, BE, and Kip Bishofberger (2006), "Simple algorithm for designing skew-quadrupole cooling configurations," New J. Phys. **8**, 286.

Carlsten, Bruce E, Kip A. Bishofberger, Leanne D. Duffy, Steven J. Russell, Robert D. Ryne, Nikolai A. Yampolsky, and Alex J. Dragt (2011a), "Arbitrary emittance partitioning between any two dimensions for electron beams," Phys. Rev. ST Accel. Beams **14** (5), 050706.

Carlsten, Bruce E, Kip A. Bishofberger, Steven J. Russell, and Nikolai A. Yampolsky (2011b), "Using an emittance exchanger as a bunch compressor," Phys. Rev. ST Accel. Beams **14**, 084403.

Carlsten, Bruce E, and Steven J. Russell (1996), "Subpicosecond compression of 0.1-1 nC electron bunches with a magnetic chicane at 8 Mev," Phys. Rev. E **53**, R2072–R2075.

Chao, AW (1993), *Physics of collective beam instabilities in high-energy accelerators* (Wiley, New York).

Charles, T K, D. M. Paganin, A. Latina, M. J. Boland, and R. T. Dowd (2017), "Current-horn suppression for reduced coherent-synchrotron-radiation-induced emittance growth in strong bunch compression," Phys. Rev. Accel. Beams **20**, 030705.

Cheon, Yoo-Lim, Seok-Ho Moon, Moses Chung, and Dong-O Jeon (2020), "Mitigation of Space-Charge Driven Fourth-Order Resonance by Beam Spinning," arXiv:2010.14080 [physics.acc-ph].

Claessens, B J, S. B. van der Geer, G. Taban, E. J. D. Vredenbregt, and O. J. Luiten (2005), "Ultracold electron source," Phys. Rev. Lett. **95**, 164801.

Cornacchia, M, S. Di Mitri, G. Penco, and A. A. Zholents (2006), "Formation of electron bunches for harmonic cascade x-ray free electron lasers," Phys. Rev. ST Accel. Beams **9**, 120701.

Cornacchia, M, and P. Emma (2002), "Transverse to longitudinal emittance exchange," Phys. Rev. ST Accel. Beams **5**, 084001.

Courant, ED (1966), *Perspectives in Modern Physics. Essays in honor of Hans A. Bethe on the occasion of his 60th birthday, July 1966. R. E. Marshak, Ed. Interscience (Wiley), New York. xii + 673 pp., illus.*

Courant, ED, and H.S Snyder (1958), "Theory of the alternating-gradient synchrotron," Annals of Physics **3** (1), 1–48.

Craievich, Paolo, and Alberto A. Lutman (2017), "Effects of the quadrupole wakefields in a passive streaker," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **865**, 55 – 59, physics and Applications of High Brightness Beams 2016.

Decking, W, *et al.* (2020), "A MHz-Repetition-Rate Hard X-Ray Free-Electron Laser Driven by a Superconducting Linear Accelerator," Nature Photonics **14** (6), 391 – 397.

Deng, Haixiao, Meng Zhang, Chao Feng, Tong Zhang, Xingtao Wang, Taihe Lan, Lie Feng, Wenyan Zhang, Xiaoqing Liu, Haifeng Yao, Lei Shen, Bin Li, Junqiang Zhang, Xuan Li, Wencheng Fang, Dan Wang, Marie-Emmanuelle Couprie, Guoqiang Lin, Bo Liu, Qiang Gu, Dong Wang, and Zhentang Zhao (2014), "Experimental demonstration of longitudinal beam phase-space linearizer in a free-electron laser facility by corrugated structures," Phys. Rev. Lett. **113**, 254802.

Derbenev, Y S, and A. N. Skrinsky (1978), "Magnetization effects in electron cooling," Fizika Plasmy **4**, 492–500.

Derbenev, Ya (1998), *Adapting Optics for High Energy Electron Cooling*, Tech. Rep. UM-HE-98-04 (University of Michigan).

Derbenev, Ya S, J. Rossbach, E.L. Saldin, and V.D. Shiltsev (1995), "Microbunch radiative tail - head interaction," TESLA-FEL Report No. 95-05, available from DESY, Hamburg, Germany.

Ding, Y, C. Behrens, R. Coffee, F.-J. Decker, P. Emma, C. Field, W. Helml, Z. Huang, P. Krejcik, J. Krzywinski, H. Loos, A. Lutman, A. Marinelli, T. J. Maxwell, and J. Turner (2015), "Generating femtosecond x-ray pulses using an emittance-spoiling foil in free-electron lasers," Applied Physics Letters **107** (19), 191104.

Dohlus, M, K Flöttmann, O.S Kozlov, T Limberg, Ph Piot, E.L Saldin, E.A Schneidmiller, and M.V Yurkov (2004), "Start-to-end simulations of SASE FEL at the TESLA Test Facility, Phase I: comparison with experimental results," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **528** (1), 448–452, in Proceedings of the 25th International Free Electron Laser Conference, and the 10th FEL Users Workshop.

Dohlus, M, E. A. Schneidmiller, and M. V. Yurkov (2011), "Generation of attosecond soft x-ray pulses in a longitudinal space charge amplifier," Phys. Rev. ST Accel. Beams **14**, 090702.

Dowell, D H, J. F Schmerge, and S. Lidia (2008), "High brightness electron injectors for light sources," in *U.S. Particle Accelerator School, Course Materials - University of California, Santa Cruz - January 2008*.

Dowell, David H (2010), "High brightness electron injectors for fel light sources course," .

Dowell, David H (2016), "Sources of emittance in rf photocathode injectors: Intrinsic emittance, space charge forces due to non-uniformities, rf and solenoid effects," arXiv:1610.01242 [physics.acc-ph].

Dowell, David H, and John F. Schmerge (2009), "Quantum efficiency and thermal emittance of metal photocathodes," Phys. Rev. ST Accel. Beams **12**, 074201.

Dragt, Alex J (2011), *Lie methods for nonlinear dynamics with applications to accelerator physics* (Maryland U., College Park) e-print only.

Du, Ying-Chao, Wen-Hui Huang, and Chuan-Xiang Tang (2012), "A new method to generate relativistic comb bunches with tunable subpicosecond spacing," Chinese Physics C **36** (2), 151–155.

Duan, Zhaoyun, Xianfeng Tang, Zhanliang Wang, Yabin Zhang, Xiaodong Chen, Min Chen, and Yubin Gong (2017), "Observation of the reversed Cherenkov radiation," Nature Communications **8**, 14901.

Duffy, Leanne Delma, and Alex J. Dragt (2016), "Utilizing the eigen-emittance concept for bright electron beams," in *Advances in Imaging and Electron Physics*, Vol. 193 (Elsevier) pp. 1–44.

Duncan, CJR, D.A. Muller, and J.M. Maxson (2020), "Lossless monochromation for electron microscopy with pulsed photoemission sources and radio-frequency cavities," Phys. Rev. Applied **14**, 014060.

Dupen, D W, H. A. Hogg, and G. A. Loew (1968), in *The Stanford two mile accelerator*, edited by R. B. Neal.

Edwards, D, H. Edwards, M. Ferrario, N. Holtkamp, S. Nagaitsev, J. Santucci, R. Brinkmann, K. Desler, K. Flöttman, and I. Bohnet (2000), "The flat beam experiment at the FNAL photoinjector," in *Proc. LINAC 2000*, pp. 122–124.

Edwards, D, H. Edwards, M. Ferrario, N. Holtkamp, S. Nagaitsev, J. Santucci, R. Brinkmann, K. Desler, K. Flöttman, and I. Bohnet (2001), "Status of flat electron beam production," in *Proc. PAC 2001*, pp. 73–75.

Egerton, R F (2008), "Electron energy-loss spectroscopy in the TEM," Reports on Progress in Physics **72** (1), 016502.

Emma, C, J. Wu, K. Fang, S. Chen, S. Serkez, and C. Pellegrini (2014a), "Terawatt x-ray free-electron-laser optimization by transverse electron distribution shaping," Phys. Rev. ST Accel. Beams **17**, 110701.

Emma, P, K. Bane, M. Cornacchia, Z. Huang, H. Schlarb, G. Stupakov, and D. Walz (2004), "Femtosecond and subfemtosecond x-ray pulses from a self-amplified spontaneous-emission-based free-electron laser," Phys. Rev. Lett. **92**, 074801.

Emma, P, Z. Huang, K.-J. Kim, and P. Piot (2006), "Transverse-to-longitudinal emittance exchange to improve performance of high-gain free-electron lasers," Phys. Rev. ST Accel. Beams **9**, 100702.

Emma, P, M. Venturini, K. L. F. Bane, G. Stupakov, H.-S. Kang, M. S. Chae, J. Hong, C.-K. Min, H. Yang, T. Ha, W. W. Lee, C. D. Park, S. J. Park, and I. S. Ko (2014b), "Experimental demonstration of energy-chirp control in relativistic electron bunches using a corrugated pipe," Phys. Rev. Lett. **112**, 034801.

England, R J, J. B. Rosenzweig, G. Andonian, P. Musumeci, G. Travish, and R. Yoder (2005), "Sextupole correction of the longitudinal transport of relativistic beams in dispersionless translating sections," Phys. Rev. ST Accel. Beams **8**, 012801.

England, R J, J. B. Rosenzweig, and G. Travish (2008), "Generation and measurement of relativistic electron bunches characterized by a linearly ramped current profile," Phys. Rev. Lett. **100**, 214802.

Faltens, A, and J.J Laslett (1975), "An estimate of limits to the longitudinal coupling impedance," in *Proceedings of the 1975 Isabelle Summer Study: July 14-25, 1975, Brookhaven (In 2 volumes)*, edited by J. C. Herrera and A. M. Thorndike, p. 486.

Ferrario, M, D. Alesini, A. Bacci, M. Bellaveglia, R. Boni, M. Boscolo, M. Castellano, L. Catani, E. Chiadroni, S. Cialdi, A. Cianchi, A. Clozza, L. Cultrera, G. Di Pirro, A. Drago, A. Esposito, L. Ficcadenti, D. Filippetto, V. Fusco, A. Gallo, G. Gatti, A. Ghigo, L. Giannessi, C. Ligi, M. Mattioli, M. Migliorati, A. Mostacci, P. Musumeci, E. Pace, L. Palumbo, L. Pellegrino, M. Petrarca, M. Quattromini, R. Ricci, C. Ronsivalle, J. Rosenzweig, A. R. Rossi, C. Sanelli, L. Serafini, M. Serio, F. Sgamma, B. Spataro, F. Tazzioli, S. Tomassini, C. Vaccarezza, M. Vescovi, and C. Vicario (2007), "Direct measurement of the double emittance minimum in the beam dynamics of the sparc high-brightness photoinjector," Phys.

Rev. Lett. **99**, 234801.

Ferrario, M, J. E. Clendenin, D. T. Palmer, J. B. Rosenzweig, and L. Serafini (2000), "HOMDYN study for the LCLS RF photoinjector," in *2nd ICFA Advanced Accelerator Workshop on the Physics of High Brightness Beams*, pp. 534–563.

Fetterman, A, D. Mihalcea, S. Benson, D. Crawford, D. Edstrom, F. Hannon, P. Piot, J. Ruan, and S. Wang (2022), "Photoinjector generation of high-charge magnetized beams for electron-cooling applications," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **1025**, 166051.

Flöttmann, K (1997), "Note on the thermal emittance of electrons emitted by cesium telluride photo cathodes," TESLA FEL-Report 1997-01, available from DESY, Hamburg, Germany.

Flöttmann, K (2014), "Generation of sub-fs electron beams at few-mev energies," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **740**, 34–38, proceedings of the first European Advanced Accelerator Concepts Workshop 2013.

Flöttmann, K, T. Limberg, and P. Piot (2001), "Generation of ultrashort electron bunches by cancellation of nonlinear distortions in the longitudinal phase space," TESLA-FEL Report No. 2001-06, available from DESY, Hamburg, Germany.

Flöttmann, Klaus (2015), "rf-induced beam dynamics in rf guns and accelerating cavities," Phys. Rev. ST Accel. Beams **18**, 064801.

Flöttmann, Klaus (2017), "Emittance compensation in split photoinjectors," Phys. Rev. Accel. Beams **20**, 013401.

Flöttmann, Klaus, and Valentin V. Paramonov (2014), "Beam dynamics in transverse deflecting rf structures," Phys. Rev. ST Accel. Beams **17**, 024001.

Forbes, Richard G (2015), "Field electron emission theory," arXiv:1801.08251 [physics.acc-ph].

Fowler, R H, and L. Nordheim (1928), "Electron emission in intense electric fields," Proc. R. Soc. Lond. A **119** (781), 173–181.

Fraser, J S, R. L. Sheffield, E. R. Gray, and G. W. Rodenz (1985), "High-brightness photoemitter injector for electron accelerators," IEEE Transactions on Nuclear Science **32** (5), 1791–1793.

Fu, Feichao, Rui Wang, Pengfei Zhu, Lingrong Zhao, Tao Jiang, Chao Lu, Shengguang Liu, Libin Shi, Lixin Yan, Haixiao Deng, Chao Feng, Qiang Gu, Dazhang Huang, Bo Liu, Dong Wang, Xingtao Wang, Meng Zhang, Zhentang Zhao, Gennady Stupakov, Dao Xiang, and Jie Zhang (2015), "Demonstration of nonlinear-energy-spread compensation in relativistic electron bunches with corrugated structures," Phys. Rev. Lett. **114**, 114801.

Gai, W, A. D. Kanareykin, A. L. Kustov, and J. Simpson (1997), "Numerical simulations of intense charged-particle beam propagation in a dielectric wake-field accelerator," Phys. Rev. E **55**, 3481–3488.

Gao, Q, G. Ha, C. Jing, S. P. Antipov, J. G. Power, M. Conde, W. Gai, H. Chen, J. Shi, E. E. Wisniewski, D. S. Doran, W. Liu, C. E. Whiteford, A. Zholents, P. Piot, and S. S. Baturin (2018), "Observation of high transformer ratio of shaped bunch generated by an emittance-exchange beam line," Phys. Rev. Lett. **120**, 114801.

Garault, Y (1964), "Etude d'une classe d'ondes electromagnetiques guidees: les ondes EH, applications aux structures deflectrices pour les separateurs a onde progressive de particules relativistes," 10.5170/CERN-1964-043.

Ginzton, E L, W. W. Hansen, and W. R. Kennedy (1948), "A linear electron accelerator," Review of Scientific Instruments **19** (2), 89–108, https://doi.org/10.1063/1.1741225.

Gluckstern, R L (1989), "Longitudinal impedance of a periodic structure at high frequency," Phys. Rev. D **39**, 2780–2783.

Golay, Marcel J E (1947), "Theoretical consideration in heat and infra-red detection, with particular reference to the pneumatic detector," Review of Scientific Instruments **18** (5), 347–356.

Goldstein, Herbert (2002), *Classical Mechanics (3rd edition)* (Addison-Wesley).

Gover, A, R. Ianconescu, A. Friedman, C. Emma, N. Sudar, P. Musumeci, and C. Pellegrini (2019), "Superradiant and stimulated-superradiant emission of bunched electron beams," Rev. Mod. Phys. **91**, 035003.

Graves, W S, S. L. Y. Chang, D. Dwyer, P. Fromme, M. Holl, B. D. A. Levin, L. E. Malin, J. L. Vincent, J. C. H. Spence, E. A. Nanni, R. K. Li, X. Shen, S. Weathersby, and A. Sandhu (2019), "Nanopatterned electron beams for temporal coherence and deterministic phase control of x-ray free-electron lasers," arXiv:1906.01525 [physics.acc-ph].

Graves, W S, F. X. Kärtner, D. E. Moncton, and P. Piot (2012), "Intense superradiant x rays from a compact source using a nanocathode array and emittance exchange," Phys. Rev. Lett. **108**, 263904.

Graves, William, *et al.* (2018), "ASU Compact XFEL," in *38th International Free-Electron Laser Conference*, p. TUB03.

Groening, L, C. Xiao, and M. Chung (2018), "Extension of Busch's theorem to particle beams," Phys. Rev. Accel. Beams **21**, 014201.

Ha, G, M. H. Cho, W. Gai, K.-J. Kim, W. Namkung, and J. G. Power (2016), "Perturbation-minimized triangular bunch for high-transformer ratio using a double dogleg emittance exchange beam line," Phys. Rev. Accel. Beams **19**, 121301.

Ha, G, M. H. Cho, W. Namkung, J. G. Power, D. S. Doran, E. E. Wisniewski, M. Conde, W. Gai, W. Liu, C. Whiteford, Q. Gao, K.-J. Kim, A. Zholents, Y.-E Sun, C. Jing, and P. Piot (2017a), "Precision control of the electron longitudinal bunch shape using an emittance-exchange beam line," Phys. Rev. Lett. **118**, 104801.

Ha, G, M. E. Conde, and J. G. Power (2019a), "Arbitrary transverse profile shaping using transverse wigglers," in *Proc. NAPAC 2019*, pp. 403–406.

Ha, G, M.E. Conde, D.S. Doran, W. Gai, and J.G. Power (2018a), "Preparations for Installation of the Double Emittance-Exchange Beamline at the Argonne Wakefield Accelerator Facility," in *Proc. of International Free Electron Laser Conference (FEL'17), Santa Fe, NM, USA, August 20-25, 2017*, International Free Electron Laser Conference No. 38 (JACoW, Geneva, Switzerland) pp. 340–343.

Ha, G, M.E. Conde, J.G. Power, and E.E. Wisniewski (2018b), "CSR Shielding Effect in Dogleg and EEX Beamlines," in *Proc. 9th International Particle Accelerator Conference (IPAC'18), Vancouver, BC, Canada, April 29-May 4, 2018*, International Particle Accelerator Conference No. 9 (JACoW Publishing, Geneva, Switzerland) pp. 1498–1500.

Ha, G, J G Power, M Conde, D S Doran, and W Gai (2017b), "Limiting effects in double EEX beamline," **874**, 012061.

Ha, G, J. G. Power, M. Conde, D. S. Doran, and W. Gai (2017c), "Preliminary simulations on chirpless bunch compression using double-EEX beamline," in *Proc. IPAC 2017*, pp. 3862–3864.

Ha, Gwanghui, Manoel Conde, John Power, Jiahang Shao, and Eric Wisniewski (2019b), "Tunable Bunch Train Generation Using Emittance Exchange Beamline With Transverse Wiggler," in *10th International Particle Accelerator Conference*, p. TUPGW089.

Ha, Gwanghui, John G. Power, Jiahang Shao, Manoel Conde, and Chunguang Jing (2020), "Coherent synchrotron radiation free longitudinal bunch shaping using transverse deflecting cavities," *Phys. Rev. Accel. Beams* **23**, 072803.

Hahn, H (1963), "Deflecting mode in circular iris-loaded waveguides," *Review of Scientific Instruments* **34** (10), 1094–1100, https://doi.org/10.1063/1.1718141.

Halavanau, A, G. Qiang, G. Ha, E. Wisniewski, P. Piot, J. G. Power, and W. Gai (2017), "Spatial control of photoemitted electron beams using a microlens-array transverse-shaping technique," *Phys. Rev. Accel. Beams* **20**, 103404.

Harms, Elvin, Helen Edwards, Markus Hüning, and Elmar Vogel (2011), "Commissioning and Early Operating Experience of the FLASH Third Harmonic RF System," in *25th International Linear Accelerator Conference*, p. TUP013.

Harris, J R, J. G. Neumann, K. Tian, and P. G. O'Shea (2007), "Longitudinal density modulation and energy conversion in intense beams," *Phys. Rev. E* **76**, 026402.

Heifets, S, G. Stupakov, and S. Krinsky (2002), "Coherent synchrotron radiation instability in a bunch compressor," *Phys. Rev. ST Accel. Beams* **5**, 064401.

Heifets, S A, and S. A. Kheifets (1991), "Coupling impedance in modern accelerators," *Rev. Mod. Phys.* **63**, 631–673.

Hemsing, Erik, Gennady Stupakov, Dao Xiang, and Alexander Zholents (2014), "Beam by design: Laser manipulation of electrons in modern accelerators," *Rev. Mod. Phys.* **86**, 897–941.

Huang, Z, A. Brachmann, F.-J. Decker, Y. Ding, D. Dowell, P. Emma, J. Frisch, S. Gilevich, G. Hays, Ph. Hering, R. Iverson, H. Loos, A. Miahnahri, H.-D. Nuhn, D. Ratner, G. Stupakov, J. Turner, J. Welch, W. White, J. Wu, and D. Xiang (2010), "Measurements of the linac coherent light source laser heater and its impact on the x-ray free-electron laser performance," *Phys. Rev. ST Accel. Beams* **13**, 020703.

Huang, Zhirong, and Kwang-Je Kim (2002), "Formulas for coherent synchrotron radiation microbunching in a bunch compressor chicane," *Phys. Rev. ST Accel. Beams* **5**, 074401.

Huang, Zhirong, Juhao Wu, and Timur Shaftan (2005), "Microbunching Instability due to Bunch Compression," *ICFA Beam Dyn. Newslett.* **38**, 37–51.

Hyun, J, P. Piot, T. Sen, and J.C. Thangaraj (2019), "Microbunching for generating tunable narrow-band THz radiation at the FAST photoinjector," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **916**, 195 – 203.

Iogansen, LV, and M.S. Rabinovich (1960), "Coherent electron radiation in the synchrotron. II," *JETP* **37** (1), 93 – 87.

Irvine, W, and D. Bouwmeester (2008), "Linked and knotted beams of light," *Nature Phys.* **4**, 716.

Jackson, John David (1998), *Classical Electrodynamics* (Wiley).

Jarvis, J D, H. L. Andrews, C. A. Brau, B. K. Choi, J. Davidson, W.-P. Kang, and Y.-M. Wong (2009), "Fabrication of self-aligned-gate diamond field-emitter-array triodes for free-electron lasers," in *Proceedings of FEL2009*, p. TUPC60.

Jarvis, J D, H. L. Andrews, B. Ivanov, C. L. Stewart, N. de Jonge, E. C. Heeres, W.-P. Kang, Y.-M. Wong, J. L. Davidson, and C. A. Brau (2010), "Resonant tunneling and extreme brightness from diamond field emitters and carbon nanotubes," *Journal of Applied Physics* **108**, 094322.

Jensen, Kevin L (2018), "A tutorial on electron sources," *IEEE Transactions on Plasma Science* **46** (6), 1881–1899.

Jiang, B, J. G. Power, R. Lindberg, W. Liu, and W. Gai (2011), "Emittance-exchange-based high harmonic generation scheme for a short-wavelength free electron laser," *Phys. Rev. Lett.* **106**, 114801.

Jiao, Yi, and Xiao-Hao Cui (2015), "Generating a uniform transverse distributed electron beam along a beam line," *Chinese Physics C* **39** (11), 117001.

Kallestrup, J, and M. Aiba (2020), "Emittance exchange in electron booster synchrotron by coupling resonance crossing," *Phys. Rev. Accel. Beams* **23**, 020701.

Kapchinskij, I M, and V. V. Vladimirskij (1959), "Limitations Of Proton Beam Current In A Strong Focusing Linear Accelerator Associated With The Beam Space Charge," in *2nd International Conference on High-Energy Accelerators*, pp. 274–287.

Kashy, E, and B. Sherrill (1987), "A method for the uniform charged particle irradiation of large targets," *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms* **26** (4), 610 – 613.

Kedia, Hridesh, Iwo Bialynicki-Birula, Daniel Peralta-Salas, and William T. M. Irvine (2013), "Tying knots in light fields," *Phys. Rev. Lett.* **111**, 150404.

Kellogg, Oliver Dimon (1967), *Foundations of Potential Theory* (Springer, Berlin, Heidelberg).

Khojoyan, M, M. Krasilnikov, and F. Stephan (2014), "Studies on the application of the 3d ellipsoidal cathode laser pulses at PITZ," in *Proc. of 5th International Particle Accelerator Conference*, p. THPRO043.

Kim, K-J (2020), *Simple Formula to Determine the Parameters of X-Y Emittance Exchange*, Tech. Rep. Light Source Note No. ANL/APS/LS-365. (Argonne National Laboratory).

Kim, K-J, Z. Huang, and R. Lindberg (2017), *Synchrotron Radiation and Free-Electron Lasers: Principles of Coherent X-Ray Generation* (Cambridge University Press).

Kim, K-J, and R. R. Lindberg (2011), "Collective and individual aspects fluctuations in relativistic electron beams for free electron lasers," in *Proceedings of the 2011 Interfnational Free-Electron Ler Conference (FEL11)*, pp. 156–158.

Kim, Kwang-Je (1989), "Rf and space-charge effects in laser-driven rf electron guns," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* **275** (2), 201 – 218.

Kim, Kwang-Je (2003), "Round-to-flat transformation of angular-momentum-dominated beams," *Phys. Rev. ST Accel. Beams* **6**, 104002.

Kim, Kwang-Je, and Vinit Kumar (2007), "Electron beam requirements for a three-dimensional Smith-Purcell

backward-wave oscillator for intense terahertz radiation," Phys. Rev. ST Accel. Beams **10**, 080702.

Kim, Kwang-Je, Yuri Shvyd'ko, and Sven Reiche (2008), "A proposal for an x-ray free-electron laser oscillator with an energy-recovery linac," Phys. Rev. Lett. **100**, 244802.

Kim, Kwang-Je, and Chun-xi Wang (2000), "Formulas for transverse ionization cooling in solenoidal focusing channels," Phys. Rev. Lett. **85**, 760–763.

Kim, Kwang-Je, and Andrew Sessler (2006), "Transverse-longitudinal phase-space manipulations and correlations," AIP Conference Proceedings **821** (1), 115–138.

Klimontovich, YuL (1995), *Statistical Theory of Open Systems* (Springer Netherlands).

Komljenovic, T, M. Davenport, J. Hulme, A. Y. Liu, C. T. Santis, A. Spott, S. Srinivasan, E. J. Stanton, C. Zhang, and J. E. Bowers (2016), "Heterogeneous silicon photonic integrated circuits," Journal of Lightwave Technology **34** (1), 20–35.

Kowalski, S, and H. Enge (1972), "Beam Rotator," in *Proceedings, 4th International Conference on Magnet Technology (MT-4)*, Vol. C720919, edited by Yvonne Winterbottom, p. 182.

Krafft, G A (1996), "Correcting M56 and T566 to obtain very short bunches at CEBAF," AIP Conference Proceedings **367** (1), 46–55.

Kur, E, G. Penn, J. Qiang, M. Venturini, R. Wells, and Zholents A. (2009), "Accelerator design study for a soft x-ray free electron laser at the Lawrence Berkeley National Laboratory," Internal Report **LBNL-2670**, 40–41.

Kuske, Peter, and Felix Kramer (2016), "Transverse Emittance Exchange for Improved Injection Efficiency," in *7th International Particle Accelerator Conference*, p. WEOAA01.

Kuzmin, I, S. Mironov, E. Gacheva, V. Zelenogorsky, A. Potemkin, E. Khazanov, A. Kanareykin, S. Antipov, M. Krasilnikov, G. Loisch, and F. Stephan (2018), "Shaping triangular picosecond laser pulses for electron photoinjectors," Laser Physics Letters **16** (1), 015001.

Kuzmin, I V, S. Yu. Mironov, E. I. Gacheva, A. K. Potemkin, E. A. Khazanov, M. A. Krasilnikov, and F. Stephan (2020), "Shaping picosecond ellipsoidal laser pulses with periodic intensity modulation for electron photoinjectors," Appl. Opt. **59** (9), 2776–2783.

Kuzmin, I V, S. Yu. Mironov, E. I. Gacheva, A. K. Poteomkin, and E. A. Khazanov (2019), "Retaining 3d shape of picosecond laser pulses during optical harmonics generation," Appl. Opt. **58** (10), 2678–2686.

Landau, L D, and E. M. Lifshitz (1969), *Mechanics. Vol. 1 (2nd ed.)* (Pergamon Press).

Laszczyk, KU (2020), "Field emission cathodes to form an electron beam prepared from carbon nanotube suspensions," Micromachines **11**, 260.

Lemery, F, and P. Piot (2015), "Tailored electron bunches with smooth current profiles for enhanced transformer ratios in beam-driven acceleration," Phys. Rev. ST Accel. Beams **18**, 081301.

Lemery, F, P. Piot, G. Amatuni, P. Boonpornprasert, Y. Chen, J. Good, B. Grigoryan, M. Groß, M. Krasilinikov, O. Lishilin, G. Loisch, A. Oppelt, S. Philipp, H. Qian, Y. Renier, F. Stephan, and I. Zagorodnov (2019), "Passive ballistic microbunching of nonultrarelativistic electron bunches using electromagnetic wakefields in dielectric-lined waveguides," Phys. Rev. Lett. **122**, 044801.

Lemons, R, J. C. Frisch, A. Fry, J. Robinson, S. Smith, and S. Carbajo (2021), "Integrated structured light architectures," Scientific Reports **11**, 796.

Lewellen, JW, S. Biedron, A. Lumpkin, S.V. Milton, A. Nassiri, S. Pasky, G. Travish, and M. White (1998), "Operation of the aps rf gun," in *1998 Linear Accelerator Accelerator Conference (LINAC98)*, TU4042, pp. 863–865.

Li, R (2001), "Sensitivity of the CSR self-interaction to the local longitudinal charge concentration of an electron bunch," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **475** (1), 498–503, FEL2000: Proc. 22nd Int. Free Electron Laser Conference and 7th FEL Users Workshop.

Li, S, S. Alverson, D. Bohler, A. Egger, A. Fry, S. Gilevich, Z. Huang, A. Miahnahri, D. Ratner, J. Robinson, and F. Zhou (2017), "Ultraviolet laser transverse profile shaping for improving x-ray free electron laser performance," Phys. Rev. Accel. Beams **20**, 080704.

Li, Yuelin, and Sergey Chemerisov (2008), "Manipulation of spatiotemporal photon distribution via chromatic aberration," Opt. Lett. **33** (17), 1996–1998.

Li, Yuelin, and Kwang-Je Kim (2008), "Nonrelativistic electron bunch train for coherently enhanced terahertz radiation sources," Applied Physics Letters **92** (1), 014101.

Li, Yuelin, and John W. Lewellen (2008), "Generating a quasiellipsoidal electron beam by 3D laser-pulse shaping," Phys. Rev. Lett. **100**, 074801.

LightSources, (2021), "Lightsources," .

Liouville, J (1838), "Note sur la théorie de la variation des constantes arbitraires," Journal de Mathématiques Pures et Appliquées **3**, 342–349.

Litvinenko, V N, G. Wang, D. Kayran, Y. Jing, J. Ma, and I. Pinayev (2018), "Plasma-cascade micro-bunching amplifier and coherent electron cooling of a hadron beams," arXiv:1802.08677 [physics.acc-ph].

Litvinenko, Vladimir N, Yichao Jing, Dmitry Kayran, Patrick Inacker, Jun Ma, Toby Miller, Irina Petrushina, Igor Pinayev, Kai Shih, Gang Wang, and Yuan H. Wu (2021), "Plasma-cascade instability," Phys. Rev. Accel. Beams **24**, 014402.

Litvinenko, Vladimir N, Gang Wang, Yichao Jing, Dmitry Kayran, Jun Ma, Irina Petrushina, Igor Pinayev, and Kai Shih (2019), "Plasma-Cascade Instability – theory, simulations and experiment," arXiv:1902.10846 [physics.acc-ph].

Liu, Fangming, Senlin Huang, Shangyu Si, Gang Zhao, Kexin Liu, and Shukui Zhang (2019), "Generation of picosecond pulses with variable temporal profiles and linear polarization by coherent pulse stacking in a birefringent crystal shaper," Opt. Express **27** (2), 1467–1478.

Loisch, G, J. Good, M. Gross, H. Huck, I. Isaev, M. Krasilnikov, O. Lishilin, A. Oppelt, Y. Renier, F. Stephan, R. Brinkmann, F. Grüner, and I. Will (2018), "Photocathode laser based bunch shaping for high transformer ratio plasma wakefield acceleration," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **909**, 107–110, 3rd European Advanced Accelerator Concepts Workshop (EAAC2017).

Lu, Chao, Tao Jiang, Shengguang Liu, Rui Wang, Lingrong Zhao, Pengfei Zhu, Dao Xiang, and Jie Zhang (2018), "Coulomb-driven relativistic electron beam compression," Phys. Rev. Lett. **120**, 044801.

Lu, Xueying, Michael A. Shapiro, Ivan Mastovsky, Richard J. Temkin, Manoel Conde, John G. Power, Jiahang Shao, Eric E. Wisniewski, and Chunguang Jing (2019), "Generation of high-power, reversed-cherenkov wakefield radiation in a metamaterial structure," Phys. Rev. Lett. **122**, 014801.

Luiten, O J, S. B. van der Geer, M. J. de Loos, F. B. Kiewiet, and M. J. van der Wiel (2004), "How to realize uniform three-dimensional ellipsoidal electron bunches," Phys. Rev. Lett. **93**, 094802.

Lumpkin, A H, R. J. Dejus, and N. S. Sereno (2009), "Coherent optical transition radiation and self-amplified spontaneous emission generated by chicane-compressed electron beams," Phys. Rev. ST Accel. Beams **12**, 040704.

Mach-Batlle, Rosa, Mark G. Bason, Nuria Del-Valle, and Jordi Prat-Camps (2020), "Tailoring magnetic fields in inaccessible regions," Phys. Rev. Lett. **125**, 177204.

MacKay, William W, and Mario Conte (2012), *Accelerator Physics Example Problems with Solutions* (World Scientific Publishing Co. Pte. Ltd., Singapore).

Majernik, N, G. Andonian, R. Roussel, S. Doran, G. Ha, J. Power, E. Wisniewski, and J. Rosenzweig (2021), "Multileaf collimator for real-time beam shaping using emittance exchange," arXiv:2107.00125 [physics.acc-ph].

Malyzhenkov, Alexander, and Alexander Scheinker (2018), "Phase space exchange-based bunch compression with reduced CSR effects," arXiv:1809.05579 [physics.acc-ph].

Mamun, MA, *et al.* (2018), "Production of Magnetized Electron Beam from a DC High Voltage Photogun," in *Proc. 9th International Particle Accelerator Conference (IPAC'18), Vancouver, BC, Canada, April 29-May 4, 2018*, International Particle Accelerator Conference No. 9 (JACoW Publishing, Geneva, Switzerland) pp. 4567–4570, https://doi.org/10.18429/JACoW-IPAC2018-THPMK108.

Marinelli, A, E. Hemsing, M. Dunning, D. Xiang, S. Weathersby, F. O'Shea, I. Gadjev, C. Hast, and J. B. Rosenzweig (2013), "Generation of coherent broadband photon pulses with a cascaded longitudinal space-charge amplifier," Phys. Rev. Lett. **110**, 264802.

Marinelli, Agostino, Erik Hemsing, and James B. Rosenzweig (2011), "Three dimensional analysis of longitudinal plasma oscillations in a thermal relativistic electron beam," Physics of Plasmas **18** (10), 103105.

Matsui, Futoshi, Shin'ichi Goriki, Yukio Shimizu, Hiromitsu Tomizawa, Sakae Kawato, and Takao Kobayashi (2008), "Genetic-algorithm-based method to optimize spatial profile utilizing characteristics of electrostatic actuator deformable mirror," Opt. Rev. **15** (3), 156–161.

Maxson, Jared, Hyeri Lee, Adam C. Bartnik, Jacob Kiefer, and Ivan Bazarov (2015), "Adaptive electron beam shaping using a photoemission gun and spatial light modulator," Phys. Rev. ST Accel. Beams **18**, 023401.

Mayet, F, R. Assmann, and F. Lemery (2020), "Longitudinal phase space synthesis with tailored 3D-printable dielectric-lined waveguides," Phys. Rev. Accel. Beams **23**, 121302.

McCulloch, A J, D. V. Sheludko, S. D. Saliba, S. C. Bell, M. Junker, K. A. Nugent, and R. E. Scholten (2011), "Arbitrarily shaped high-coherence electron bunches from cold atoms," Nature Physics **7**, 785.

McDonald, KT (1988), "Design of the laser-driven rf electron gun for the bnl accelerator test facility," IEEE Transactions on Electron Devices **35** (11), 2052–2059.

Meads, PF, Jr (1983), "A Nonlinear Lens System to Smooth the Intensity Distribution of a Gaussian Beam," IEEE Trans. Nucl. Sci. **30**, 2838.

Meot, F, and T Aniel (1996), "Principles of the nonlinear tuning of beam expanders," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **379** (2), 196 – 205.

Merminga, N, John Irwin, Richard H. Helm, and R.D. Ruth (1991), "Optimizing a nonlinear collimation system for future linear colliders," in *Conference Record of the 1991 IEEE Particle Accelerator Conference*, Vol. 910506, edited by Loretta Lizama and Joe Chew, pp. 219–221.

Miginsky, SV (2009), "Emittance compensation of elliptical beams," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **603** (1), 32–34, proceedings of the XVII International Synchrotron Radiation Conference.

Mihalcea, D, L. Faillace, J. Hartzell, H. Panuganti, S. Boucher, A. Murokh, P. Piot, and J. C. T. Thangaraj (2015), "Measurement of Ampère-class pulsed electron beams via field emission from carbon-nanotube cathodes in a radiofrequency gun," Applied Physics Letters **107** (3), 033502.

Mironov, S Yu, A. K. Potemkin, E. I. Gacheva, A. V. Andrianov, V. V. Zelenogorskii, M. Krasilnikov, F. Stephan, and E. A. Khazanov (2016), "Shaping of cylindrical and 3D ellipsoidal beams for electron photoinjector laser drivers," Appl. Opt. **55** (7), 1630–1635.

Mitchell, Chad, Ji Qiang, and Paul Emma (2013), "Longitudinal pulse shaping for the suppression of coherent synchrotron radiation-induced emittance growth," Phys. Rev. ST Accel. Beams **16**, 060703.

Muggli, P, B. Allen, V. Yakimenko, M. Fedurin, K. Kusche, and M. Babzien (2010a), "Resonant excitation of plasma wakefields," AIP Conference Proceedings **1299** (1), 495–499.

Muggli, P, B. Allen, V. E. Yakimenko, J. Park, M. Babzien, K. P. Kusche, and W. D. Kimura (2010b), "Simple method for generating adjustable trains of picosecond electron bunches," Phys. Rev. ST Accel. Beams **13**, 052803.

Muggli, P, V. Yakimenko, M. Babzien, E. Kallos, and K. P. Kusche (2008), "Generation of trains of electron microbunches with adjustable subpicosecond spacing," Phys. Rev. Lett. **101**, 054801.

Murphy, JB, S. Krinsky, and R. L. Gluckstern (1997), "Longitudinal wake field for an electron moving on a circular orbit," Part. Accel. **57**, 9–64.

Mustonen, A, P. Beaud, E. Kirk, T. Feurer, and S. Tsujino (2011), "Five picocoulomb electron bunch generation by ultrafast laser-induced field emission from metallic nanotip arrays," Appl. Phys. Lett. **99**, 103504.

Musumeci, P, R. K. Li, and A. Marinelli (2011), "Nonlinear longitudinal space charge oscillations in relativistic electron beams," Phys. Rev. Lett. **106**, 184801.

Musumeci, P, R. K. Li, K. G. Roberts, and E. Chiadroni (2013), "Controlling nonlinear longitudinal space charge oscillations for high peak current bunch train generation," Phys. Rev. ST Accel. Beams **16**, 100701.

Musumeci, P, J. T. Moody, R. J. England, J. B. Rosenzweig, and T. Tran (2008), "Experimental generation and characterization of uniformly filled ellipsoidal electron-beam distributions," Phys. Rev. Lett. **100**, 244801.

Nagaitsev, S, Z. Huang, J. Power, J. L. Vay, P. Piot, L. Spentzouris, J. Rosenzweig, Y. Cai, S. Cousineau, M. Conde, M. Hogan, A. Valishev, M. Minty, T. Zolkin, X. Huang, V. Shiltsev, J. Seeman, J. Byrd, Y. Hao, B. Dunham,

B. Carlsten, A. Seryi, and R. Patterson (2021), "Accelerator and beam physics research goals and opportunities," arXiv:2101.04107 [physics.acc-ph].

Nagayama, Kuniaki (2011), "Another 60 years in electron microscopy: development of phase-plate electron microscopy and biological applications," Journal of Electron Microscopy **60** (suppl 1), S43–S62.

Nanni, E A, and W. S. Graves (2015), "Aberration corrected emittance exchange," Phys. Rev. ST Accel. Beams **18**, 084401.

Nanni, E A, W. S. Graves, and D. E. Moncton (2018), "Nanomodulated electron beams via electron diffraction and emittance exchange for coherent x-ray generation," Phys. Rev. Accel. Beams **21**, 014401.

Neil, G R, C. L. Bohn, S. V. Benson, G. Biallas, D. Douglas, H. F. Dylla, R. Evans, J. Fugitt, A. Grippo, J. Gubeli, R. Hill, K. Jordan, R. Li, L. Merminga, P. Piot, J. Preble, M. Shinn, T. Siggins, R. Walker, and B. Yunn (2000), "Sustained kilowatt lasing in a free-electron laser with same-cell energy recovery," Phys. Rev. Lett. **84**, 662–665.

Neil, GR, C. Behre, S.V. Benson, M. Bevins, G. Biallas, J. Boyce, J. Coleman, L.A. Dillon-Townes, D. Douglas, H.F. Dylla, R. Evans, A. Grippo, D. Gruber, J. Gubeli, D. Hardy, C. Hernandez-Garcia, K. Jordan, M.J. Kelley, L. Merminga, J. Mammosser, W. Moore, N. Nishimori, E. Pozdeyev, J. Preble, R. Rimmer, M. Shinn, T. Siggins, C. Tennant, R. Walker, G.P. Williams, and S. Zhang (2006), "The jlab high power erl light source," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **557** (1), 9–15.

Neil, V Kelvin, and Andrew M. Sessler (1965), "Longitudinal resistive instabilities of intense coasting beams in particle accelerators," Review of Scientific Instruments **36** (4), 429–436.

Neri, Filippo, and Govindan Rangarajan (1990), "Kinematic moment invariants for linear Hamiltonian systems," Phys. Rev. Lett. **64**, 1073–1075.

Neumann, J G, R. B. Fiorito, P. G. O'Shea, H. Loos, B. Sheehy, Y. Shen, and Z. Wu (2009), "Terahertz laser modulation of electron beams," Journal of Applied Physics **105** (5), 053304.

Neumann, JG, P.G. O'Shea, D. Demske, W.S. Graves, B. Sheehy, H. Loos, and G.L. Carr (2003), "Electron beam modulation using a laser-driven photocathode," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **507** (1), 498 – 501, proceedings of the 24th International Free Electron Laser Conference and the 9th Users Workshop.

Nguyen, DC, and B.E Carlsten (1996), "Amplified coherent emission from electron beams prebunched in a masked chicane," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **375** (1), 597 – 601, proceedings of the 17th International Free Electron Laser Conference.

Nichols, K E, H. L. Andrews, D. Kim, E. I. Simakov, M. Conde, D. S. Doran, G. Ha, W. Liu, J. F. Power, J. Shao, C. Whiteford, E. E. Wisniewski, S. P. Antipov, and G. Chen (2020), "Demonstration of transport of a patterned electron beam produced by diamond pyramid cathode in an rf gun," Applied Physics Letters **116** (2), 023502.

Nielsen, C E, and A. M. Sessler (1959), "Longitudinal space charge effects in particle accelerators," Review of Scientific Instruments **30** (2), 80–89.

Nielsen, CE, A.M. Sessler, and K.R. Symon (1959), "Longitudinal instabilities in intense relativistic beams," in Proc. 2nd International Conference on High-Energy Accelerators, pp. 239–252.

Noether, Emmy (1971), "Invariant variation problems," Transport Theory and Statistical Physics **1** (3), 186–207.

Novokhatski, A (2015), "Wakefield potentials of corrugated structures," Phys. Rev. ST Accel. Beams **18**, 104402.

O'Shea, Brendan, James B. Rosenzweig, Galina Asova, Jürgen Bähr, Marc Hänel, Yevgeniy Ivanisenko, Martin Khojoyan, Mikhail Krasilnikov, Lazar Staykov, Frank Stephan, and Juliane Rönsch-Schulenburg (2011), "Measurement of self-shaped ellipsoidal bunches from a photoinjector with postacceleration," Phys. Rev. ST Accel. Beams **14**, 012801.

van Oudheusden, T, E. F. de Jong, S. B. van der Geer, W. P. E. M. Op 't Root, O. J. Luiten, and B. J. Siwick (2007), "Electron source concept for single-shot sub-100 fs electron diffraction in the 100 kev range," Journal of Applied Physics **102** (9), 093501, https://doi.org/10.1063/1.2801027.

Panofsky, W K H, and W. A. Wenzel (1956), "Some considerations concerning the transverse deflection of charged particles in radio-frequency fields," Review of Scientific Instruments **27** (11), 967–967, https://doi.org/10.1063/1.171542.

Paramonov, Valentin, and Klaus Flöttmann (2019), "Fundamental characteristics of transverse deflecting field," arXiv:1806.11023 [physics.acc-ph].

Parkhomchun, V, and A. Skrinskii (2000), "Electron cooling: 35 years of development," Physics-Uspekhi **43 (5)**, 433–452.

Pasquali, M, A. Bertarelli, C. Accettura, E. Berthome, L. Bianchi, P. Bolz, F. Carra, C. Fichera, M. I. Frankl, T. Furness, G. Gobbi, P. Grosclaude, J. Guardia-Valenzuela, M. D. Jedrychowsky, F. J. Harden, A. Lechner, P. Mollicone, P. D. Pastuszak, M. Portelli, S. Redaelli, E. Rigutto, O. Sacristan de Frutos, and P. Simon (2019), "Dynamic response of advanced materials impacted by particle beams: The multimat experiment," Journal of Dynamic Behavior of Materials **5**, 266–295.

Penco, G, E. Allaria, I. Cudin, S. Di Mitri, D. Gauthier, S. Spampinati, M. Trovó, L. Giannessi, E. Roussel, S. Bettoni, P. Craievich, and E. Ferrari (2017), "Passive linearization of the magnetic bunch compression using self-induced fields," Phys. Rev. Lett. **119**, 184802.

Penco, G, M. Danailov, A. Demidovich, E. Allaria, G. De Ninno, S. Di Mitri, W. M. Fawley, E. Ferrari, L. Giannessi, and M. Trovó (2014), "Experimental demonstration of electron longitudinal-phase-space linearization by shaping the photoinjector laser pulse," Phys. Rev. Lett. **112**, 044801.

Petrushina, I, Y. Jing, V. N. Litvinenko, J. Ma, I. Pinayev, K. Shih, G. Wang, and Y. H. Wu (2019), "First experimental observations of the plasma-cascade instability in the CeC PoP accelerator," in Proc. of North America Particle Accelerator Conference 2019, pp. 379–382.

Pines, David, and David Bohm (1952), "A collective description of electron interactions: II. collective vs individual particle aspects of the interactions," Phys. Rev. **85**, 338–353.

Piot, P (2009), "Alternative lattice options for energy recovery in high-average-power high-efficiency free-electron lasers," in 2nd Workshop on High Average Power and High Brightness Beams, Los Angeles, California, 14-16 Jan 2009

(unpublished).

Piot, P, C. Behrens, C. Gerth, M. Dohlus, F. Lemery, D. Mihalcea, P. Stoltz, and M. Vogt (2012), "Generation and characterization of electron bunches with ramped current profiles in a dual-frequency superconducting linear accelerator," Phys. Rev. Lett. **108**, 034801.

Piot, P, C. A. Brau, B. K. Choi, B. Blomberg, W. E. Gabella, B. Ivanov, J. Jarvis, M. H. Mendenhall, D. Mihalcea, H. Panuganti, P. Prieto, and J. Reid (2014), "Operation of an ungated diamond field-emission array cathode in a L-band radiofrequency electron source," Applied Physics Letters **104** (26), 263504.

Piot, P, D. R. Douglas, and G. A. Krafft (2003), "Longitudinal phase space manipulation in energy recovering linac-driven free-electron lasers," Phys. Rev. ST Accel. Beams **6**, 030702.

Piot, P, and Y. E. Sun (2014), "Generation and dynamics of magnetized electron beams for high-energy electron cooling," International Workshop on Accelerator Science & Technologies for future electron-ion colliders (EIC'14), Jefferson Lab, March 17-21, 2014 .

Piot, P, Y.-E Sun, and K.-J. Kim (2006), "Photoinjector generation of a flat electron beam with transverse emittance ratio of 100," Phys. Rev. ST Accel. Beams **9**, 031001.

Piot, P, Y.-E Sun, T. J. Maxwell, J. Ruan, A. H. Lumpkin, M. M. Rihaoui, and R. Thurman-Keup (2011a), "Observation of coherently enhanced tunable narrow-band terahertz transition radiation from a relativistic sub-picosecond electron bunch train," Applied Physics Letters **98** (26), 261501.

Piot, P, Y.-E Sun, T. J. Maxwell, J. Ruan, E. Secchi, and J. C. T. Thangaraj (2013), "Formation and acceleration of uniformly filled ellipsoidal electron bunches obtained via space-charge-driven expansion from a cesium-telluride photocathode," Phys. Rev. ST Accel. Beams **16**, 010102.

Piot, P, Y.-E Sun, J. G. Power, and M. Rihaoui (2011b), "Generation of relativistic electron bunches with arbitrary current distribution via transverse-to-longitudinal phase space exchange," Phys. Rev. ST Accel. Beams **14**, 022801.

Piot, P, Y.-E. Sun, A. S. Johnson, A. H. Lumpkin, T. Maxwell, J. Ruan, and R. Thurman-Keup (2010), "Transverse-to-longitudinal phase space exchange: a versatile tool for shaping the current and energy profiles of relativistic electron bunches," AIP Conference Proceedings **1299** (1), 592–597.

Power, John G, and Chunguang Jing (2009), "Temporal laser pulse shaping for RF photocathode guns: The cheap and easy way using UV birefringent crystals," AIP Conference Proceedings **1086** (1), 689–694.

Prat, Eduard, Eugenio Ferrari, Sven Reiche, and Thomas Schietinger (2017), "Using the optical-klystron effect to increase and measure the intrinsic beam energy spread in free-electron-laser facilities," Phys. Rev. Accel. Beams **20**, 040702.

Qin, W, *et al.* (2016), "Beam Optimization Study for an X-ray FEL Oscillator at the LCLS-II," in *Proc. of International Particle Accelerator Conference (IPAC'16), Busan, Korea, May 8-13, 2016*, International Particle Accelerator Conference No. 7 (JACoW, Geneva, Switzerland) pp. 3020–3023, doi:10.18429/JACoW-IPAC2016-WEPOY019.

Ranada, A F (1990), "Knotted solutions of the Maxwell equations in vacuum," Journal of Physics A: Mathematical and General **23** (16), L815–L820.

Rand, RE, and T.I. Smith (1980), "Beam optical control of beam breakup in a recirculating electron accelerator," Part. Accel. **11**, 1–13.

Rao, Triveni, and David H Dowell (2015), "An engineering guide to photoinjectors," arXiv:1403.7539 [physics.acc-ph].

Ratner, D (2013), "Microbunched electron cooling for high-energy hadron beams," Phys. Rev. Lett. **111**, 084802.

Ratner, D, C. Behrens, Y. Ding, Z. Huang, A. Marinelli, T. Maxwell, and F. Zhou (2015), "Time-resolved imaging of the microbunching instability and energy spread at the Linac Coherent Light Source," Phys. Rev. ST Accel. Beams **18**, 030704.

Reiser, M (1994), *Theory and Design of Charged Particle Beams* (John Wiley & Sons, inc.).

Richardson, O W (1913), "The emission of electron from tugsten at high temperatures: an experimental proof that the electron current in metals is carried by electrons," Science **38** (967), 57–61, publisher: American Association for the Advancement of Science _eprint: https://science.sciencemag.org/content/38/967/57.full.pdf.

Rihaoui, M, P. Piot, J. G. Power, Z. Yusof, and W. Gai (2009), "Observation and simulation of space-charge effects in a radio-frequency photoinjector using a transverse multibeamlet distribution," Phys. Rev. ST Accel. Beams **12**, 124201.

Rosenzweig, J, C Pellegrini, L Serafini, C Ternieden, and G Travish (1997), "Space-charge oscillations in a self-modulated electron beam in multi-undulator free-electron lasers," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **393** (1), 376 – 379, Free Electron Lasers 1996.

Rosenzweig, J B, A. Cahill, V. Dolgashev, C. Emma, A. Fukasawa, R. Li, C. Limborg, J. Maxson, P. Musumeci, A. Nause, R. Pakter, R. Pompili, R. Roussel, B. Spataro, and S. Tantawi (2019), "Next generation high brightness electron beams from ultrahigh field cryogenic rf photocathode sources," Phys. Rev. Accel. Beams **22**, 023403.

Rosenzweig, JB, A. M. Cook, R. J. England, M. Dunning, S .G. Anderson, and Massimo Ferrario (2006), "Emittance compensation with dynamically optimized photoelectron beam profiles," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **557** (1), 87 – 93, Energy Recovering Linacs 2005.

Roussel, R, G. Andonian, W. Lynn, K. Sanwalka, R. Robles, C. Hansel, A. Deng, G. Lawler, J. B. Rosenzweig, G. Ha, J. Seok, J. G. Power, M. Conde, E. Wisniewski, D. S. Doran, and C. E. Whiteford (2020), "Single shot characterization of high transformer ratio wakefields in nonlinear plasma acceleration," Phys. Rev. Lett. **124**, 044802.

Ruan, J, A. S. Johnson, A. H. Lumpkin, R. Thurman-Keup, H. Edwards, R. P. Fliller, T. W. Koeth, and Y.-E Sun (2011), "First observation of the exchange of transverse and longitudinal emittances," Phys. Rev. Lett. **106**, 244801.

Ruth, R D (1986), "Single Particle Dynamics and Nonlinear Resonances in Circular Accelerators," Lect. Notes Phys. **247**, 37.

Saldin, E L, E. A. Schneidmiller, and M. V. Yurkov (1997), "On the coherent radiation of an electron bunch moving in an arc of a circle," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **398** (2), 373 – 394.

Saldin, E L, E. A. Schneidmiller, and M. V. Yurkov (2002a), "An analytical description of longitudinal phase space distortions in magnetic bunch compressors," Nuclear Instru-

ments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **483**, 516 – 520.

Saldin, E L, E. A Schneidmiller, and M. V Yurkov (2004), "Longitudinal space charge-driven microbunching instability in the TESLA Test Facility linac," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **528** (1), 355 – 359.

Saldin, EL, E.A. Schneidmiller, and M.V. Yurkov (1996), "Calculation of energy diffusion in an electron beam due to quantum fluctuations of undulator radiation," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **381** (2), 545–547.

Saldin, EL, E.A. Schneidmiller, and M.V. Yurkov (2002b), "Klystron instability of a relativistic electron beam in a bunch compressor," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **490** (1), 1–8.

Sands, Matthew (1955), "Synchrotron oscillations induced by radiation fluctuations," Phys. Rev. **97**, 470–473.

Sands, Matthew (1969), "The Physics of Electron Storage Rings: An Introduction," Conf. Proc. C **6906161**, 257–411.

Schneidmiller, E A, and M. V. Yurkov (2010), "Using the longitudinal space charge instability for generation of vacuum ultraviolet and x-ray radiation," Phys. Rev. ST Accel. Beams **13**, 110701.

Schott, G A (1912), *Electromagnetic Radiation and the Mechanical Reactions Arising from It,* (Cambridge University Press, Cambridge).

Seeman, John T (1992), "Observation and cures of wakefield effects in the slc linac," AIP Conference Proceedings **255** (1), 339–346, https://aip.scitation.org/doi/pdf/10.1063/1.42287.

Seok, J, M. Chung, M. Conde, G. Ha, and J. Power (2019a), "Double-horn suppression in EEX based bunch compression," in *2019 North American Particle Accelerator Conference (NAPAC2019)*, pp. 407–410.

Seok, J, G. Ha, J. Power, and M. Chung (2021), "Longitudinal phase space manipulation using double emittance exchange to generate multi-color x-ray," arXiv:2104.07296 [physics.acc-ph].

Seok, J, G. Ha, J. Power, M. Conde, and M. Chung (2019b), "Suppression of correlated energy spread using emittance exchange," in *2019 International Particle Accelerator Conference (IPAC2019)*, pp. 3275–3278.

Seok, Jimin, Moses Chung, Heung-Sik Kang, Chang-Ki Min, and Donghyun Na (2018), "Use of a corrugated beam pipe as a passive deflector for bunch length measurements," Phys. Rev. Accel. Beams **21**, 022801.

Serafini, Luca, and James B. Rosenzweig (1997), "Envelope analysis of intense relativistic quasilaminar beams in rf photoinjectors:ma theory of emittance compensation," Phys. Rev. E **55**, 7565–7590.

Shchegolkov, Dmitry, Sergey Antipov, Mikhail Fedurin, and Evgenya Simakov (2015), "Dielectric Wakefield Accelerator Experiments at ATF," in *6th International Particle Accelerator Conference*, p. WEPJE006.

Shiloh, Roy, Peng-Han Lu, Roei Remez, Amir H. Tavabi, Giulio Pozzi, Rafal E. Dunin-Borkowski, and Ady Arie (2019), "Nanostructuring of electron beams," Physica Scripta **94** (3), 034004.

Shiltsev, Vladimir D (2016), *Electron Lenses for Super-Colliders*, Particle Acceleration and Detection (Springer).

Siders, Craig W, Jennifer L. W. Siders, Antoinette J. Taylor, Sang-Gyu Park, and Andrew M. Weiner (1998), "Efficient high-energy pulse-train generation using a 2n-pulse michelson interferometer," Appl. Opt. **37** (22), 5302–5305.

Simakov, Evgenya I, Sergey A. Arsenyev, Cynthia E. Buechler, Randall L. Edwards, William P. Romero, Manoel Conde, Gwanghui Ha, John G. Power, Eric E. Wisniewski, and Chunguang Jing (2016), "Observation of wakefield suppression in a photonic-band-gap accelerator structure," Phys. Rev. Lett. **116**, 064801.

Smedley, John, and Matt Poelker (2012), "Cathode physics," in *U.S. Particle Accelerator School, Course Materials - University of Texas at Austin - January 2012*.

Smirnova, E I (2004), "Photonic band gap structures for accelerator applications," AIP Conference Proceedings **737** (1), 309–319.

Smith, TI (1986), "Intense Low Emittance Linac Beams for Free Electron Lasers," in *13th International Linear Accelerator Conference*, pp. TH2–4.

Sørensen, A H (1988), "Liouville's theorem and emittance," in *CERN Accelerator School: General Accelerator Physics* (CERN, Geneva, Switzerland) pp. 18–36.

Spampinati, S, E. Allaria, L. Badano, S. Bassanese, S. Biedron, D. Castronovo, P. Craievich, M. B. Danailov, A. Demidovich, G. De Ninno, S. Di Mitri, B. Diviacco, M. Dal Forno, E. Ferrari, W. M. Fawley, L. Fröhlich, G. Gaio, L. Giannessi, G. Penco, C. Serpico, C. Spezzani, M. Trovò, M. Veronese, S. V. Milton, and M. Svandrlik (2014), "Laser heater commissioning at an externally seeded free-electron laser," Phys. Rev. ST Accel. Beams **17**, 120705.

Stupakov, G V (2001), "Wake and impedance," AIP Conference Proceedings **592** (1), 205–230.

Sun, Y (2005), "Angular-momentum-dominated electron beams and flat-beam generation," Ph.D. dissertation, Department of Physics, The Univerity of Chicago, 10.2172/15017103.

Sun, Y, P. Piot, A. Johnson, A. Lumpkin, J. Ruan, and R. Thurman-Keup (2010a), "Conversion of a transverse density modulation into a longitudinal phase space modulation using an emittance exchange technique," arXiv:1003.3126 [physics.acc-ph].

Sun, Y, P. Piot, A. Johnson, A. H. Lumpkin, T. J. Maxwell, J. Ruan, and R. Thurman-Keup (2010b), "Tunable subpicosecond electron-bunch-train generation using a transverse-to-longitudinal phase-space exchange technique," Phys. Rev. Lett. **105**, 234801.

Sun, Y, P. Piot, K.-J. Kim, N. Barov, S. Lidia, J. Santucci, R. Tikhoplav, and J. Wennerberg (2004), "Generation of angular-momentum-dominated electron beams from a photoinjector," Phys. Rev. ST Accel. Beams **7**, 123501.

Sun, Y, J. G. Power, K. J. Kim, P. Piot, and M. M. Rihaoui (2007), "Design study of a transverse-to-longitudinal emittance exchange proof-of-principle experiment," in *2007 IEEE Particle Accelerator Conference (PAC)*, pp. 3441–3443.

Sun, Y-E, J. W. Lewellen, and D. W. Feldman (2006), "Photothermal cathode measurements at the advanced photon source," in *2006 Linear Accelerator Accelerator Conference (LINAC06)*, TUP045, pp. 349–351.

Swanwick, Michael E, Phillip D. Keathley, Arya Fallahi, Peter R. Krogen, Guillaume Laurent, Jeffrey Moses, Franz X.

Kärtner, and Luis F. Velásquez-García (2014), "Nanostructured ultrafast silicon-tip optical field-emitter arrays," Nano Lett. **14** (9), 5035–5043.

Talman, Richard (1995), "A proposed Möbius accelerator," Phys. Rev. Lett. **74**, 1590–1593.

Tan, Wei Hou, Philippe Piot, and Alexander Zholents (2021), "Formation of temporally shaped electron bunches for beam-driven collinear wakefield accelerators," Phys. Rev. Accel. Beams **24**, 051303.

Tennant, Christopher D, Kevin B. Beard, David R. Douglas, Kevin C. Jordan, Lia Merminga, Eduard G. Pozdeyev, and Todd I. Smith (2005), "First observations and suppression of multipass, multibunch beam breakup in the Jefferson Laboratory free electron laser upgrade," Phys. Rev. ST Accel. Beams **8**, 074403.

Tiefenback, M G, and G. A. Krafft (1993), "Compensation of rf-induced energy spread in the CEBAF injector chopping system," in *Proceedings of International Conference on Particle Accelerators*, Vol. 1, pp. 426–428.

Tournois, Pierre (1997), "Acousto-optic programmable dispersive filter for adaptive compensation of group delay time dispersion in laser systems," Optics Communications **140** (4), 245–249.

Tsoupas, N, L. Ahrens, S. Bellavia, R. Bonati, K. A. Brown, I-Hung Chiang, C. J. Gardner, D. Gassner, S. Jao, W. W. Mackay, I. Marneris, W. Meng, D. Phillips, P. Pile, R. Prigl, A. Rusek, L. Snydstrup, and K. Zeno (2007), "Uniform beam distributions at the target of the NASA Space Radiation Laboratory's beam line," Phys. Rev. ST Accel. Beams **10**, 024701.

Tsoupas, N, R. Lankshear, C. L. Snead, T. E. Ward, M. Zucker, and H. A. Enge (1991), "Uniform beam distributions using octupoles," in *Conference Record of the 1991 IEEE Particle Accelerator Conference*, Vol. 3, pp. 1695–1697.

Tsuno, K, and E. Munro (1997), "Design of an omega filter for a 200 kV electron microscope," Review of Scientific Instruments **68** (1), 109–115.

Uchida, Masaya, and Akira Tonomura (2010), "Generation of electron beams carrying orbital angular momentum," Nature **464**, 737.

van der Geer, S B, E. J. D. Vredenbregt, O. J. Luiten, and M. J. de Loos (2014), "An ultracold electron source as an injector for a compact SASE-FEL," Journal of Physics B: Atomic, Molecular and Optical Physics **47** (23), 234009.

Venturini, Marco (2008), "Models of longitudinal space-charge impedance for microbunching instability," Phys. Rev. ST Accel. Beams **11**, 034401.

Verluise, Frédéric, Vincent Laude, Jean-Pierre Huignard, Pierre Tournois, and Arnold Migus (2000), "Arbitrary dispersion control of ultrashort optical pulses with acoustic waves," J. Opt. Soc. Am. B **17** (1), 138.

Wang, Chun-xi (2006), "Hamiltonian analysis of transverse beam dynamics in high-brightness photoinjectors," Phys. Rev. E **74**, 046502.

Wang, D X, G. A. Krafft, and C. K. Sinclair (1998), "Measurement of femtosecond electron bunches using a rf zero-phasing method," Phys. Rev. E **57**, 2283–2286.

Weiner, A M, J. P. Heritage, and E. M. Kirschner (1988), "High-resolution femtosecond pulse shaping," J. Opt. Soc. Am. B **5** (8), 1563–1572.

Werner, Douglas H, and Do-Hoon Kwon, Eds. (2014), *Transformation Electromagnetics and Metamaterials: Fundamental Principles and Applications* (Springer-Verlag London).

Wideröe, R (1928), "Über ein neues prinzip zur herstellung hoher spannungen," Archiv f. Elektrotechnik **21**, 387–406.

Wiedemann, H (1999), *Particle Accelerator Physics I* (Springer).

Wigner, E (1934), "On the interaction of electrons in metals," Phys. Rev. **46**, 1002–1011.

Will, Ingo, and Guido Klemz (2008), "Generation of flat-top picosecond pulses by coherent pulse stacking in a multicrystal birefringent filter," Opt. Express **16** (19), 14922.

Williamson, John (1936), "On the algebraic problem concerning the normal forms of linear dynamical systems," American Journal of Mathematics **58** (1), 141.

Wilson, M A, R. I. Cutler, D. L. Mohr, S. Penner, and L. M. Young (1985), "Performance of the 100 keV chopper/buncher system of the NBS-Los Alamos RTM Injector," IEEE Transactions on Nuclear Science **32** (5), 3089–3091.

Wilson, P B (1989), "Introduction to wakefields and wake potentials," AIP Conference Proceedings **184** (1), 525–564.

Wisniewski, E E, C. Li, W. Gai, and J. Power (2012), "Generation of annular, high-charge electron beams at the Argonne Wakefield Accelerator," AIP Conference Proceedings **1507** (1), 956–958.

Wollnik, H, and M. Berz (1985), "RELATIONS BETWEEN ELEMENTS OF TRANSFER MATRICES DUE TO THE CONDITION OF SYMPLECTICITY," Nucl. Instrum. Meth. A **238**, 127–140.

Wu, Juhao, Zhirong Huang, and Paul Emma (2008), "Analytical analysis of longitudinal space charge effects for a bunched beam with radial dependence," Phys. Rev. ST Accel. Beams **11**, 040701.

Xiang, Dao (2010), "Laser assisted emittance exchange: Downsizing the x-ray free electron laser," Phys. Rev. ST Accel. Beams **13**, 010701.

Xiang, Dao, and Alex Chao (2011), "Emittance and phase space exchange for advanced beam manipulation and diagnostics," Phys. Rev. ST Accel. Beams **14**, 114001.

Xu, T, *et al.* (2019), "Generation High-Charge of Flat Beams at the Argonne Wakefield Accelerator," in *Proc. 10th International Particle Accelerator Conference (IPAC'19), Melbourne, Australia, 19-24 May 2019*, International Particle Accelerator Conference No. 10 (JACoW Publishing, Geneva, Switzerland) pp. 3337–3340.

Xu, T, *et al.* (2021), "A damping-ring-free electron injector for future linear colliders," in *Proc. 12th International Particle Accelerator Conference (IPAC'21), Campinas, Brazil*.

Yampolsky, Nikolai, Bruce Carlsten, Robert Ryne, Kip Bishofberger, Steven Russell, and Alex Dragt (2010), "Controlling electron-beam emittance partitioning for future x-ray light sources," arXiv:1010.1558 [physics.acc-ph].

Yampolsky, Nikolai, Evgenya I. Simakov, and Alexander Malyzhenkov (2020), "Imposing strong correlated energy spread on relativistic bunches with transverse deflecting cavities," Phys. Rev. Accel. Beams **23**, 054403.

Yang, J, F. Sakai, T. Yanagida, M. Yorozu, Y. Okada, K. Takasago, A. Endo, A. Yada, and M. Washio (2002), "Low-emittance electron-beam generation with laser pulse shaping in photocathode radio-frequency gun," Journal of Applied Physics **92** (3), 1608–1612.

Yokoya, Kaoru, and Pisin Chen (1992), "Beam-beam phenomena in linear colliders," in *Frontiers of Particle Beams: Intensity Limitations*, edited by M. Dienes, M. Month, and

S. Turner (Springer Berlin Heidelberg, Berlin, Heidelberg) pp. 415–445.

Yuri, Yosuke, Mitsuhiro Fukuda, and Takahiro Yuyama (2019), "Formation of hollow ion beams of various shapes using multipole magnets," Progress of Theoretical and Experimental Physics **2019** (5), 10.1093/ptep/ptz024, 053G01, https://academic.oup.com/ptep/article-pdf/2019/5/053G01/28599734/ptz024.pdf.

Yuri, Yosuke, Nobumasa Miyawaki, Tomihiro Kamiya, Watalu Yokota, Kazuo Arakawa, and Mitsuhiro Fukuda (2007), "Uniformization of the transverse beam profile by means of nonlinear focusing method," Phys. Rev. ST Accel. Beams **10**, 104001.

Zeitler, Benno, Klaus Flöttmann, and Florian Grüner (2015), "Linearization of the longitudinal phase space without higher harmonic field," Phys. Rev. ST Accel. Beams **18**, 120102.

Zhang, Zhen, Lixin Yan, Yingchao Du, Zheng Zhou, Xiaolu Su, Lianmin Zheng, Dong Wang, Qili Tian, Wei Wang, Jiaru Shi, Huaibi Chen, Wenhui Huang, Wei Gai, and Chuanxiang Tang (2016), "Tunable high-intensity electron bunch train production based on nonlinear longitudinal space charge oscillation," Phys. Rev. Lett. **116**, 184801.

Zhao, Lingrong, Tao Jiang, Chao Lu, Rui Wang, Zhe Wang, Pengfei Zhu, Yanchao Shi, Wei Song, Xiaoxin Zhu, Chunguang Jing, Sergey Antipov, Dao Xiang, and Jie Zhang (2018a), "Few-femtosecond electron beam with terahertz-frequency wakefield-driven compression," Phys. Rev. Accel. Beams **21**, 082801.

Zhao, Lingrong, Zhe Wang, Chao Lu, Rui Wang, Cheng Hu, Peng Wang, Jia Qi, Tao Jiang, Shengguang Liu, Zhuoran Ma, Fengfeng Qi, Pengfei Zhu, Ya Cheng, Zhiwen Shi, Yanchao Shi, Wei Song, Xiaoxin Zhu, Jiaru Shi, Yingxin Wang, Lixin Yan, Liguo Zhu, Dao Xiang, and Jie Zhang (2018b), "Terahertz streaking of few-femtosecond relativistic electron beams," Phys. Rev. X **8**, 021061.

Zholents, A, P. Heimann, M. Zolotorev, and J. Byrd (1999), "Generation of subpicosecond x-ray pulses using rf orbit deflection," Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **425** (1), 385 – 389.

Zholents, A A, and M. S. Zolotorev (2011), *A new type of bunch compressor and seeding of short wave length coherent radiation*, Tech. Rep. ANL/APS/LS-327 (Argonne National Laboratory).

Zhou, Shian, Dimitre Ouzounov, Heng Li, Ivan Bazarov, Bruce Dunham, Charles Sinclair, and Frank W. Wise (2007), "Efficient temporal shaping of ultrashort pulses with birefringent crystals," Appl. Opt. **46** (35), 8488–8492.

Zolotorev, M, E. D. Commins, and F. Sannibale (2007), "Proposal for a quantum-degenerate electron source," Phys. Rev. Lett. **98**, 184801.