REVIEWS OF MODERN PHYSICS, VOLUME 94, APRIL–JUNE 2022

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
Department of Physics and Northern Illinois Center for Accelerator and Detector Development, Northern Illinois University, DeKalb, Illinois 60115, USA and 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, such as 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.

DOI: 10.1103/RevModPhys.94.025006

CONTENTS

---

[*] jp@anl.gov

I. INTRODUCTION

The ability to control the electron beam distribution produced by a linear accelerator (linac) evolved during the 20th century. The first generation of electron linacs (*e*-linacs) that begun with Wideröe'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, Alvarez *et al.* (1955) 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, Hansen, and Kennedy, 1948). This generation of *e*-linacs culminated in the construction of the 100-GeV electron-positron SLAC linear collider (Dupen, Hogg, and Loew, 1968). Electron bunches in these linacs can be approximated using a Gaussian distribution in phase space that 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, Huang, and Lindberg (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: toward a control at a level finer than the macroscopic scale but coarser than the microscopic scale, which is referred to in the review as the *mesoscopic* level, and toward multidimensional 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 an interesting and useful distribution at the finest level: the *microscopic scale*, where the distribution is described as a "granular" Klimontovich-distribution function (Klimontovich, 1995). Such an ultimate degree of control would open the path toward producing structured beams, such as Wigner-crystal beams (Wigner, 1934) with arbitrary shapes; see Fig. 1(c).

The development of multidimensional shaping began in the late 1990s and early 2000s. The flat-beam generation (Derbenev, 1998; Brinkmann, Derbenev, and Flöttmann, 2001) 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 of

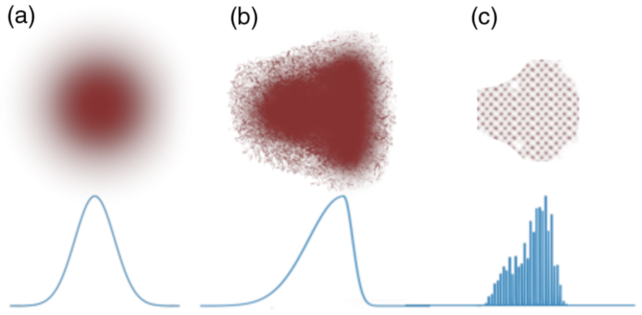

FIG. 1. Definition of the level of control on a beam distribution: (a) macroscopic, (b) mesoscopic, and (c) microscopic scales. The shaded distributions depict the projected beam distribution in a plane defined by any pair of coordinates associated with the beam. The blue traces describe the projection of the distribution along one arbitrary axis.

Edwards *et al.* (2000) and Piot, Sun, and Kim (2006) and the emittance exchange of Ruan *et al.* (2011). The emittance exchange opened up the possibility of shaping a beam in the transverse dimension and transferring it to the temporal dimension, and vice versa. The technique was used to produce a train of subpicosecond bunchlets (Sun *et al.*, 2010a) and bunches with linearly ramped current profile (Ha *et al.*, 2017)

Advanced beam phase-space shaping is needed to enable many accelerator applications, such as improving the efficiency of beam-driven advanced acceleration techniques (Bane, Chen, and Wilson, 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 a 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 (Seeman, 1992; Derbenev *et al.*, 1995)

This review is devoted mainly 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 were reviewed by 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 1 degree of freedom, such as 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 2 degrees of freedom.

This review is organized as follows. Section II introduces the fundamental concepts necessary to the understanding of phase-space manipulations, and Secs. III–VI discuss the various previously mentioned classes of manipulation. Finally, Sec. VII offers some perspective on likely research directions motivated by recent developments.

II. GENERAL PRINCIPLES

A. Hamiltonian formalism

1. Equation of motion under electromagnetic (EM) fields

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

$$\begin{aligned} \nabla\cdot\mathbf{E} &= \frac{\rho_S}{\varepsilon_0}, & \nabla\times\mathbf{E} &= -\frac{\partial}{\partial t}\mathbf{B}, \\ \nabla\cdot\mathbf{B} &= 0, & \nabla\times\mathbf{B} &= \mu_0\mathbf{J}_S + \frac{1}{c^2}\frac{\partial}{\partial t}\mathbf{E}. \end{aligned} \tag{1}$$

In Eq. (1) 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 Sec. II.D. We use the meter-kilogram-second units throughout this review. Introducing the vector and the scalar potential $\mathbf{A}$ and ϕ, respectively, the electromagnetic fields can be written as follows:

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

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

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

In Eq. (3) c is the velocity of light, $\mathbf{x}$ is the coordinate vector, and $\mathbf{p}$ is the canonical momentum conjugate to $\mathbf{x}$.

The Hamiltonian equations of motion are

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

Equations (4) reproduce the following Lorentz force equation:

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

In Eq. (5) $\mathbf{p}_{\rm kin}$ is the kinetic momentum related to the canonical momentum $\mathbf{p}$ as follows:

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

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

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. The position of any particle in the beam can then 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, $\rho(s)$ is 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 along the horizontal plane and perpendicular to the reference trajectory, and the unit vector $\mathbf{e}_y$ is directed toward the reader. The line perpendicular to the reference trajectory at s

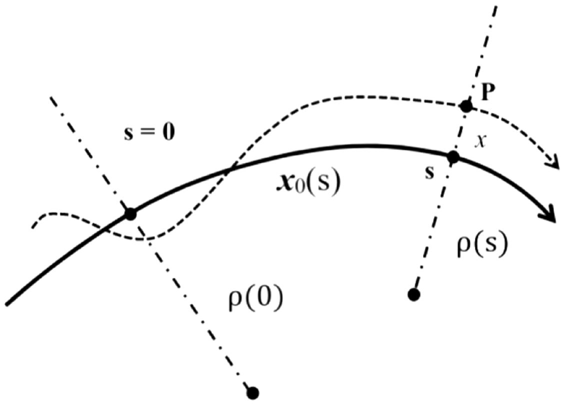

FIG. 2. Curvilinear coordinate system for particle motion. See the text for details.

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 (Landau and Lifshitz, 1969; Goldstein, 2002) 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 that the particle arrives at the transverse plane at s. The new canonically conjugate variables are

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

where U is the energy of the particle. The new Hamiltonian is

$$\begin{aligned}\mathcal{H}(\mathbf{x}_{\text{new}}, \mathbf{p}_{\text{new}}; s) \\ = -eA_s - \left(1 + \frac{x}{\rho(s)}\right) \\ \times \sqrt{\left(\frac{U - e\phi}{c}\right)^2 - m^2c^2 - (p_x - eA_x)^2 - (p_y - eA_y)^2}.\end{aligned} \tag{9}$$

In Eq. (9) $\rho(s)$ 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 replacements $t \to s$, $\mathbf{x} \to \mathbf{x}_{\text{new}}$, $\mathbf{p} \to \mathbf{p}_{\text{new}}$, 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 $A_x = A_y = \phi = 0$. In addition, 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 $(x, y, -v_0(t - t_0); p_x/p_0, p_y/p_0, (U - U_0)/(v_0p_0))$ (Ruth, 1986; MacKay and Conte, 2012). Note that 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

$$\frac{p_x}{p_0} \approx \frac{dx}{ds} = x', \qquad \frac{p_y}{p_0} \approx \frac{dy}{ds} = y', \qquad \frac{U - U_0}{v_0p_0} \approx \frac{p_s - p_0}{p_0} = \delta. \tag{10}$$

We 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. Therefore, the canonical variables in this case are

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

The corresponding scaled Hamiltonian is

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

In Eq. (13) $[\cdots]$ implies removal of 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, we need to go back to the original Hamiltonian equation of motion (4) or the Lorentz force equation (5), as done in Sec. II.B.2. The presence of a longitudinal magnetic field can be treated by working in a rotating frame, as discussed in Sec. II.B.4.

3. Symplecticity

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

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

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

It is often useful to introduce the subspace as follows:

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

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

When the gradient vector ∇ is introduced in 6D phase space as

$$\nabla_j=\frac{\partial}{\partial\zeta_j}, \tag{16}$$

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

$$\frac{d}{ds}\mathcal{Z}=J\nabla\mathcal{H}_S(\mathcal{Z};s). \tag{17}$$

In Eq. (17) we have introduced the unit symplectic matrix

$$J=\begin{pmatrix}J_{2D}&0&0\\ 0&J_{2D}&0\\ 0&0&J_{2D}\end{pmatrix},\qquad J_{2D}=\begin{pmatrix}0&1\\ -1&0\end{pmatrix}. \tag{18}$$

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

$$\bar{\mathcal{Z}}=M(\mathcal{Z}),\quad \bar{\zeta}_i=M_i(\mathcal{Z}). \tag{19}$$

The inverse map is

$$\mathcal{Z}=M^{-1}(\bar{\mathcal{Z}}). \tag{20}$$

We introduce the Jacobian matrix R whose components are

$$R_{ij}(\mathcal{Z})=\frac{\partial\bar{\zeta}_i}{\partial\zeta_j}. \tag{21}$$

For Hamiltonian dynamics, the Jacobian matrix is symplectic:

$$R^TJR=RJR^T=J. \tag{22}$$

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

$$\det(R)=1. \tag{23}$$

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

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

$$\bar{\mathcal{Z}}=R\mathcal{Z}. \tag{24}$$

In Eq. (24) R is the Jacobian matrix given by Eq. (21), whose elements are independent of $\mathcal{Z}$.

B. Single-particle motion in an external field

In this 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, the components of the static magnetic fields corresponding to a bending magnet and quadrupole are given by

$$B_x=B_1(s)y,\qquad B_y=-B_0(s)+B_1(s)x. \tag{25}$$

In Eqs. (25) $B_0=p_0/e\rho$ is the strength of the dipole magnet bending the particle horizontally. In computing the vector potential A_s, one finds the scaled Hamiltonian equation (13) up to the quadratic terms:

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

In Eq. (26) $K_q=eB_1/p_0c$ 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 (17). The equations in the transverse directions, after reducing first-order differential equations to second-order ones, become

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

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

In deriving Eqs. (27) and (28), we use the symbol x' as both the canonical momentum and the slope dx/ds. We indicate that ρ and K_q are functions of s. However, we assume that the functions are piecewise constant and neglect the transition effects. Equation (27) shows that the motion in the x direction is influenced both the quadrupole force and the centripetal force due to the curvature. For historical reasons, 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 the opposite of that in the y direction, focusing in both directions can be achieved either by the centripetal focusing (weak focusing) or by arranging the adjacent quadrupoles to have opposite signs (strong focusing). For details, see Wiedemann (1999).

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

$$\begin{aligned}y&=\sqrt{2J_y\beta_y}\cos(\psi_y),\\ y'&=\frac{\beta_y'}{\beta_y}y-\sqrt{\frac{2J_y}{\beta_y}}\sin(\psi_y).\end{aligned} \tag{29}$$

In Eqs. (29) J_y is a constant and

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

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

$$\alpha_y = -\frac{1}{2}\beta_y', \qquad \gamma_y = \frac{1+\alpha_y^2}{\beta_y}. \tag{31}$$

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

$$\gamma_y y^2 + 2\alpha_y y y' + \beta_y y'^2 = J_y. \tag{32}$$

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

$$2\beta_y\beta_y'' - \beta_y'^2 + 4\beta_y^2 K_q = 4. \tag{33}$$

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 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: $C_{s\tau}$ (cosinelike), which $C_{\tau\tau} = 1$ and $C'_{\tau\tau} = 0$, and $S_{s\tau}$ (sinelike), with $S_{\tau\tau} = 0$ and $S'_{\tau\tau} = 1$. The full solution can then be written as follows (Brown, 1968; Wiedemann, 1999):

$$\begin{aligned} x_s &= C_{s\tau}x_\tau + S_{s\tau}x'_\tau + \eta_{s\tau}\delta, \\ x'_s &= C'_{s\tau}x_\tau + S'_{s\tau}x'_\tau + \eta'_{s\tau}\delta. \end{aligned} \tag{34}$$

In Eq. (34)

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

The longitudinal motion is given by

$$\begin{aligned} \frac{dz}{ds} &= -\frac{x}{\rho(s)}, \\ \frac{d\delta}{ds} &= 0, \end{aligned} \tag{36}$$

with the solution

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

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

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

The $R_{5,j}$ elements in Eq. (38) are

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

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

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

$$D_\ell = \begin{pmatrix} 1 & \ell \\ 0 & 1 \end{pmatrix}. \tag{40}$$

In Eq. (40) $\ell = s - \tau$ is the length of the free space. If we have a thin quadrupole of focal length f_C, then the block becomes

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

When Eq. (39) is used, the matrix for a sector bending magnet of constant ρ and deflection angle θ can be found; see Eq. (77) of 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 by the result as follows:

$$R_{B\mathrm{rect}}(\rho) = \begin{pmatrix} 1 & \rho\sin(\theta) & 0 & \rho[1-\cos(\theta)] \\ 0 & 1 & 0 & 2\tan(\theta/2) \\ -2\tan(\theta/2) & -\rho[1-\cos(\theta)] & 1 & -\rho[\theta-\sin(\theta)] \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{42}$$

Equation (42) was also given by 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:

$$R_{\mathrm{DL}} = \begin{pmatrix} 1 & \ell & 0 & \eta \\ 0 & 1 & 0 & 0 \\ 0 & \eta & 1 & \xi \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{43}$$

---

[1] However, the sign convention in this reference is such that the z coordinate has the opposite sign of ours. Thus, the $(5, j)$ and $(j, 5)$ elements there should be multiplied by -1. Otherwise, the matrix is not symplectic.

In Eq. (44)

$$\begin{aligned} \ell &= d + 2\rho\, \sin(\theta), \\ \eta &= 2[d + \rho\, \sin(\theta)]\tan\left(\frac{\theta}{2}\right), \\ \xi &= -2\rho[\theta + \sin(\theta)] + 8\rho\, \tan\left(\frac{\theta}{2}\right) + 4d\tan^2\left(\frac{\theta}{2}\right). \end{aligned} \tag{44}$$

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

2. rf photocathode cavity

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

The nonvanishing components of the EM field, keeping terms quadratic in radial coordinate r, take the following form:

$$\begin{aligned} E_s &= \left(\mathcal{E}(s) - \frac{r^2}{4}[\mathcal{E}''(s) - k^2\mathcal{E}(s)]\right)\sin(\omega t), \\ E_r &= -\tfrac{1}{2} r\mathcal{E}'(s)\, \sin(\omega t), \\ B_\varphi &= \frac{k}{2c} r\mathcal{E}(s)\, \cos(\omega t). \end{aligned} \tag{45}$$

In Eqs. (45) $k = 2\pi/\lambda = \omega/c$, λ is the rf wavelength and

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

In Eq. (46) 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 unit step function. Inside the cavity $0 \le s < s_f$, 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 shown in Fig. 3 can be chosen to produce fields close to those in Eq. (45) (McDonald, 1988).

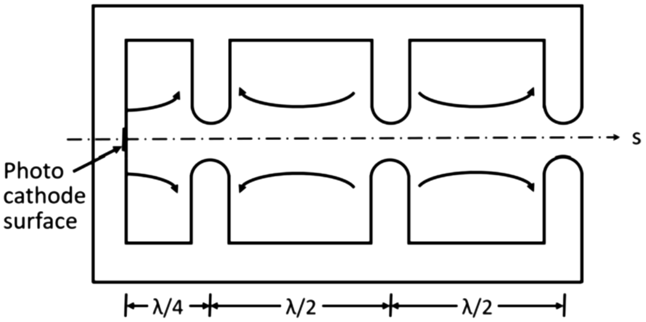

FIG. 3. Schematics of a rf photocathode cavity. The first cell where the photocathode is located is a half cell. From Kim, 1989.

a. Longitudinal motion

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

$$\psi = \omega t - ks = k(ct - s) = \psi_0 - kz. \tag{47}$$

In Eq. (47) ψ_0 is the reference phase. Note that z is the same as that in Eq. (11) for particles moving with relativistic velocity c. The equations of motion are

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

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

$$\begin{aligned} \psi &= \frac{1}{2\alpha\, \sin(\psi_i)}\left[\sqrt{\tilde{\gamma}^2 - 1} - (\tilde{\gamma} - 1)\right] + \psi_i, \\ \gamma &= 1 + \alpha\{ks\, \sin(\psi_i) + \tfrac{1}{2}[\cos(\psi) - \cos(\psi + 2ks)]\}. \end{aligned} \tag{49}$$

In Eqs. (49) ψ_i is the initial phase and

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

and we introduced the acceleration strength parameter

$$\alpha = \frac{eE_0}{2mc^2k}. \tag{51}$$

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

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

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

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

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

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

In Eq. (55) $v_s = ds/dt$. 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 the exit is the time integral of Eq. (56). 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 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. The increase can be corrected using the emittance compensation technique (Carlsten, 1989; Serafini and Rosenzweig, 1997; Ferrario *et al.*, 2000, 2007; Wang, 2006; Miginsky, 2009; Flöttmann, 2017).

3. Transverse deflecting cavity

In the rf photocathode cavity considered in Sec. II.B.2, the transverse deflection occurs only at the cavity exit. To see how a sustained deflection can be induced, we write the transverse part of the third and fourth of Maxwell equations [Eq. (1)] as follows:

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

The first line of Eqs. (58) is the differential form of the Panofsky-Wenzel theorem (Panofsky and Wenzel, 1956), and the second line was noted by Paramonov and Flöttmann (2019). We now consider a traveling wave with fields in the following form:

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

Assuming that the phase velocity of the wave and the particle velocity are both equal to c, we can show that the lhs's of both equations are proportional to the transverse force. From the expressions on the rhs, it then follows 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 that shown in Fig. 4, supports hybrid electromagnetic (HEM) modes that can provide a sustained transverse force. The mode has been extensively studied since 1963 (Hahn, 1963; Garault, 1964; Flöttmann and Paramonov, 2014; 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

$$\begin{aligned}E_x &= -\hat{E}\frac{k^2}{4}(a^2+x^2-y^2)\sin(\psi),\\ E_y &= -\hat{E}\frac{k^2}{2}xy\,\sin(\psi),\\ E_s &= \hat{E}kx\,\cos(\psi),\\ cB_x &= \hat{E}\frac{k^2}{2}xy\,\sin(\psi),\\ cB_y &= -\hat{E}\frac{k^2}{4}(a^2-4/k^2+x^2-y^2)\sin(\psi),\\ cB_s &= -\hat{E}ky\,\cos(\psi).\end{aligned} \quad (60)$$

The Lorentz force is then

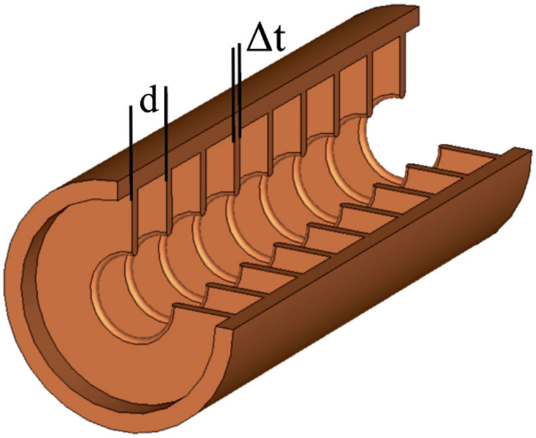

FIG. 4. Iris-loaded structure producing HEM modes for transverse deflection. Δt is the thickness of the disks and d is the structure periodicity. From Paramonov and Flöttmann, 2019.

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

For beam-shaping applications, we want the transverse deflection to change 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 that $kz \ll 1$, the force to first order in the deviation variables (x, z) becomes

$$\mathbf{F} \approx e\hat{E}(kz, 0, kx). \quad (62)$$

The equation of motion inside the HEM structure is then

$$\begin{aligned} \frac{dx}{ds} &= x', & \frac{dx'}{ds} &= \frac{\kappa_T}{\ell}z, \\ \frac{dz}{ds} &= 0, & \frac{d\delta}{ds} &= \frac{\kappa_T}{\ell}x. \end{aligned} \quad (63)$$

In Eq. (63) ℓ is the length of the cavity and

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

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 using the following matrix connecting the entrance and exit values of the vector $(\mathbf{X}, \mathbf{Z}) = (x, x', z, \delta)$ (Cornacchia and Emma, 2002):

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

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

The $(6,5)$ element in R_{TDC} wreaks havoc in the transverse-to-longitudinal emittance exchange discussed in Sec. VI. However, it can be removed using 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 purposes. Assuming cylindrical symmetry around the s axis perpendicular to the cathode surface, the vector potential can be written as

$$\mathbf{A}(r, s) = \tfrac{1}{2} r B(s)\mathbf{e}_\varphi. \quad (66)$$

In Eq. (66) we are using the cylindrical coordinates (r, φ, s) with an origin at the cathode center, and s is the distance away from the cathode. The magnetic field is

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

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

The transverse motion is given by

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

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:

$$\mathbf{x} = M_\varphi \mathbf{x}_R, \qquad M_\varphi = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ -\sin(\varphi) & \cos(\varphi) \end{pmatrix},$$

$$\text{with } \varphi(s) = \int_0^s \kappa(\tau)d\tau. \quad (69)$$

By differentiating, we obtain

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

The first term is

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

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

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

The second derivative is

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

Using Eq. (73) with Eq. (68) and choosing

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

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

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

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

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

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

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

In Eq. (79) 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 the field part to the 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

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

In Eq. (80)

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

Note that 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 follows 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

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

In Eq. (84) R is the Jacobian matrix given by Eq. (21). Comparing this to the rhs of Eq. (83), we find the following transformation of the distribution function corresponding to the transform of the phase-space variable in Eq. (19):

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

Equation (85) is Liouville's theorem (Liouville, 1838), which is fundamental to many beam-shaping schemes involving external fields.

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

We now return to the case where the map is linear and can be represented by matrix R. As a simple example, consider the following distribution in the (z, δ) subspace:

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

In Eq. (86) ℓ is the bunch length, over which the distribution is regarded as constant.

The corresponding line density in z is constant:

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

The matrix for the bunch compression section consisting of chirping (energy change linear in z) followed by 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

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

In Eq. (89) we introduced the compression factor C:

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

Therefore, the distribution after the beamline when Eq. (85) is applied becomes

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

The corresponding line density is

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

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

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

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

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

The first-order moments can be made to vanish if we note that ζ's are deviation variables; that is, they are 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:

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

The second-order moments are

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

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

Note that the R matrix is symplectic, thereby 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

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

In Eq. (97) $\langle\mathbf{X}\mathbf{X}^T\rangle$, etc., are 2×2 matrices and $\langle\cdots\rangle$ is the averaging operation given 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 as follows:

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

The projected emittances are in general not invariant.

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

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

In Eq. (100), $A$is a symplectic matrix and the two-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

$$\varepsilon_j = \sqrt{\det\langle\mathbf{X}_j\mathbf{X}_j{}^T\rangle} \tag{101}$$

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

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

From Eq. (101), it follows that the parameters are related by $\beta_j\gamma_j - \alpha_j^2 = 1$. Note that this relation is equivalent to the second function of Eqs. (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, we find that the principal emittance in the jth subspace is the statistical average of the particles' action variables in that subspace:

$$\varepsilon_j = \langle J_j\rangle. \tag{103}$$

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

The beam matrix equation (102) can be diagonalized with the matrix

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

Indeed,

---

[2] Dragt (2011) referred to this emittance as the *eigenemittance*.

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

In Eq. (105) we used the notation

$$\beta_j{}^* \equiv 1/\gamma_j. \tag{106}$$

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 that the second function of Eqs. (31) can be written as

$$\beta_j = (1 + \alpha_j{}^2)/\gamma_j = \beta_j{}^*[1 + (\ell/\beta_j{}^*)^2]. \tag{107}$$

Equation (107) is a well-known and useful equation describing how the beta function changes away from the waist.

Performing diagonalization in other dimensions, we obtain

$$\begin{aligned} &BA\Sigma A^T B^T \\ &\quad = \mathrm{diag}(\varepsilon_1\beta_1{}^*, \varepsilon_1/\beta_1{}^*, \varepsilon_2\beta_2{}^*, \varepsilon_2/\beta_2{}^*, \varepsilon_3\beta_3{}^*, \varepsilon_3/\beta_3{}^*). \end{aligned} \tag{108}$$

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 $\mathrm{diag}(\varepsilon_1, \varepsilon_1, \varepsilon_2, \varepsilon_2, \varepsilon_3, \varepsilon_3)$. 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 the z axis is an example of such a transformation. If particles are accelerated in the z direction, 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 also invariant under acceleration is

$$\varepsilon_j{}^n = \beta\gamma\varepsilon_j. \tag{109}$$

Note that β and γ in Eq. (109) 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 in Eq. (109) is referred to as *normalized* emittance, while that defined in Eq. (101) is known as the *unnormalized* emittance. The emittances in this review are unnormalized emittances unless otherwise specified.

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). Note first that the quantities

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

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 equation (108). Thus, we obtain

$$\Gamma^n = 2(-1)^n[(\varepsilon_1)^{2n} + (\varepsilon_2)^{2n} + (\varepsilon_3)^{2n}]. \tag{111}$$

Equation (111) gives three equations for three principal emittances. We also have

$$\det(\Sigma) = (\varepsilon_1\varepsilon_2\varepsilon_3)^2. \tag{112}$$

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, is 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 repartition of the emittances is not possible by means of symplectic transformation. Thus, for example, we can exchange $(\varepsilon_x, \varepsilon_y, \varepsilon_z)$ for $(\varepsilon_z, \varepsilon_y, \varepsilon_x)$ but cannot repartition it to $(\varepsilon_x/4, 4\varepsilon_y, \varepsilon_z)$. This fact appears to have first been noted in an accelerator physics context by Courant (1966).

Although the set of three principal emittances does not change, the ordering in the set can 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 EEX, we consider a beam matrix of the form

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

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

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 that a beamline gives rise to the following transformation:

$$R\begin{pmatrix} \Sigma_x & 0 \\ 0 & \Sigma_z \end{pmatrix}R^T = \begin{pmatrix} \Sigma_{z'} & 0 \\ 0 & \Sigma_{x'} \end{pmatrix}. \tag{114}$$

In Eq. (114) $\Sigma_{x'}$ is a 2×2 beam matrix symplectically connected to Σ_x and $\Sigma_{z'}$ is similarly connected to Σ_z. It can be shown that R is of the following form:

$$R = \begin{pmatrix} 0 & B \\ C & 0 \end{pmatrix}, \tag{115}$$

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

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

Therefore, an EEX transformation is much more than an exchange of the magnitude of the subspace area: it exchanges the entire subspace $\mathbf{X}$ for the entire subspace $\mathbf{Z}$. A phase-space shape in $\mathbf{X}$ will be transformed into 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 PSE.

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

4. Emittance repartitioning

Although emittances can be entirely exchanged only in a symplectic transformation, an emittance repartitioning is possible if a nonsymplectic 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 Sec. II.B.4. We construct the following beam matrix of the rotating frame-vector equation (80):

$$\Sigma_B = \begin{pmatrix} \langle \mathbf{X}_R {\mathbf{X}_R}^T \rangle & \langle \mathbf{X}_R {\mathbf{Y}_R}^T \rangle \\ \langle \mathbf{Y}_R {\mathbf{X}_R}^T \rangle & \langle \mathbf{Y}_R {\mathbf{Y}_R}^T \rangle \end{pmatrix}. \tag{117}$$

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 emittances of the beam in the absence of the magnetic field ε_{01} and ε_{02}. In view of Eq. (82), we have

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

In Eq. (118) Σ_0 is the beam matrix for $B = 0$. Thus,

$$\varepsilon_{B1}\varepsilon_{B2} = \varepsilon_{01}\varepsilon_{02}. \tag{119}$$

Invoking Busch's theorem for a beam, it is shown in Sec. 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 to calling this example an emittance repartition. 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 repartitioning schemes always involve nonsymplectic elements, such as beam masks or tapered absorbing blocks. A generalization of Busch's theorem to a noncylindrically symmetric system was discussed by Groening, Xiao, and Chung (2018). A general emittance repartitioning was investigated by Carlsten, Bishofberger, Duffy *et al.* (2011).

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

5. Nonlinear case

If the Hamiltonian contains polynomials of an 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 in Eq. (8) can be expressed in Taylor series in $\mathcal{Z}$. The map from $\mathcal{Z} \to \bar{\mathcal{Z}}$ can be found by solving the equation of motion using the original Hamiltonian equation (9) in the following form:

$$\overline{\zeta_i} = R_{ij}\zeta_j + T_{ijk}\zeta_j\zeta_k + \cdots. \tag{120}$$

The coefficients $T_{ijk}, \ldots$ were worked out in detail by Brown (1968) (however, see footnote 1 for the sign convention). The constraints on these coefficients from symplectic property were discussed by 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 Ruth (1986). A powerful method using Lie canonical transformation with polynomial generators has been developed that can handle high-order polynomial terms with numerical computations (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 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 $\beta c, \beta \approx 1$ 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 (Wilson, 1989; Heifets and Kheifets, 1991; Chao, 1993; Stupakov, 2001):

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

The minus sign in Eq. (121) 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 wakefield is V/(C/m). Here we are considering the case where the wake is uniform along the beam chamber. When the wake is localized, the wake function is defined as 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]

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

In Eq. (122) 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 (Nielsen, Sessler, and Symon, 1959; Chao, 1993):

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

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

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

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 while keeping only the $\mathbf{Z}$. In the more general case, we should write

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

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 follows as an integral in 6D phase space $\mathcal{Z}$:

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

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

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

In Eq. (129) r_e is the classical electron radius and N is the total number of particles. The subscript coll emphasizes that this is the beam-generated, collective force.

The meaning of Eqs. (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 Sec. II.D.5. Next we discuss three representative cases of impedance.

2. Fields due to boundary perturbation

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

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

A wakefield device with corrugated walls referred to as a *dechirper* can be useful in correcting the energy chirp that may arise during compression of a bunch with high peak current (Bane and Stupakov, 2012, 2016; Deng *et al.*, 2014; Emma *et al.*, 2014). 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 in the following form:

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

with $Z_0 = 377\ \Omega$ being the free-space impedance. For a flattop charge density, we see readily in Eq. (122) that the energy loss is linear in z, with 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) as

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

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)

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

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 $\lambda(z) = \text{const}$ and a uniform cross section of radius b traveling at velocity βc, $\beta \approx 1$ along the axis of a circular, perfectly conducting pipe of radius a. The nonvanishing EM field components in cylindrical coordinates

---

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

are the radial electric field E_r and the azimuthal magnetic field B_φ given by

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

Now consider a density with a variation such 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 the 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 illustrated by Fig. 28(b) in Sec. V.A.1. If there is a periodic modulation in the density, on the other hand, then the modulation could become enhanced, as discussed in Sec. V.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 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 the z variable, and observe that only the 0th component in the azimuthal series will contribute. The impedance per unit length can then be computed with the following result (Rosenzweig *et al.*, 1997; Venturini, 2008):

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

When $K_1(kb/\gamma)$ for small kb/γ (Huang, Wu, and Shaftan, 2005; Venturini, 2008) is expanded, one obtains

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

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 results known in 1912 (Schott, 1912), the wakefield at a distance s ahead of the emitting particle (note that s is negative) can be written as follows (Derbenev *et al.*, 1995; Murphy, Krinsky, and Gluckstern, 1997; Saldin, Schneidmiller, and Yurkov, 1997):

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

In Eq. (141) ρ is the radius of curvature. The function $G(s)$ vanishes for $s \ge 0$. 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 Eq. (142) we used integration by parts and also assumed that $|s - s'|$ is much greater than ρ/γ^3. The impedance associated with this force is referred to as *coherent synchrotron radiation* (CSR) impedance:

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

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 (Iogansen and Rabinovich, 1960; Faltens and Laslett, 1975). CSR impedance can be significant when bending magnets are used in linacs, such as in the final chicane magnet of a bunch compressor where the current is high (Heifets, Stupakov, and Krinsky, 2002; Huang and Kim, 2002).

Synchrotron radiation also has a high-frequency incoherent part referred to as incoherent synchrotron radiation (ISR) (Sands, 1955, 1969; Saldin, Schneidmiller, and Yurkov, 1996). 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 acts back on the beam, limiting ourselves to the 4D phase space $\mathcal{Z} = (\mathbf{X}, \mathbf{Z})$ for simplicity. The first question to ask is whether 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 nonrelativistic plasmas (Pines and Bohm, 1952) and were extended to relativistic beams (Sørensen, 1988; Rosenzweig *et al.*, 1997; Kim and Lindberg, 2011). 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 certain cases, such as generation of ultrashort bunch trains via nonlinear plasma oscillation, as discussed in Sec. V.A.3.

We now discuss how a beam-generated field affects the evolution of the phase-space distribution $f(\mathcal{Z};s)$ in a beamline beginning at $s=0$ (Heifets, Stupakov, and Krinsky, 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 Liouville's theorem as follows:

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

Equation (147) is the same as Eq. (85), with $R_{\tau s}$ a combination of any transformation matrix discussed in Sec. II.B. For notational clarity, a variable at location τ was given the same subscript, e.g., $\mathcal{Z}_\tau$. Let $d\delta_{\mathrm{coll}}(z_\tau;\tau)$ 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

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

From Eq. (148) the bunching factor at s [Eq. (128)] can be obtained as follows after some mathematical manipulation (Huang and Kim, 2002):

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

In Eq. (149)

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

On the first line of Eq. (150), $\zeta_{j,\tau}$ denote the components of $\mathcal{Z}_\tau$; see Eq. (14). The second line of Eq. (150) is for the case $\mathcal{Z} = (\mathbf{X},\mathbf{Z})$. Equation (149) is an integral equation for the evolution of the bunching factor. We consider the initial distribution in the following form:

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

The first part of Eq. (151) $\bar{f}_0$ 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 as 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, $\mathcal{Z} = \mathbf{Z} = (z,\delta)$. 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 $n_0 = 1/\ell$ the line density. The quantities in the exponent of Eq. (149) are

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

In Eq. (153) we simplified the notation as $R_{s0,56} \to R_{56}(s)$ and $(z_0,\delta_0) \to (x,\delta)$. The integral over δ and z can be performed to obtain

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

In Eq. (154) $I_A = ec/r_e$ is the Alfven current, $I = ecn_0C(\tau)$ is the peak current at τ, and

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

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 as either *mesoscopically* shaped or *macroscopically* smooth bunches; see Fig. 1 and the related discussion in Sec. 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 that is subsequently shaped by the methods presented in Secs. IV and V. 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 that is injected into the linac. The emission of the initial electron-bunch distribution from the cathode surface and its evolution through the gun are complex and varied. Electron gun physics was discussed in depth by Dowell, Schmerge, and Lidia (2008), Rao and Dowell (2013), and Dowell (2016) and in references therein. In the introduction to this section, we present only the 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 (such as the 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 the initial cathode distribution. Sections IVand V include the other effects.

Electron guns used in *e*-linacs can be classified by their electric field (dc and rf) or their cathode type. There are a variety of cathodes in use, but they can be conveniently classified by their emission mechanism: field emission (FE), thermionic emission (TE), or photoemission (PE). 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 shortened to “photocathode” gun. However, as previously stated, the impact of the gun fields on shaping is ignored (in this section), so we drop the dc and rf labels unless otherwise noted. 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 workhorses of storage rings for light sources around the world.[4] 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 research and development (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 self-amplified spontaneous emission FEL facilities, and they are being used for both transverse- and longitudinal-shaping applications. FE and TE cathodes are not used for longitudinal shaping because the electron emission from these cathodes follows the time structure of the applied electric field (ranging from continuous in a dc gun to nanosecond scale in a rf gun), which is too long for the timescales of interest (femtosecond to picosecond scale) 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, see Jensen (2018).

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

B. Cathode review

In this section, we review different cathodes types and summarize their properties that are 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 are introduced in the corresponding sections.

1. Current density

There is great variation in the magnitude of the current density J generated by the three cathode types. The emission time row in Table I shows the response times of the cathodes. Note that, while the current density does not play a direct role

---

[4] See https://lightsources.org/lightsources-of-the-world/.

TABLE I. Cathode characteristics.

	Field emission (large area)	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 $\sigma_{x,y}$	1–3	1–3	0.1–10	mm
Typical emittance $\varepsilon_{x,y}$	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

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 which the electrons are extracted. The local FE current density is given by the Fowler-Nordheim equation (Fowler and Nordheim, 1928)

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

with material constant A_{FE}. Lowercase letters are used to differentiate local (j_{FE} and w) from macroscopic (J_{FE} and W) parameters. 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 ($j_{\mathrm{FE}} \sim 10^6$ 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 ~0.1 mA per tip, which is inadequate for operation of the linacs that we consider in this review since they require larger current (at least 1 A 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 ultrananocrystalline diamonds (UNCDs) (Baryshev *et al.*, 2014). These surfaces are composed of a quasicontinuous distribution of electron emitters of macroscopic area ~1 mm^2 and achieve macroscopic current densities of $J_{\mathrm{FE}} \sim 1$ A/mm^2, which are 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 is operated in guns with applied electric field on the cathode E_c ($\sim$100 MV/m). Dispenser cathodes have work functions W ($\sim$1.6 eV), are representative of TE cathodes, and are operated with a temperature T ($\sim$1400 K) to allow electron emission. The TE macroscopic current density is given by Richardson's law (Richardson, 1913) as

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

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 ($\sim$100 MV/m) can achieve current densities of $J_{\mathrm{TE}} \sim 1$ 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

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

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 ($\sim$100 MV/m) and achieve the highest macroscopic current density of the three cathode types ($J_{\mathrm{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 seen in Sec. III.B.3.b, 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, and z.

The average kinetic energy of electrons emitted from a FE cathode (Forbes, 2016) is

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

where m_e is the electron mass and typically $E_k = 0.3$ eV for FE cathode tips. Electrons emitted from a TE cathode have an average kinetic energy

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

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

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

where a copper cathode (work function of 4.65 eV) illuminated by a laser of wavelength $\lambda = 248$ nm has $E_k = 0.18$ eV.

An examination of the three expressions for E_k reveals two noteworthy points. 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 $\phi_{\rm eff}$ to account for the Schottky effect (Dowell and Schmerge, 2009), which is the reduction of the work function by the applied field. 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 on the fourth row 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 the end,

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

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 that we are ignoring the collective effects of the beam and higher-order magnetic optics of the beamline. [For the derivation of Eq. (163), see Eq. 6.45 of Carey (1987) and the discussion therein. Note also that the smallest spot size is not necessarily the waist of Eq. (106).] A similar expression for the minimum longitudinal bunch length ($\sigma_{z,\rm min}$) at the end of the beamline can be derived as

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

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., a 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: increasing the initial spot size ($\sigma_{x,0}, \sigma_{z,0}$) or decreasing the elements (R_{12} and R_{56}). The third way to increase the resolution is to decrease the emittance ($\varepsilon_x, \varepsilon_z$); 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 that 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 six dimensions is given by the following product of the three normalized intrinsic 2D emittances:

$$\varepsilon^n_{\rm 6D} = \varepsilon^n_x \varepsilon^n_y \varepsilon^n_z. \tag{165}$$

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

$$\varepsilon^n_x = \sigma_x \sigma_{p_x}, \tag{166}$$

where $\sigma_x = \sqrt{\langle x^2 \rangle}$ and $\sigma_{p_x} = \sqrt{\langle p_x^2 \rangle}/m_0 c$. As corresponding expressions apply for each coordinate, we can rewrite Eq. (166) to give the intrinisic rms normalized emittance at the cathode for each of the 2D phase-space planes as

$$\varepsilon^n_c = \sigma_c \sigma_{p_c}, \tag{167}$$

where σ_c and σ_{p_c} are the initial rms bunch size and the dimensionless rms momentum at the cathode in the horizontal, vertical, or 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 (such as when taking 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

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

where c is the speed of light and 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 intrinisic rms normalized emittance at the cathode per rms spot size (i.e., $\sigma_{p_x} = \varepsilon^n_x/\sigma_x$) by virtue of Eq. (166). It is given in units of μm/mm. Equation (168) 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. 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 intrinisic 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 Sec. 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 for the transverse parameters ($\varepsilon_{x,y}$ and $\sigma_{x,y}$) listed in 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 ($\mathrm{MTE} = (2/3)E_k$) of the electrons emitted from the cathode. By equipartition of energy, we know that the mean longitudinal energy is $(1/3)E_k$, and we can find the following expression of the intrinisic rms normalized longitudinal emittance at the cathode per rms bunch length:

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

4. Response time

The response time of a cathode is defined as the temporal lag between the excitation of the cathode and the 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 that 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 femtosecond scale for metallic photocathodes to tens of picoseconds for certain semiconductor photocathodes (Smedley and Poelker, 2012). The emittance time row in Table I shows the 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 previously mentioned considerations, the source designer must choose one of the three electron cathodes for the 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 the current density. The larger the spot size on the cathode, the more the 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 by the linac. TE cathodes do not require laser systems but have the added complexity of operating at $T > 1000\,°\mathrm{C}$. FE cathodes do not require laser systems or heating and thus are the simplest of all.

We end this section with the caveat that we left out hybrid cathodes that combine two of the three basic emission mechanisms: photoassisted field emission cathodes (Mustonen *et al.*, 2011; Swanwick *et al.*, 2014) and photoassisted thermionic emission cathodes (Sun, Lewellen, and Feldman, 2006). These are potentially important cathodes since they open up the possibility of longitudinal shaping for FE and TE cathodes; however, we omit 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 Secs. IV–VI. In this section, we describe how the smooth distributions are produced 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 Sec. III.B.1, large-area engineered FE cathodes can be based on FEAs, CNTs, or UNCDs. All of their surfaces are composed of a quasicontinuous 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 case of the randomly oriented CNTs, the emission sites are shown in the top center image of Fig. 5 (Mihalcea *et al.*, 2015). The emission site of the UNCD cathode, shown in the top right image of Fig. 5, 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,

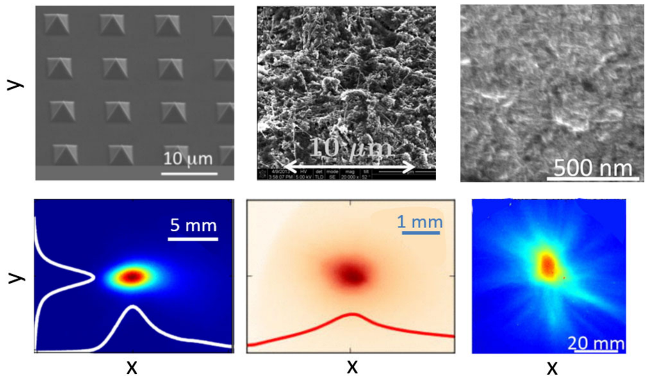

FIG. 5. Three types of large-area FE cathodes. Surface images (top row from left to right) of FEA, CNT, and UNCD cathodes are displayed. The distance between the emitters is largest for FEAs (10 μm), smaller for CNTs (0.1–1 μm), and smallest for UNCDs (10–100 nm). Smooth beam images (bottom row from left to right) generated by FEA, CNT, and UNCD cathodes are displayed. Adapted from Baryshev *et al.*, 2014, Piot *et al.*, 2014, and Mihalcea *et al.*, 2015.

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 nonuniformity of the emitters across the cathode. The discreteness of the emitters becomes less noticeable as the separation between sites decreases. Another limitation of the FE cathode rf gun is the large energy spread and long 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. 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.[5] 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 in achieving a short pulse. However, researchers at SACLA (Asaka *et al.*, 2017) developed a low-emittance thermionic-gun-based injector. In the injector, electron beams are emitted 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 compressed to produce an electron beam with a high peak current (3 to 4 kA) and a low transverse normalized-slice emittance (below 1 μm) sufficient to drive a compact free-electron laser.

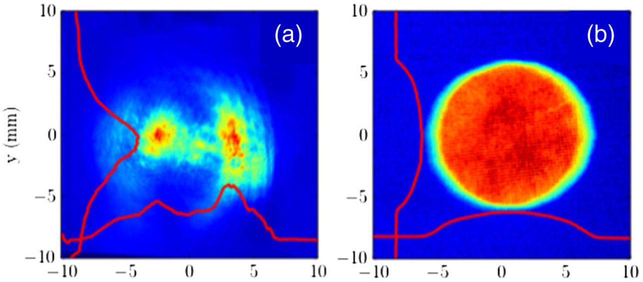

FIG. 6. Comparison of (a) inhomogeneous and (b) homogeneous laser profiles at the plane of the PE cathode. From Halavanau *et al.*, 2017.

3. PE-based smooth transverse distributions

Many applications of the electron source require smooth transverse distributions, such as the uniform flattop or the Gaussian profile. While the typical IR output from the photocathode laser (TEM_{00}) is nearly Gaussian at the output, its shape can become distorted during the frequency up-conversion process to generate UV light. 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 (such as a microlens array) and active (such as a deformable mirror). In general, passive elements are simpler, but active elements have greater capabilities for obtaining complicated transverse distributions, as we later describe.

Microlens arrays (MLAs) are passive optical elements used to homogenize the laser's transverse profile (Bich *et al.*, 2008). Figure 7(a) shows a schematic of homogenizing optics where an UV laser beam (248 nm) passes 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 by Halavanau *et al.* (2017), who imaged the laser from the homogenization plane to a PE cathode plane located approximately 3.5 m away with an imaging system [Figs. 7(b) and 7(c)]. Another advantage of passive homogenization systems (such as 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 timescale of the feedback loop. On the other hand, the passive system will instantaneously transform an arbitrary inhomogeneous laser profile into a homogeneous one.

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

---

[5] See https://lightsources.org/lightsources-of-the-world/.

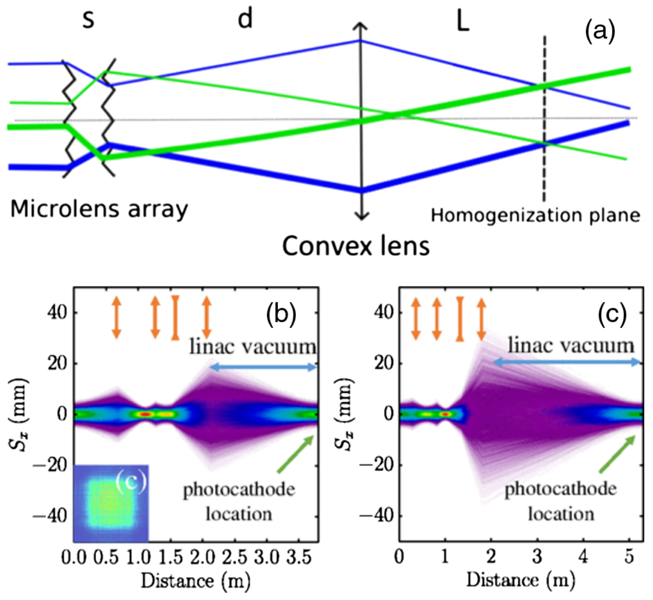

FIG. 7. Microlens-array- (MLA-) based homogenization system. (a) Diagram of several rays passing through the MLA pair onto the homogenization plane. (b),(c) Ray tracing from the homogenization plane through imaging optics and to the photocathode plane for two different systems. From Halavanau *et al.*, 2017.

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 of the beam 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 under active development by researchers (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 that 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 a SLM act as a polarization rotator with spatial dependence, which shapes the light when used with a polarizing beam splitter. The surface of the SLM is then $4f$ imaged ($f = 100$ mm lens pair) onto an intermediate plane to preserve the beam divergence, and this intermediate plane is then imaged with a single long focal length lens ($f = 750$ mm) onto either the photocathode or a CCD. An ultrahigh vacuum 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 works only with IR and visible light, while DM-based methods can work in the UV spectrum (Li *et al.*, 2017), thus avoiding the distortions in the upconversion process. In other words, this continues to be a lively area of research.

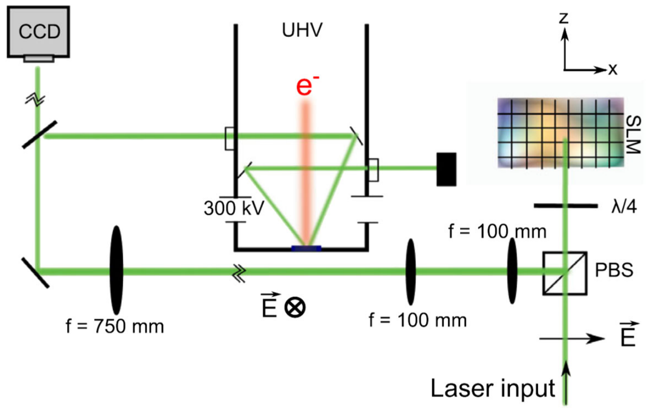

FIG. 8. Spatial light modulator (SLM) homogenization system. A SLM is actively adjusted via a computer (not shown) in a feedback loop to homogenize the laser profile on the photocathode by monitoring the profile at the virtual cathode location with a camera (CCD). From Maxson *et al.*, 2015.

D. Shaped distributions

Transverse bunch-shaping methods have been demonstrated for both FE and PE cathodes. In this section, 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 that we consider here. Nonetheless, because of 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 (such as triangular) and modulated (such as in an 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 ∼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 (such as in 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 that it degraded due to the nonuniformity of the emitters. Note that as the spacing of the emitters gets closer (such as in a nanoengineered FEA) (Graves *et al.*, 2012), it becomes more difficult to maintain the modulation.

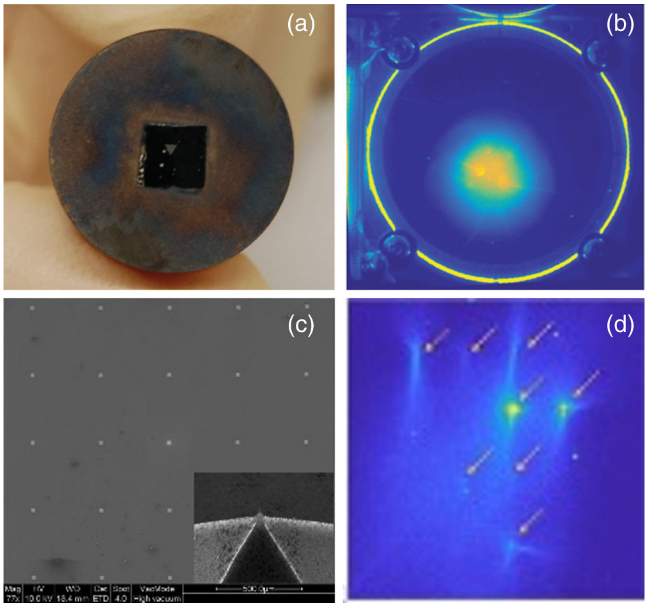

FIG. 9. FEA cathodes for transverse shaping. (a) Photograph of an FEA cathode with a triangular array inside a central square area of a round cathode plug and (b) its corresponding electron-bunch image on a downstream screen showing a slight triangular shape. (c) SEM image of rectangular array on a cathode plug with diamond tips (inset), and (d) its corresponding electron-bunch image on a downstream screen showing a grid-shaped pattern. From Andrews *et al.*, 2020, and Nichols *et al.*, 2020.

To reproduce the initial source modulation out of the gun, the charge must be kept low to avoid space-charge dilution, and the FEA must be gated to keep the 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 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 [Figs. 10(a) and 10(b)] or hollow beams [Fig. 10(c)]. To date these patterns have been generated only with passive optical systems in PE cathode guns based on an optical mask (Rihaoui *et al.*, 2009; Wisniewski *et al.*, 2012) or a 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 used to generate a positive array of laser spots [Fig. 10(a)] and a painted quartz plate used to create a negative hollow laser ring [Fig. 10(c)]. In the latter case, a MLA system was used to create a large array of spots [Fig. 10(b)] by changing the location of the convex lens [Fig. 7(a)] as described by Halavanau *et al.* (2017).

3. PE-based shaped longitudinal distributions

As previously stated, owing to the difficulties in gating FE and TE cathodes on a picosecond (or 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 a laser pulse duration ranging from tens of femtoseconds to tens of picoseconds, while PE cathode response times range from the femtosecond scale for metallic cathodes to tens of picoseconds for some semiconductor cathodes (Dowell, Schmerge, and Lidia, 2008). For longitudinal-shaping applications via the electron gun, one chooses a PE cathode with a response time that is much less than 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 here and categorized as either frequency domain or time domain.

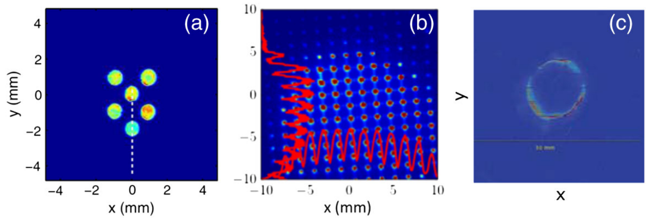

FIG. 10. Transversely shaped electron-bunch distributions from a PE cathode. An array of spots generated with (a) an optical mask and (b) a MLA. (c) A hollow beam generated with an optical mask. From Rihaoui *et al.*, 2009, Wisniewski *et al.*, 2012, and Halavanau *et al.*, 2017.

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 filters 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 megahertz 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 megahertz (Tournois, 1997; Verluise *et al.*, 2000).

In Fourier-transform pulse shaping, a $4f$ 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

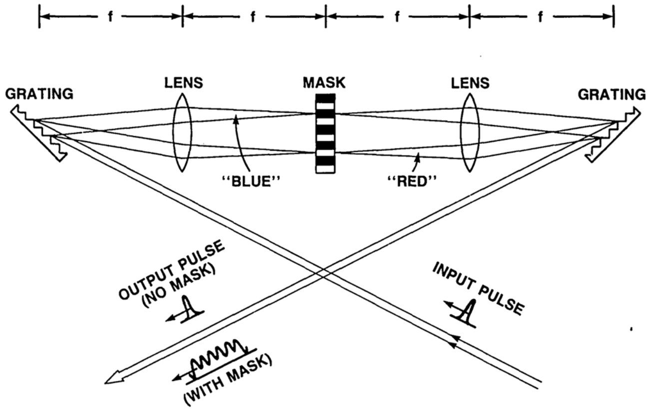

FIG. 11. Fourier-transform laser pulse-shaping system. An input laser pulse enters a $4f$ grating stretcher, and its frequency spectrum is exposed at the Fourier-transform plane. A mask described by a frequency-domain transfer function $H(\omega)$ is placed on this plane to modify the spectrum of the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

by $X(\omega)$ and the frequency-domain transfer function of the optical element at the Fourier plane be represented by $H(\omega)$. The spectrum of the output laser pulse is then given by their convolution

$$Y(\omega) = X(\omega)H(\omega) \tag{171}$$

such that the output laser pulse in the time domain is simply the following inverse Fourier transform of this convolution:

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

The first systems used for longitudinal shaping with a rf photocathode gun (Neumann *et al.*, 2003, 2009) 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 following: 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, Heritage, and Kirschner, 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 a 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 supertriangular laser pulses with a SLM-based system (Kuzmin *et al.*, 2019).

The advantage of the frequency-domain method is its high resolution and flexibility in producing various pulse 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 nonlinearities that cause 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, while another would be the development of PE cathodes that can respond to IR; this is an additional area of active research.

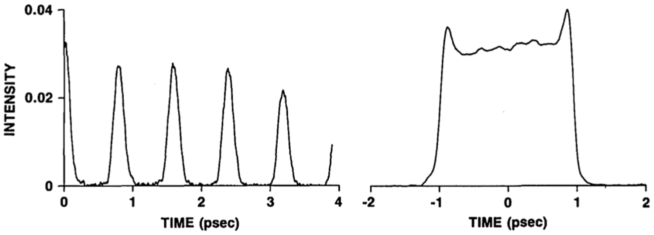

FIG. 12. Demonstrated laser pulse temporal shapes generated with the Fourier-transform laser pulse-shaping system of Fig. 11 using an amplitude mask. Laser pulse train (left panel) and quasiflattop laser pulse (right panel). From Weiner, Heritage, and Kirschner, 1988.

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 terahertz spacing with low loss. More recently pulse stacking with birefringent crystals has been used almost exclusively. It exploits the group velocity mismatch between the ordinary and extraordinary axes of a birefringent crystal [such as alpha barium borate (α -BBO)] (Zhou *et al.*, 2007; Will and Klemz, 2008; Power and Jing, 2009). 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

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

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 length controls the delay between the different polarizations, while the 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 (such as a double triangle) by passing the input laser pulse through a stack of crystals rotated at optimized angles relative to the incident laser pulse.

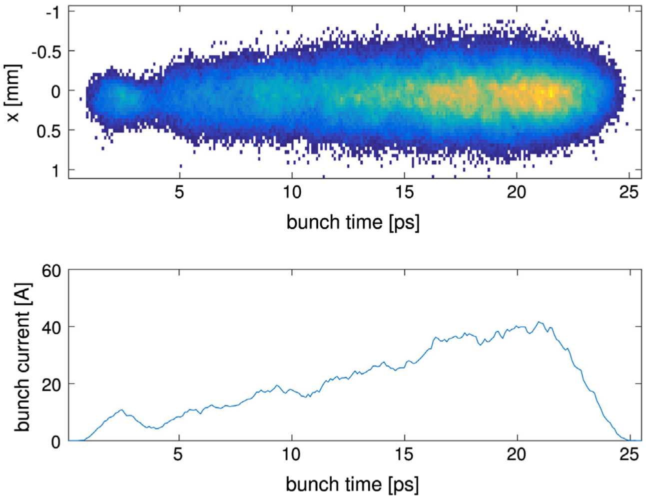

FIG. 13. Measured double-triangular electron bunch generated with α-BBO laser pulse stacking. Top panel: x-z projection of the bunch on a screen. Bottom panel: corresponding current profile. From Loisch *et al.*, 2018.

Loisch *et al.* (2018) and Liu *et al.* (2019) used this method to generate a double-triangular laser pulse that was then used in a 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 the shapes of the previously discussed "continuous" distributions (i.e., triangle and flat-top), birefringent crystals have also been used to produce bunch trains. Li and Kim (2008) proposed a simple and more compact system specifically adapted to bunch train generation. The setup consists of a concentric stairstep 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 such that the method is well adapted to the generation of pulses with picosecond-scale temporal separation. Li and Kim (2008) numerically investigated the technique for the formation of a train comprising 20- to 100-fs bunches with an associated bunching factor peaking at ~0.5 THz for application to a coherent Smith-Purcell terahertz source based on a 50-kV electron beam.

4. PE-based spatiotemporal (3D) shaping

A natural continuation of the progress with programmable 2D transverse and 1D longitudinal laser shaping are methods for shaping the complete 3D spatiotemporal 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 indicate 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. 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 are of interest. For example, one could combine a 2D circular transverse distribution (for low emittance) with a triangular 1D longitudinal distribution (for a high transformer ratio) to simultaneously achieve different objectives.

Recent progress in 3D laser shaping with IR laser pulses has been achieved by several groups (Mironov *et al.*, 2016; Kuzmin, Mironov *et al.*, 2019; Kuzmin *et al.*, 2020). For example, researchers used multiple programmable SLMs (Fig. 14) to generate both a 3D quasicylinder and 3D quasiellipsoidal 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 was used to introduce a complex spectral structure that, in

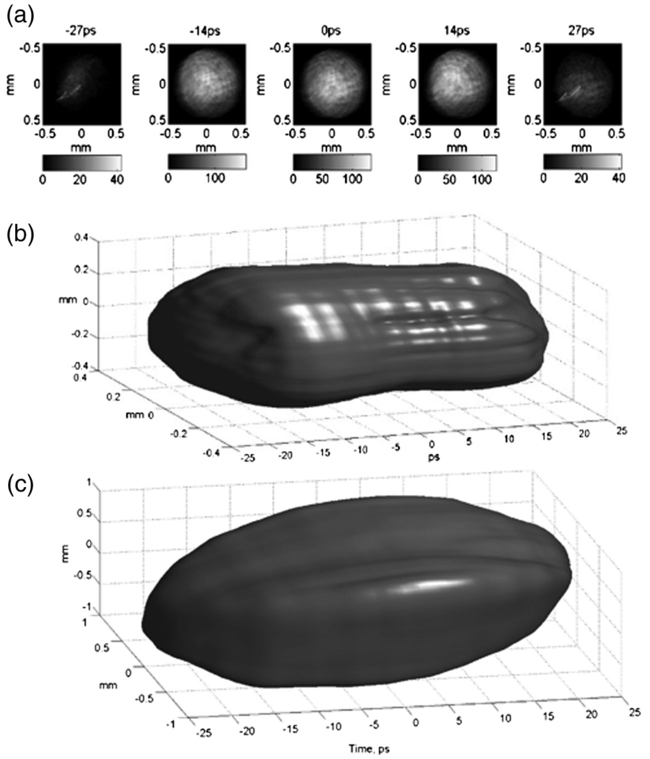

FIG. 14. Demonstrated 3D laser distributions based on SLMs. (a) Measured transverse intensity distributions of quasicylindrical pulses. (b) Reconstructed 3D quasicylindrical distribution. (c) Reconstructed 3D quasiellipsoidal distribution. From Mironov *et al.*, 2016.

combination with a highly dispersive section, resulted in the generation of a 3D ellipsoidal bunch (Liouville, 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 conversion of the IR pulse shapes to the UV, where they can be used to excite PE cathodes. Nonetheless, this is an exciting and important area of research. Further, note that the demand for 3D laser pulse shaping, and subsequent progress in that area, has been led by the PE cathode gun community, not the laser community.

IV. BEAM CONTROL WITHIN 1 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 the longitudinal or transverse direction 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 such that the overall transformation is still within 1 degree of freedom (d.o.f.) 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 d.o.f. is commonly employed to control the beam properties along any of the d.o.f.'s. For instance, time-dependent fields are commonly used to introduce 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 as 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 other d.o.f.'s, such as in the transverse plane to focus the beam.

Generally, this type of manipulation can be split into two stages referred to as the modulator and the converter. The modulator introduces a position-dependent momentum, whereas the converter consists of a beamline providing a momentum-dependent change in the position. For simplicity, we consider the case of uncoupled motion between the three d.o.f.'s and focus on one of the d.o.f.'s, 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

$$\begin{aligned} \zeta_{i,0} &\to \zeta_{i,m} = \zeta_{i,0}, \\ \zeta_{i+1,0} &\to \zeta_{i+1,m} = \zeta_{i+1,0} + F_{i+1}(\zeta_{i,0}), \end{aligned} \tag{174}$$

under an impulse approximation (where we ignore change in the particle position). In the latter equation of Eqs. (174), 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 such that the final position $\zeta_{i,f}$ is related to the upstream coordinate as

$$\begin{aligned} \zeta_{i,m} &\to \zeta_{i,f} = \zeta_{i,m} + G_i(\zeta_{i+1,m}), \\ \zeta_{i+1,m} &\to \zeta_{i+1,f} = \zeta_{i+1,m}. \end{aligned} \tag{175}$$

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

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

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 (Sec. II.C.1) $\Phi_f(\zeta_{i,f}, \zeta_{i+1,f}) d\zeta_{i,f} d\zeta_{i+1,f} = \Phi_0(\zeta_{i,0}, \zeta_{i+1,0}) \times d\zeta_{i,0} d\zeta_{i+1,0}$, we can write the final phase-space distribution as

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

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

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

By virtue of the charge conservation we now have

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

where the rhs can be written solely in terms 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 Sec. 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, using $F(z) = hz$ and $G(\delta) = R_{56}\delta$. Here for simplicity we take both F and G to be linear functions of s such that the overall transformation is described using

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

Therefore, 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 $\lambda_0(z_0) = \left(1/\sqrt{2\pi\sigma_{z,0}^2}\right)\exp[-z_0^2/(2\sigma_{z,0}^2)]$ and assuming that δ_0 represents a random uncorrelated fraction energy spread, we obtain

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

which reflects the well-known results associated with the linearized bunch compression that was mentioned in Sec. 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 of the form (Li, 2001)

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

where $\Theta(\cdots)$ is the Heaviside 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 magnetic compression subjected to strong nonlinearities; see Dohlus *et al.* (2004).

The examples considered thus 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 nonlinearity 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 Sec. II.C.5. This is a typical process in optics where the energy-dependent path length introduced by a dispersive section, such as the previously discussed bunch compressor, is now written as $\zeta_{5,f} = \sum_{i=1}^{6} R_{5i}\zeta_{i,0} + \sum_{i=1}^{6}\sum_{j\geq 1}^{6} T_{5ij}\zeta_{i,0}\zeta_{j,0} + \mathcal{O}(\zeta_0^3)$, and likewise for any beamline; see Eq. (120). For simplicity, we assume a pencil bunch such 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, such as the limited set of klystron frequencies.

1. Interceptive beam shaping

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

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

where we assume the mask to be thin and to affect the beam according to the transverse coordinate (and not the momentum). Consequently, the transverse profile along ζ_1 can be determined from

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

where we have taken the mask to be a binary function with a 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 (such as via a function similar to F or G). Finally, there is 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 in which a collimator is used to tailor the energy distribution (such as to remove the 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 cw electron sources where a combination of masks and rf cavities (often called "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 FEL (Nguyen and Carlsten, 1996) applications. The method was eventually demonstrated at the Accelerator Test Facility (ATF) (Muggli *et al.*, 2008), where it was also extended to shaping beyond microbunch generation (Shchegolkov *et al.*, 2015). 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 a 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 are

$$\begin{aligned} x_m &= x_\beta + \eta\delta_0, \\ z_m &= z_0 + \xi_-\delta_0, \end{aligned} \tag{186}$$

where x_β is the geometric contribution to the beam spot. An intercepting mask in (x_m, y_m) with a 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 the longitudinal dispersion ξ_+ 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 (Weiner, Heritage, and Kirschner, 1988; Kuzmin *et al.*, 2019); see Sec. III.D.3.a.

This method was implemented by Muggli *et al.* (2008) to form a train of a subpicosecond bunch. In such a case the mask consists of a set of N slits such that the transmission can be given 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 the interslit spacing on the mask, the final modulation period is shown to be (Muggli *et al.*, 2008; Hyun *et al.*, 2019)

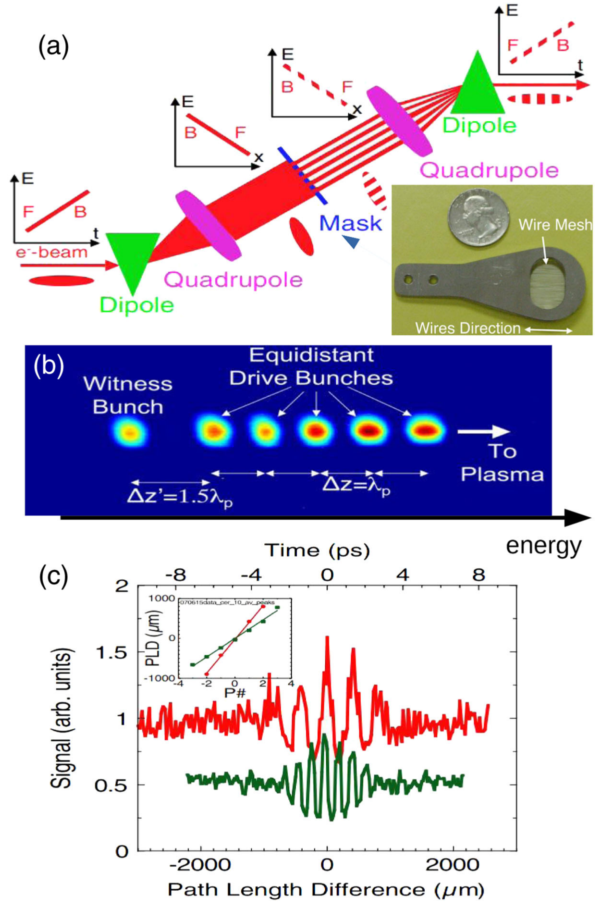

FIG. 15. Experimental generation of a subpicosecond bunch train at the ATF using (a) a dogleg beamline with final beam distribution (b) in a downstream spectrometer and (c) a measured temporal distribution via interferometry of coherent transition radiation. Adapted from Muggli *et al.*, 2008, and Muggli, Allen, Yakimenko, Park *et al.*, 2010.

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

A practical implementation of the method used for a proof-of-principle experiment at the ATF appears in Fig. 15(a). The beam was locally dispersed in a dogleg beamline with a vanishing net dispersion. The dogleg included an optical lattice providing an antisymmetric dispersion function with maximum value attaining $\eta \simeq 1.5$ m in proximity to the mask location. The mask consisted of submillimeter tungsten wire; see the inset 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 the 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 submillimeter periods were produced, which is consistent with the resonant excitation of wakefield in plasmas (Muggli, Allen, Yakimenko, Fedurin *et al.*, 2010). Likewise, this method was extended to produce bunches with triangular beam distribution at the ATF (Shchegolkov *et al.*, 2015) for wakefield excitation in dielectric-lined waveguides (Antipov *et al.*, 2012).

More recently a current-shaping method based on the direct introduction of a spatiotemporal coupling was proposed using a pair of transverse deflecting cavities (Zholents *et al.*, 1999; Kur *et al.*, 2009). 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 a rf linac (Wilson *et al.*, 1985).

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

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

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

$$\begin{aligned} m_{12} &= -\frac{L[\kappa_d + \kappa_u(Lm_{21} + m_{11} + 2m_{22})]}{2\kappa_u}, \\ \kappa_d &= -\frac{\kappa_u(Lm_{21} + 2m_{22})}{2}, \\ m_{11} &= -\frac{Lm_{21}}{2} - \frac{\kappa_u}{\kappa_d}. \end{aligned} \tag{189}$$

The matching conditions in Eqs. (189) ensure that 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

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

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

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

which is consistent with the results of Kur *et al.* (2009), where the choice $m_{11} = m_{22} = -1$ was made such that $\kappa_u = \kappa_d$ (given that the 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 that 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. Application of this method to produce a modulated relativistic electron bunch was first discussed by Du, Huang, and Tang (2012) in the context of superradiant terahertz radiation and was further investigated numerically by Ha *et al.* (2020) to support the formation of various current shapes to support a beam-driven wakefield accelerator with an 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 [Figs. 16(a) and 16(b)] for beam-driven acceleration and subpicosecond bunch trains with variable microbunch spacing [Figs. 16(c) and 16(d)].

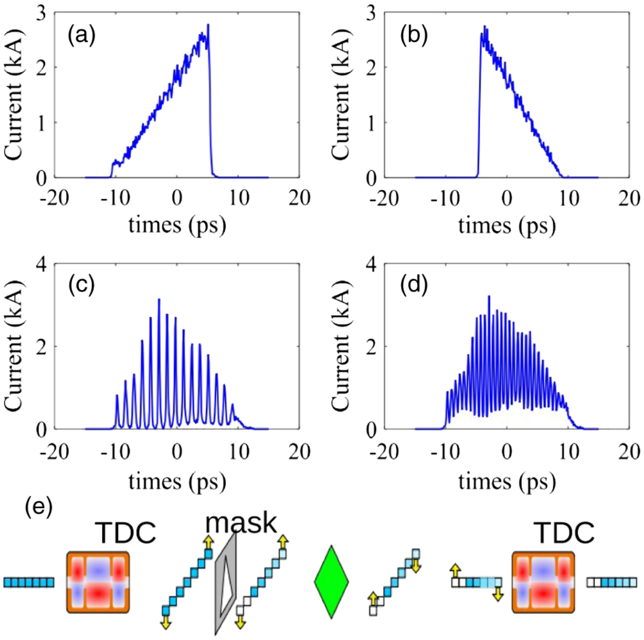

FIG. 16. Numerical modeling of current-profile shaping showcasing the generation of (a) ramped, (b) reverse-triangle, and (c), (d) modulated bunch distributions using local correlations (e) imparted by a pair of transverse deflecting cavities (TDCs) combined with a mask. Adapted from Ha *et al.*, 2020.

The masks described thus 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 such an implementation uses a slotted foil at the SLAC 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 suffers 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 $V(z) = \hat{V}\cos(kz+\psi)$, where $\hat{V}$, ψ, and k are, respectively, the peak accelerating voltage and phase and the accelerating-mode wave vector. The LPS transformation associated with the downstream dispersive sections is described via the Taylor expansion of the longitudinal map with the first- and second-order coefficients $R_{56} \equiv \partial z/\partial\delta$ and $T_{566} \equiv \partial^2 z/\partial\delta^2$, as detailed in Sec. II.C.5. The two widely used approaches are to either (i) introduce a linear energy modulation and control the nonlinear dispersion in the dispersive section or (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 by 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

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

(where $\bar{E}$ is the final beam energy), resulting in significant quadratic correlation in the LPS with an 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 ~12-MeV bunches with linearly ramped current profiles (England, Rosenzweig, and Travish, 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 consists of truncating the head and tail to suppress the current spikes and improve the FEL lasing performances. It was shown that current spikes can be strongly suppressed via control of the $U_{5666} \equiv (1/6)\partial^3 z/\partial\delta^3$ 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 the total accelerating voltage $V(z) = V_1\cos(k_1 z + \psi_1) + V_n\cos(k_n z + \psi_n)$, where $V_{1,n}$ 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 $a_l \equiv a_0 - e(k_1 V_1 \sin\psi_1 + k_n V_n \sin\psi_n)/\bar{E}_l$ and $b_l \equiv b_0 - e(k_1^2 V_1 \cos\psi_1 + k_n^2 V_n \cos\psi_n)/2\bar{E}_l$, with e 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

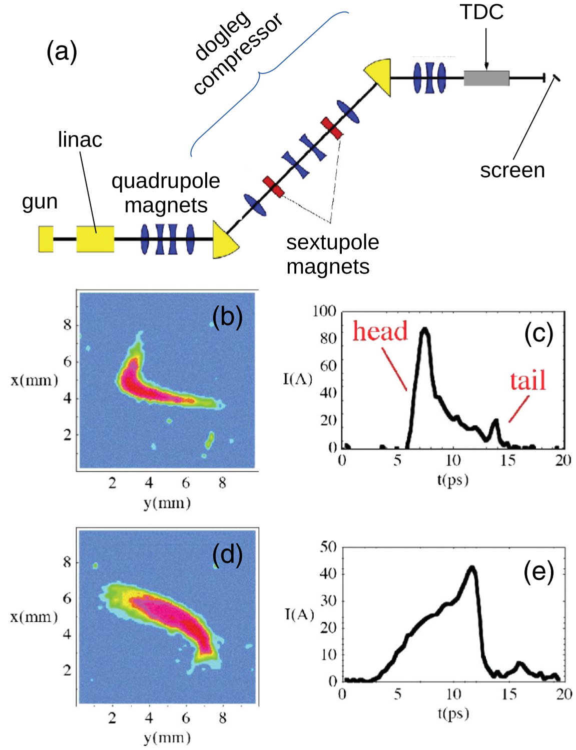

FIG. 17. (a) Experimental setup for phase-space shaping via introduction of nonlinear longitudinal dispersion at the NEPTUNE facility and (b),(d) the resulting LPS with (c),(e) the associated current distribution experimentally generated. Adapted from England, Rosenzweig, and Travish, 2008.

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

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

An example of the previously discussed method was experimentally implemented in the Free-Electron Laser in Hamburg (FLASH) facility at DESY (Piot *et al.*, 2012), where a linac composed of 1.3- and 3.9-GHz superconducting-linac modules and two magnetic bunch compressors was used; see Fig. 18(a). Specifically, the accelerator settings were

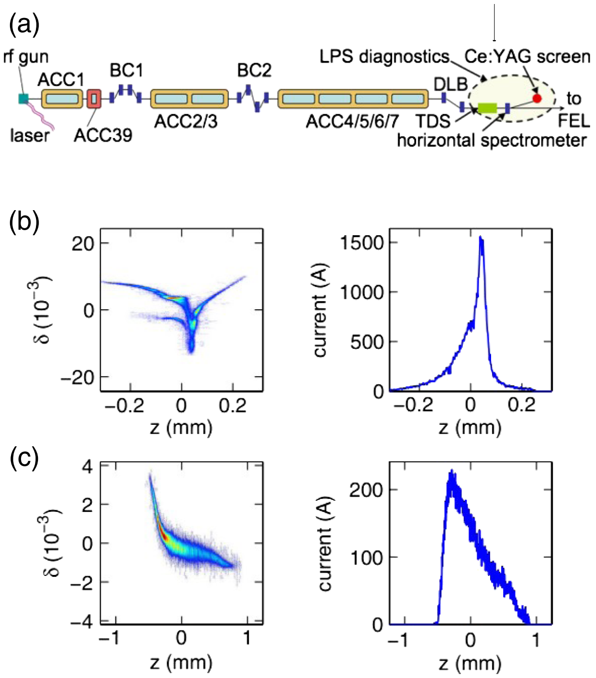

FIG. 18. (a) Experimental results on phase-space shaping using a dual-frequency linac at the FLASH facility at DESY with (b) corresponding nominally compressed beam and (c) linearly ramped bunch generation obtained via proper control of the accelerating-voltage phase and amplitude of the linac sections ACC1 (1.3 GHz) and ACC39 (3.9 GHz) prior to the bunch compressor 1 (BC1). Adapted from Piot *et al.*, 2012.

optimized to form a linearly ramped current profile; see Figs. 18(b) and 18(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 FLASHForward 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 the prohibitive associated cost. However, it could be implemented passively, as discussed in Sec. V.

As a final note, we point out that most of the time current shaping relies on an integration of different shaping techniques, as explored by Cornacchia *et al.* (2006) and 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) demonstrated the generation of uniform current distribution based on numerical simulations; see Figs. 19(a) and 19(b). Likewise, numerical simulations indicate that a properly shaped photocathode laser pulse combined with a multifrequency linac and a two-stage nonlinear compression process could provide precise control over the LPS and current distribution (Tan, Piot, and Zholents, 2021), ultimately realizing the sought-after doorstep distribution without the requirement of any collimation; see Figs. 19(c) and 19(d).

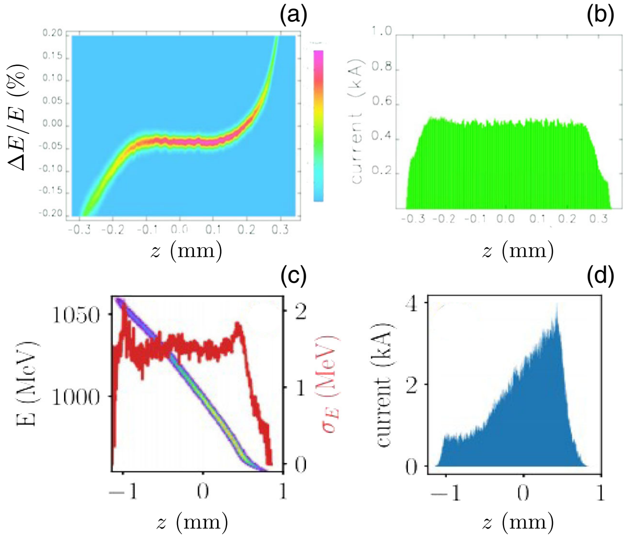

FIG. 19. Examples of (a),(c) LPS and (b),(d) associated current distribution obtained via a full integration combining the photocathode laser-shaping method discussed in Sec. III with multifrequency linacs and nonlinear dispersive sections. In (c), the superimposed trace corresponds to the slide rms energy spread (right label). (a),(b) From Cornacchia *et al.*, 2006. (c),(d) From Tan, Piot, and Zholents, 2021

C. Realizing ultralow energy spread

A recurrent topic associated with the use of bright electron beams regards the generation of ultralow longitudinal emittance. One application of such a capability is the production of a low-energy-spread beam, ultrashort electron bunch for use in ultrafast electron-scattering experiments. There has been a renewed interest in meV energy-spread, sub-MeV electron bunches for applications in electron energy-loss spectroscopy (Egerton, 2009).

In principle, the previously described method, which was detailed by Zeitler, Flöttmann, and Grüner (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, Muller, and Maxson, 2020) that demonstrated via numerical simulation the ability to produce a 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 fractional energy spread $\sigma_{\delta,f}$ under Liouville's theorem as $\sigma_{\delta,f} = \sigma_{\delta,0}(\sigma_{z,0}/\sigma_{z,f}) \geq \hbar/2$, where the subscript 0 refers to the initial bunch length and fractional energy spread. Duncan, Muller, and Maxson (2020) noted 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 for by the pair of accelerating cavities.

D. Controlling LPS correlations

As discussed, 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 a more monoenergetic bunch.

Although linear energy-spread control is conventionally achieved using an accelerating cavity operated off the crest, the method introduced by Yampolsky, Simakov, and Malyzhenkov (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 a longitudinal 2×2 block given by

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

where L_c is the transverse deflecting cavity length; see Fig. 20. Equation (194) 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, Simakov, and Malyzhenkov, 2020).

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 nonrelativistic 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 by Smith (1986) and later adapted for the LCLS and FLASH FELs (Flöttmann, Limberg, and Piot, 2001). The method relies on introducing a harmonic frequency such that the final fractional energy offset associated with an electron located at z is

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

where δ_0 is the initial relative energy offset, ψ_1 and V_1 (ψ_n and V_n) are, respectively, the phase and amplitude of the fundamental (harmonic) section of the linac, and E_f is 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 $V_n = V_1(k_1/k_n)^2$. Such an approach was first demonstrated at the FLASH facility at 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 displays a measurement of the LPS-linearization process at the LCLS where a fourth-harmonic cavity ($f_n = 11.424\ \text{GHz} = 4 \times 2.856$ 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 nonultrarelativistic bunches. Such a scheme was introduced by Krafft (1996), who showed via numerical simulations and experimentally confirmed that proper control of the phase and amplitude of a buncher and that capture cavities could control the R_{56} and T_{566} in the Continuous Electron Beam Accelerator Facility injector (Wang, Krafft, and Sinclair, 1998). A similar technique was explored theoretically and via simulation for an ultrafast electron source using a rf photoinjector; see Zeitler, Flöttmann, and Grüner (2015). Specifically, they showed 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

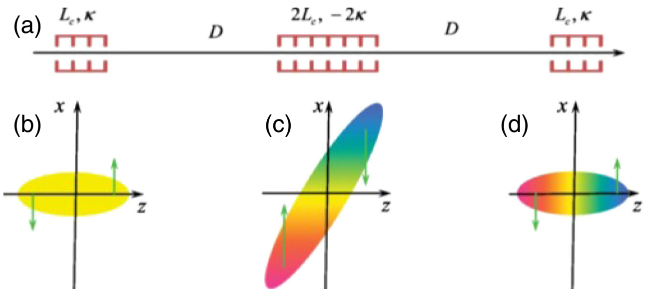

FIG. 20. (a) A TDC-based beamline for the generation of strong LPS chirps. (b)–(d) The $x - y$ electron-bunch distribution inside each TDC with the arrows gives the direction of the transverse kick x'. The color map indicates the energy distribution. Adapted from Yampolsky, Simakov, and Malyzhenkov, 2020.

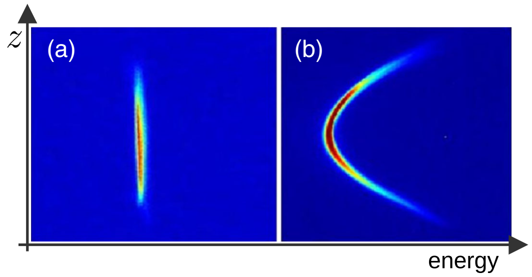

FIG. 21. Example of longitudinal phase-space linearization at the LCLS facility. The LPS measurement (a) without and (b) with operation of the fourth-harmonic linearizing cavity (operating at 11.424 GHz). Adapted from Akre *et al.*, 2008.

the electron gun. Although the method was implemented as part of a ballistic-compression scheme to demonstrate the generation of subfemtosecond 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 in Sec. IV.B, can also be used to control LPS nonlinearity. Such a method was successfully implemented in a high-power energy-recovery linac (Piot, Douglas, and Krafft, 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 a FEL oscillator (Neil *et al.*, 2000); see Fig. 22. Higher-order correction using octupole magnets was also implemented in the Jefferson Lab (JLab) 10-kW FEL facility (Neil *et al.*, 2006).

Finally, a laser pulse shaping similar to the one discussed in Secs. II and IV.B that was initially proposed by 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 for 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, Penco *et al.* (2014) demonstrated that the beam dynamics of the ramped bunch did not significantly affect the beam transverse emittance compared to their nominal operating point.

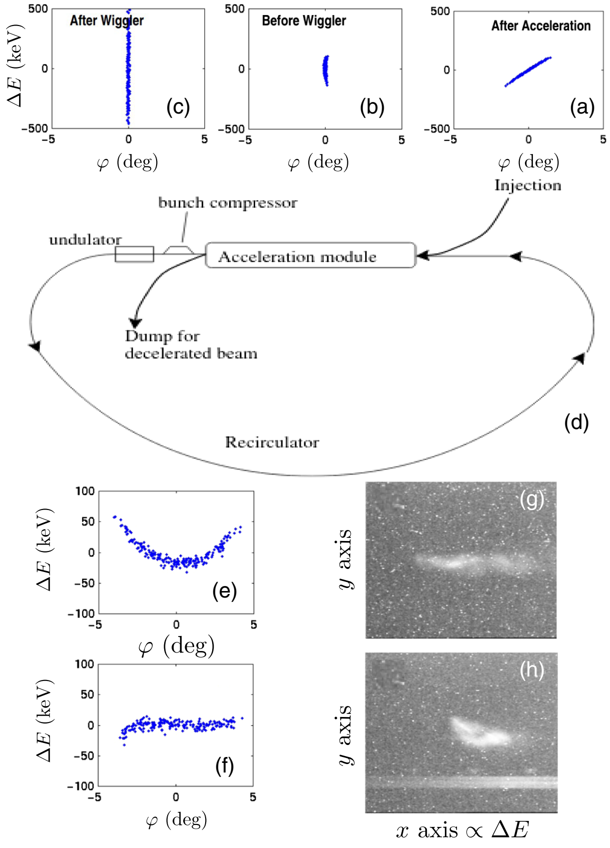

FIG. 22. Nonlinear control of LPS with high-order longitudinal dispersion in JLab IR-FEL Demo for two settings of the longitudinal lattice along with resulting measured energy spread measured after energy recovery [(g) and (h)]. The simulated LPS is displayed (a) at acceleration-module exit, (b) before and (c) after the FEL iteration, and after energy recovery for the cases of (e) linear and (f) nonlinear correction (via independent control of the R_{56} and T_{566} of the recirculator) of the LPS. The measured $(\delta E, y)$ distributions in (g) and (h) correspond, respectively, to simulated LPS presented in (e) and (f). Adapted from Piot, Douglas, and Krafft, 2003.

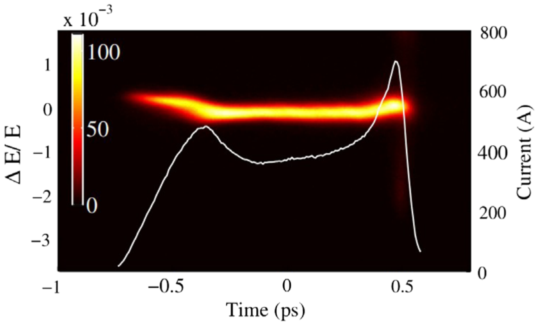

FIG. 23. Experimental demonstration of control of LPS correlation at the FERMI@ELETTRA accelerator using a 1.4-GeV beam. Electron-bunch LPS (image) with superimposed current profile (white line, right axis). These data were obtained using a shaped photocathode laser with a ramped temporal profile configuration. 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 to produce x rays, as it provides a uniform dose and mitigates thermomechanical 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, Wu *et al.*, 2014). 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 Sec. 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 by Merminga *et al.* (1991) to remove the beam tail via nonlinear focusing (implemented with sextupole magnets) as part of a "nonlinear collimation" scheme proposed for future linear colliders. Likewise, the use

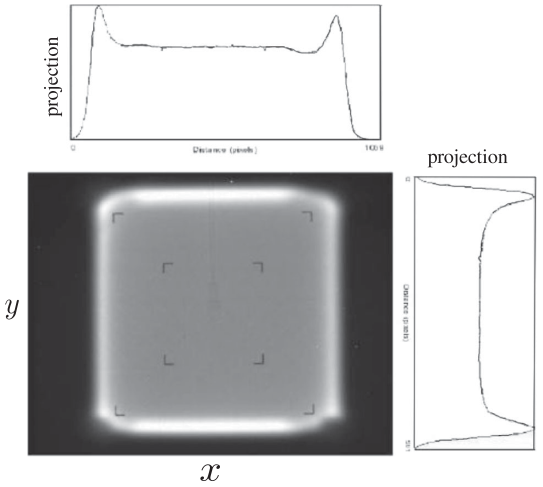

FIG. 24. Measured beam distribution obtained downstream of a uniformizing beamline composed of octupole magnets. Adapted from Tsoupas *et al.*, 2007.

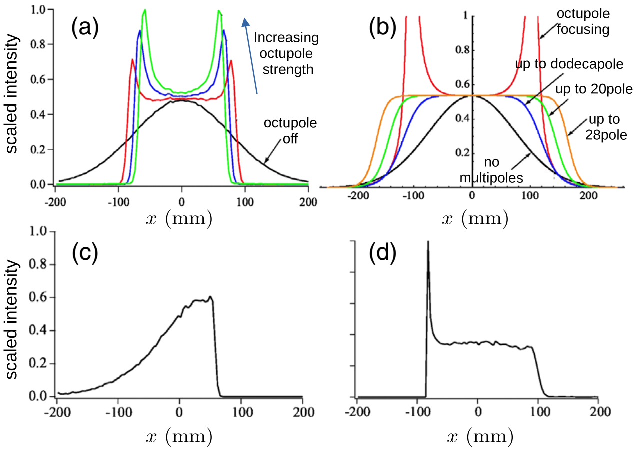

FIG. 25. Example of computed transverse profiles downstream of a nonlinear-transport beamline composed of multipole field. The upper row demonstrates (a) the impact of an octupole magnet with increasing strength and (b) the effects associated with various multipole orders on the horizontal beam distribution. The lower row displays the use of two sextupole magnets to control the beam distribution. (c),(d) distributions recorded at two locations downstream of the sextupole magnets. Adapted from Yuri *et al.*, 2007.

of nonlinear focusing to form uniform beam distributions was discussed by Meads (1983). Numerical simulations (Kashy and Sherrill, 1987) demonstrated the formation of a uniform distribution using an octupole magnet implemented in a practical beamline. Likewise, Batygin (1993) derived 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 by Meot and Aniel (1996) analytically showed that the odd-order multipole fields, such as octupole and dodecapole components, are required for transverse uniformization. Efforts to produce transversely uniform distributions over appreciable distances along an accelerator beamline have been investigated for possible use in combination with undulator tapering in x-rays 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 Brookhaven National Laboratory (BNL) (Tsoupas *et al.*, 1991), and an example of distribution measured at the NASA Space Radiation Laboratory, using an electron beam at BNL (Tsoupas *et al.*, 2007), appears in Fig. 24. Similarly, uniform beam distributions are sometimes used in medical accelerators.

Yuri *et al.* (2007) developed 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 of the following 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 Sec. II.C.1 indicates that the final distribution along one of the position coordinates can be written as

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

Equation (196) provides the values of the magnetic multipole strength K_{2n} required to transform the initial distribution into the desired one. Yuri *et al.* (2007) also extended 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. They concluded that using a combination of even-order multipoles may ultimately provide a 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 a 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. Further development demonstrated the generation of patterned hollow beams (Yuri, Fukuda, and Yuyama, 2019). The latter capabilities could be taken advantage of and combined with phase-space-exchange methods discussed in Sec. VI.

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 by Ha, Conde, Power, Shao, and Wisniewski (2019) and illustrated in Fig. 26. When we take the example of the horizontal phase space (x, x') passing through a set of transverse wigglers (a wiggler oriented transversely to the beam direction) with wiggler parameter $K_{w,i}$, the divergence of an electron at initial position x_0 will be given by

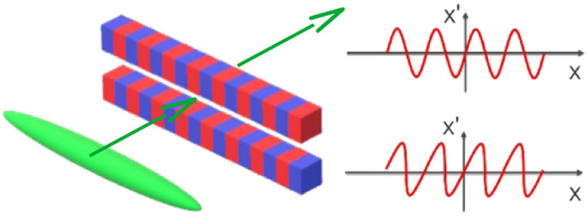

FIG. 26. Configuration for transverse beam shaping using a "transverse"-wiggler magnet (left) and horizontal phase space (right) immediately downstream (upper plot) the wiggler and after a subsequent drift (lower plot). Adapted from Ha, Conde, and Power, 2019.

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

where $k_i \equiv 2\pi/\lambda_i$ is the transverse wave vector with λ_i 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 Sec. IV.B, an interceptive "binary" mask with an 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 radiological activation and damage of beamline hardware.

For ultralow emittance beams where quantum coherence is achieved and the beam can be described by its wave function, optical photon techniques have been adapted to shape the beam with a high degree of control, such as 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 nanoengineered spiral-like phase plate made of stacked graphite thin film.

Although electron beams produced in conventional accelerators are incoherent, ultralow-emittance relativistic electron beams can be manipulated with variable-transmission masks, as discussed by Nanni, Graves, and Moncton (2018), provided that their transverse coherent $L_\perp = (\hbar/mc)\sigma_\perp/\varepsilon_\perp$ (van Oudheusden *et al.*, 2007) is comparable to or larger than the intercepting-mask crystalline structure (where $\sigma_\perp$ and $\varepsilon_\perp$, respectively, refer to the transverse beam size and emittance). van Oudheusden *et al.* (2007) proposed a diffraction contrast modulation technique to realize nanopatterned 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 the 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 [Figs. 27(c) and 27(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 submicrometric spatial period combined with demagnifying optics could produce beams with nanometric modulations.

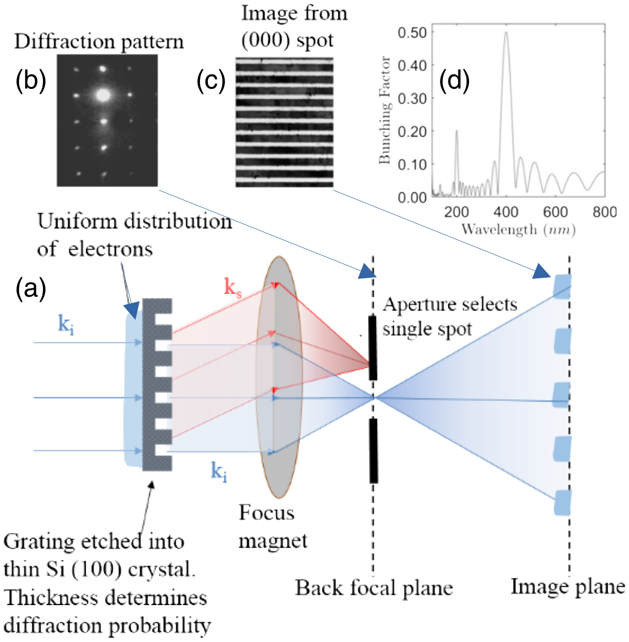

FIG. 27. Principle of diffraction modulation to produce (a) a nanopatterned beam with (b) a diffraction pattern downstream of the grating and (c) a final transverse distribution of the transmitted beam with (d) an associated spatial spectrum along the vertical coordinate. Adapted from Graves *et al.*, 2019.

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 nonlinear correlations on the beam. The nonlinearity on the beam's phase space not only increases the emittance but also introduces 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 here 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 Secs. V.A and V.B. We then describe the mechanism using wakefields in Sec. V.C. Each section provides a description of the results and their applications.

A. Shaping profiles using the 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. Sections V.A.1–V.A.5 describe these shaping methods and give the relevant results.

1. Space-charge field with a single bunch

Owing 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.

First, the space-charge force applied to the bunch is not linear, except for a few profiles. Thus, as with 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

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

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 similar to octupole-based transverse shaping; see Fig. 24. Thus, if the beam traverses a long enough distance (i.e., at large enough R_{56}), then this Gaussian distribution will evolve into a uniform distribution. Figure 28(b) shows a particle tracking simulation result that demonstrates the idea. As we can see in 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, to date this type of shaping has not been experimentally demonstrated.

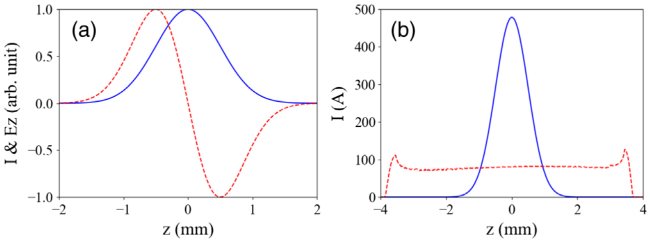

FIG. 28. Current-profile shaping using space-charge force. (a) Gaussian longitudinal profile (solid line) and corresponding longitudinal space-charge field from Eq. (198). (b) Longitudinal profile of the initial particle distribution (solid line), and the longitudinal profile after a 3-m-long drift (dashed line). (b) is a simulation result of a beam with 1-nC charge and 3-MeV energy.

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 an understanding of this mechanism may help to improve the shaping quality of the emission-based shaping methods introduced in Sec. III.

The other usage is exploiting the symmetric lengthening feature of the space-charge field. Luiten *et al.* (2004) suggested a new concept, the so-called blowout regime, to generate a 3D ellipsoidal distribution with uniform density. In this concept, an ultrashort 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 ultrathin sheet of electrons that evolves into a uniform ellipsoidal shape. Here the ultrashort pulse length plays a key role because it allows 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. Luiten *et al.* (2004) 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 a 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.

Rosenzweig *et al.* (2006) performed further theoretical work combining the blowout regime and emittance compensation. Later this blowout regime was experimentally demonstrated by Musumeci *et al.* (2008), O'Shea *et al.* (2011), and Piot *et al.* (2013). The first experiment of Musumeci *et al.* (2008) demonstrated the blowout regime using a 15-pC beam; see Fig. 29. Later the experiment of Piot *et al.* (2013) demonstrated the blowout 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 the target charge. This limits the intrinsic emittance, which scales with laser spot size [Eq. (168)]. The emittance demonstrated thus far does not show a clear advantage for this method. There is another emission scheme, based on a cigar beam and introduced by Rosenzweig *et al.* (2019), that

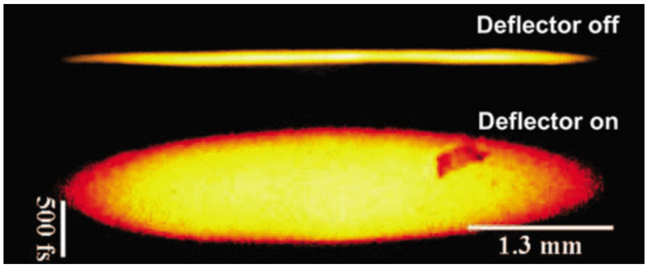

FIG. 29. Experimental demonstration of ellipsoidal beam generation using the blowout method. The beam images were measured after a deflecting cavity. The beam was focused vertically, and the deflecting cavity kicked the beam vertically. The x and y axes correspond to the beam's horizontal and longitudinal distributions, respectively. The beam's energy was 3.75 MeV, and the charge was 15 pC. From Musumeci *et al.*, 2008.

achieved lower emittance. While the blowout regime uses a short but large radius laser profile that produces a longitudinal expansion of the beam, the cigar-beam regime uses a long but small radius laser profile that produces radial expansion of the beam. Simulations showed a noticeable improvement in beam brightness of the cigar compared to the blowout regime. Because of these limitations, research efforts were shifted to using ellipsoidal laser profiles instead of the blowout regime to generate an ellipsoidal beam; see Khojoyan, Krasilnikov, and Stephan (2014). See Sec. III for further details on laser shaping.

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.

Lu *et al.* (2018) used this concept 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 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 in the middle so that the nonlinearity of the space-charge field coming from the field-shaping bunches would almost cancel out. This provided a nearly linear longitudinal chirp to the target bunch; the relative separation was adjusted using a laser delay line. The laser was injected into the rf photocathode gun, which accelerated the beam to about 3.4 MeV. Because of 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 through the velocity difference of the particles in a drift. Velocity bunching is a more general term, but ballistic bunching is used when acceleration is not included in the compression.

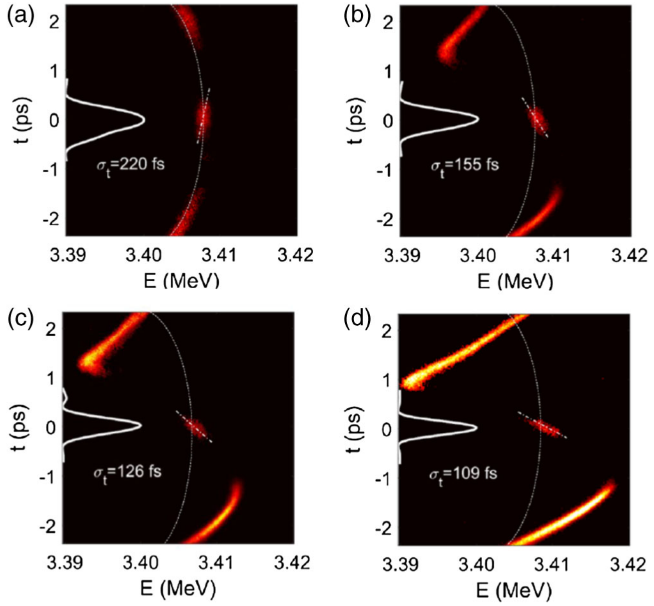

FIG. 30. Space-charge field shaping for bunch compression. Measured longitudinal phase spaces are displayed. Each panel corresponds to a different charge level of field-shaping bunches. (a)–(d) Charges of 0.2, 3.2, 4.5, and 6.7 pC, respectively. $t > 0$ corresponds to the head. From Lu *et al.*, 2018.

During the experiment, they fixed the target bunch's charge level to 50 fC while varying the field-shaping 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 a LPS measurement, as shown in Fig. 30. In addition, the main bunch length was controlled from 220 to 109 fs by varying the charge. As we see in Fig. 30(d), there is 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 results in an increase of the bunch-length jitter and arrival time jitter. On the other hand, all three bunches originate from a single laser pulse, and there are no rf cavities other than the gun in this scheme. Thus, this approach can provide better stability than other methods. This is an interesting method for applications requiring a short bunch with good stability such as ultrafast electron diffraction, ultrafast electron microscopy, and FELs.

A similar method was experimentally demonstrated for transverse phase space. Rihaoui *et al.* (2009) introduced extra bunches surrounding a central main bunch in the transverse space; see Sec. III.D.2. Because of 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 other than a Gaussian distribution. This could be an interesting option for a system in which it is difficult to use 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 one 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 applying this method to more general bunch-shaping applications. For example, nonlinearity is a key for asymmetric bunch shaping, as we saw in Sec. IV. We can imagine exploiting nonlinearities that space-charge force introduces. Bunch charges and spacing can be used to control the shape of the LPS. The space-charge field of the main bunch may be controlled using additional bunches, which may provide a new way to apply a laser-shaping technique to high-charge bunches.

We note that beam shaping with extra bunches always causes a problem by 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 thus far. A beam cutting process using a mask and an additional beamline element such as chicane or rf deflecting cavities may occur. However, these beamlines can easily add extra timing jitter to the beam, thereby 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 microbunch due to the space-charge force, so 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 microbunch continue to move in the same direction due to their momentum. These particles build another density modulation whose phase is 180° from the original modulation; thus, the direction of the space-charge force is then 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 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

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

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

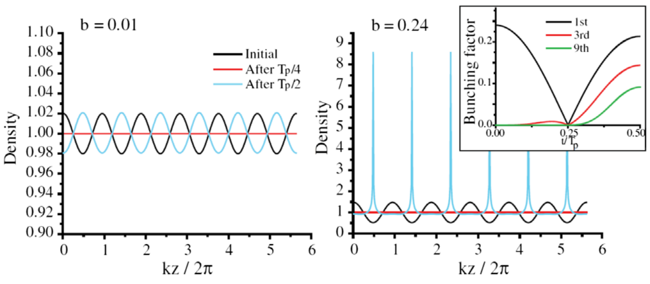

FIG. 31. Nonlinear plasma oscillation of the beam for generating a bunch train with a high bunching factor. The beam's longitudinal density profiles and their evolution are displayed, with bunching factors starting at 0.01 (left panel) and 0.24 (right panel). These are numerical solutions of the system of equations describing a coasting beam's density profile evolution. From Musumeci, Li, and Marinelli, 2011.

In the previously describe linear oscillation situation, a sinelike shape on the phase space does not break and keeps oscillating because the energy modulation induced by the initial density modulation is so small 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 large 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 becomes 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 in the following equation showing the electron density [more details were provided by Musumeci et al. (2013)]:

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

where $c_m(t) = [(-1)^{m+1}/m][2/\alpha(t)]J_m[m\alpha(t)b]$ 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 (Neumann et al., 2003, 2009; Harris et al., 2007). Here we describe more details of the experiment done by Musumeci, Li, and Marinelli (2011), who 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 ($\lambda = 300\ \mu$m) pulse train. They tested the bunch charge up to 40 pC, and a gun solenoid was used to focus the beam. To demonstrate the method, they varied the laser intensity, which controlled 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

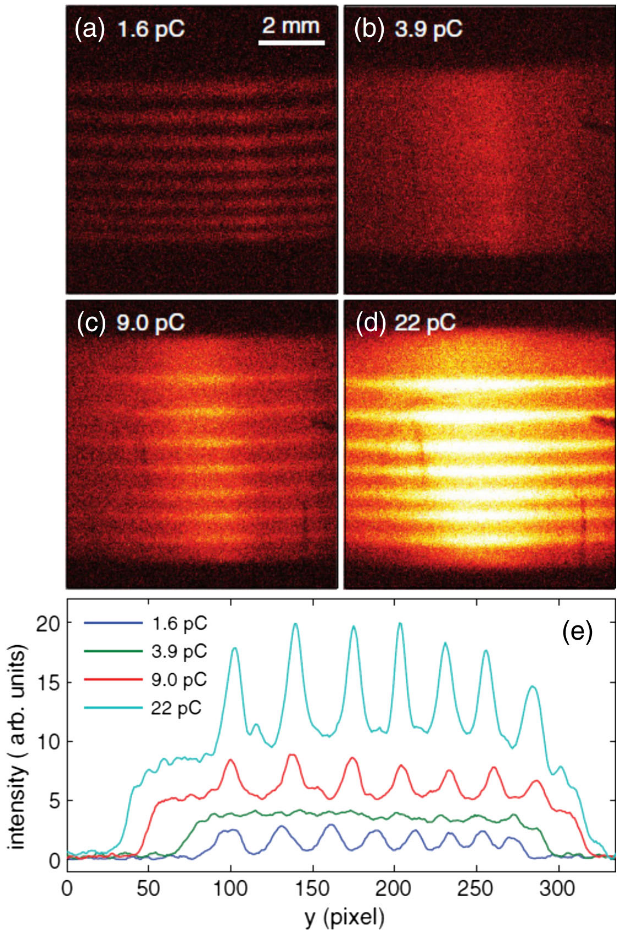

FIG. 32. Experimental demonstration of bunch train generation using nonlinear plasma oscillation. (a)–(d) Temporally streaked beam images. (e) The corresponding longitudinal profiles. The streaking direction of the beam images is vertical. From Musumeci, Li, and Marinelli, 2011.

gun phase was set to 35° to preserve the original modulation period. The results are summarized in Fig. 32.

According to the simulation of Musumeci, Li, and Marinelli (2011), 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 then 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 with at a higher charge level to consider an application to terahertz radiation. Zhang *et al.* (2016) set up an experimental beamline similar to that of Musumeci, Li, and Marinelli (2011) and added a chicane to broaden the frequency tuning range. They used a charge of up to 1 nC, obtaining results similar to those of Musumeci, Li, and Marinelli (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 terahertz 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 to 0.7 THz. The second parameter was the chicane's bending angle, which controlled R_{56}. Compared 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 to 1.6 THz.

Because of the high charge level, they achieved 2 μJ of terahertz 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 terahertz radiation at a 0.7-THz frequency. This intense and tunable terahertz radiation from a compact beamline could be useful for a typical spectroscopy-type equipment, pump-probe measurements in x-ray free-electron lasers (XFELs), or terahertz 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 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 have been published to explain and suppress this phenomenon (Heifets, Stupakov, and Krinsky, 2002; Huang and Kim, 2002; Saldin, Schneidmiller, and Yurkov, 2002a, 2002b; Saldin, Schneidmiller, and Yurkov, 2004; Wu, Huang, and Emma, 2008; Lumpkin, Dejus, and Sereno, 2009; Huang *et al.*, 2010; Marinelli, Hemsing, and Rosenzweig, 2011; Spampinati *et al.*, 2014; Ratner *et al.*, 2015; Prat *et al.*, 2017). Ratner *et al.* (2015) showed a picture of the longitudinal phase space with various conditions, and Fig. 33 shows the modulation amplified by the gain process; see Sec. II.D.5.

As described in Sec. II.D.5, Saldin, Schneidmiller, and Yurkov (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. R_{56} of the given beamline then converted an energy modulation to a density modulation. For the microbunching instability, the longitudinal space-charge force from the shot noise made 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).

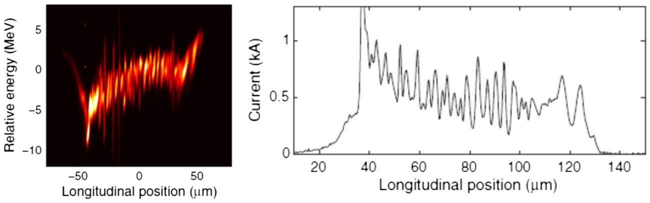

FIG. 33. Measurement of microbunching instability. Left panel: the measured longitudinal phase space without a laser heater. The strong modulation originated from the shot noise and its amplification. Right panel: the corresponding current profile. Adapted from Ratner *et al.*, 2015.

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). With this concept, they used a drift with a focusing channel to introduce a momentum modulation from the density modulation, and a chicane was then used to introduce an R_{56} optimized for the modulation conversion with maximum gain. This focusing channel and chicane formed a single cell of the amplifier, and each cell provided a 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 are close to typical FEL parameters. Thus, two cascades would provide a $G > 1000$. This is high enough to saturate the shot noise in this example.

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 using an undulator, with an undulator parameter of 0.58, a period of 1.9 cm, and a total of 11 periods. The chicanes provided R_{56} of 4, 2.5, and 1.5 mm, and they 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. Marinelli *et al.* (2013) 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 considered only 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 to generate high-power coherent radiation, large radiation bandwidth, or attosecond pulses for FEL facilities, as described by Schneidmiller and Yurkov (2010) and Dohlus, Schneidmiller, and Yurkov (2011). Another use of this cascade method was proposed by Ratner (2013), who suggested using the cascade method to achieve a high cooling rate for coherent electron cooling. Here the modulation imprinted by a hadron beam on the electron beam is amplified with the cascade amplifier, and the field from the amplified modulation of the 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 strongly consider transverse behavior, Litvinenko *et al.* (2018) introduced a new method called the plasma-cascade amplifier (PCA) that generates modulation at the plasma frequency using solenoid focusing. As previously mentioned, $G > 1$ occurs when the space-charge-induced momentum is strong enough. It is then converted to the density using 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 that of a chicane $(2L_{\mathrm{DL}}\theta^2,\ \theta \ll 1)$. However, the final density modulation is governed by how many particles gather in 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 ways to achieve this 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., $\omega_0(t) = \omega_0[1 + A\cos(\omega t)]$). 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. Because of the required relationship between the plasma oscillation frequency and its modulation frequency, system parameters such as the beam envelope, the solenoid-to-solenoid distance, and the charge level should be carefully designed. Thorough theoretical work was done by Litvinenko *et al.* (2018), and further work is ongoing; see Blaskiewicz (2019).

The method was experimentally demonstrated by Litvinenko *et al.* (2019) and 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 five 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 of around 200 for 0.6 THz. Figure 34 shows measured longitudinal beam profiles. The density modulation is 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 of around 0.4 THz, as Litvinenko *et al.* (2019) and Petrushina *et al.* (2019) expected from their simulations. See Litvinenko *et al.* (2021) for further details.

As the 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

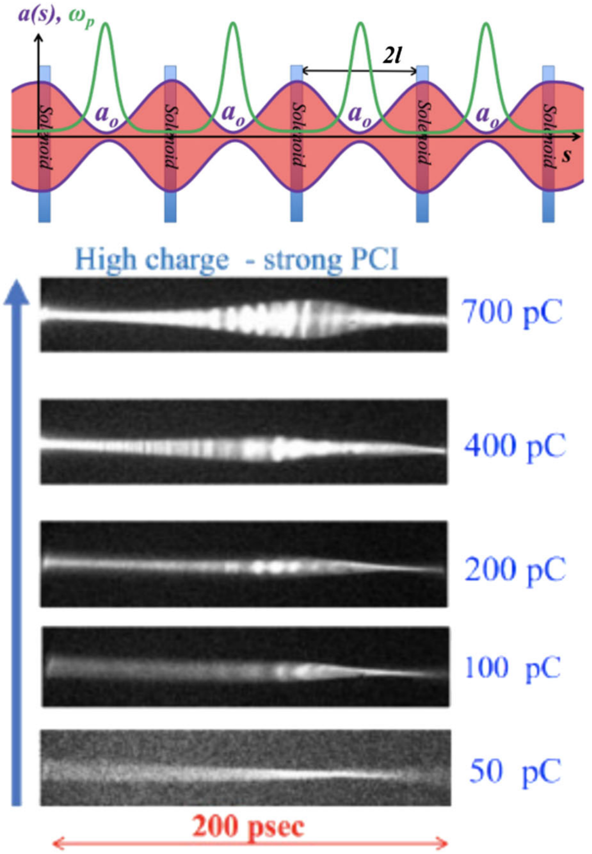

FIG. 34. Experimental demonstration of plasma-cascade amplification. The conceptual figure at the top shows (top) the configuration of a PCA with the corresponding plasma wavelength modulation and (bottom) the measured longitudinal density distribution. Adapted from Petrushina *et al.*, 2019.

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. As with 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 in Eq. (142). CSR cannot be generated without a dispersive element such as a dipole magnet. This means that the transverse and longitudinal phase space will be correlated along the path, and it makes analysis and control more difficult than that of 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 thus far.

However, CSR is indeed a field that can change the longitudinal profile. It is normally considered an obstacle to longitudinal beam manipulation. Several research efforts have 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.

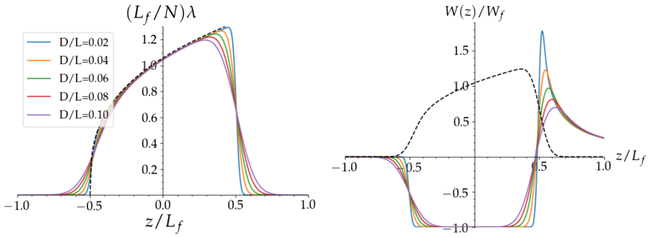

FIG. 35. Theoretical estimation of aberrations of an ideal current profile flattening the CSR wake and the aberration's impact on the CSR wake. Left panel: current profile with aberrations. Right panel: the corresponding CSR wakes. The dashed curve in the left panel shows the ideal profile from Eq. (201). The 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, Qiang, and Emma, 2013.

Mitchell, Qiang, and Emma (2013) theoretically showed that a certain beam shape can flatten the CSR wake, and that 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 domain ($a \leq z \leq b$). This shape is given in the left panel of Fig. 35, and the corresponding CSR field is in the right panel.

Here preserving the initial shape is critical to minimizing the CSR's impact on the emittance, so Mitchell, Qiang, and Emma (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

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

and $\mathcal{H}$ is given by

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

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 indicates that the final profile has the same shape as the initial profile but is compressed. Thus, the profile does 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}, an initially low energy spread, and a small transverse emittance are preferred. Beam matching in the beamline is another important factor.

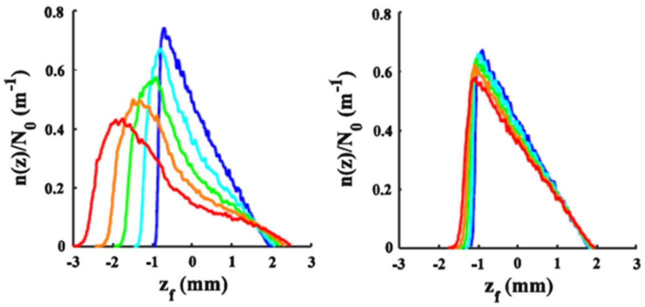

FIG. 36. Correction of CSR-induced aberration in a start-to-end simulation. The x axis is the beam's longitudinal position after an EEX beamline, while the y axis is the normalized longitudinal density. The colors represent charges of 1.1, 3.3, 5.5, 7.7, and 11 nC. The left panel depicts a simulation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

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 Secs. II.C.3 and 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. Figure 36 (left panel) corresponds to an EEX beamline with a 20° bend, and Fig. 36 (right panel) corresponds to a bending angle of 12°. The R_{56} of the doglegs were 0.29 and 0.18 m, respectively. The profile of the 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, Piot, and Zholents (2021) performed an optimization 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, they found that nonlinearities introduced by each part of the beamline could cancel each other out. The beamline that they used in the simulation consisted of a superconducting rf gun, 650-MHz linac cavities, a 3.9-GHz linac cavity, and two chicane compressors. Optimization variables included an accelerating gradient, a phase, frequencies of the 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 provides the desired final longitudinal profile. Even this setting included quite a high R_{56} for both chicanes, 0.129 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 (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 for 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 a 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 further manipulate the final profile.

C. Shaping profiles using wakefields

The longitudinal wakefield inside a given structure can be expressed as the convolution of the current distribution and the 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:

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

where k_i is the wave number of the ith mode and $\mathcal{K}_i$ is the loss factor of the 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 a 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., the 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. As with the methods described in Sec. IV, wakefields control the time-energy correlation and the subsequent beamline (such as the 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 follow the same principles as those in Sec. 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.

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.

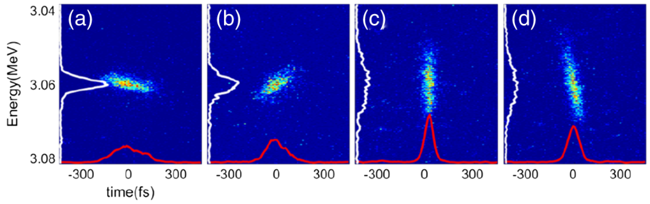

FIG. 37. Experimental demonstration of wakefield-based bunch compression using terahertz structure with a drive bunch. A measured longitudinal phase space of the target bunch and corresponding projections are shown. Each panel corresponds to a different drive bunch charge: (a) 0, (b) 0.6, (c) 0.9, and (d) 1.3 pC. Adapted from Zhao *et al.*, 2018a.

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 terahertz 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 Fig. 37(c).

While a similar compression can be easily accomplished with conventional rf cavities, wakefield-based ballistic bunching offers a few advantages. First, this compression requires a driving bunch that can be easily 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 the energy stability at 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. Zhao *et al.* (2018a) measured the arrival time jitter using a terahertz streaking method that was 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 in high demand in many modern accelerators. It is especially important for wakefield accelerators. Wakefield accelerators usually operate at high frequencies, 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 at a few picoseconds or less 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 the 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 slightly later than 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 for 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 for.

2. Bunch train generation

The bunch compression discussed in Sec. V.C.1 requires a condition $\sigma_z/\lambda_{\text{wake}} \ll 1$. If we go to the other regime ($\sigma_z/\lambda_{\text{wake}} \gg 1$), 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 another 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), who used a low-energy beam with drift, and by Antipov *et al.* (2013), who used a compact chicane.

Lemery *et al.* (2019) used a beam with a charge of 1 nC and an energy of 6.2 MeV. Two structures were prepared (with inner diameters of 450 and 750 μm and lengths of 5 and 8 cm), followed by an accelerating cavity providing 14 MeV of energy gain. Downstream diagnostics measured the phase space of both structures, and the modulation was successfully imprinted in the LPS. The measured wavelengths of the final modulation were 1.01 and 1.81 mm for the two structures. They showed good agreement with the expected fundamental wavelength 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 macrochirp, on top of the sinusoidal modulation, can provide either a frequency upshift or a frequency 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. In addition, depending on the linac phase,

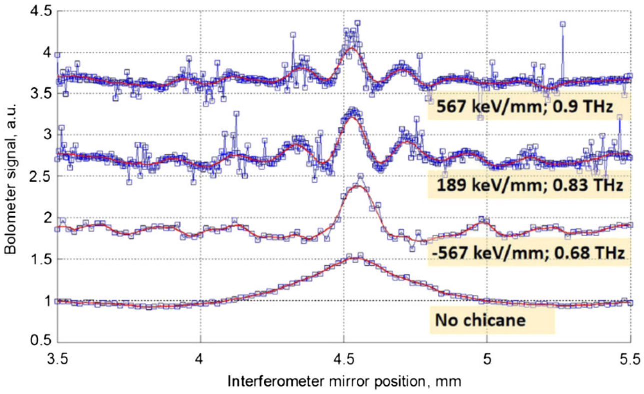

FIG. 38. Experimental demonstration of bunch train generation using a dielectric structure followed by a chicane. Shown are the measured CTR signals, which are the autocorrelation results. Each curve corresponds to a different initial energy chirp, as described in the label next to the curve. The squares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

the chirp was changed anywhere from −567 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 the methods of both Antipov *et al.* (2013) and Lemery *et al.* (2019). In the case of the wakefield, the modulation amplitude induced by the initial energy modulation and R_{56} was $A_{\mathrm{ind}} = Ck|R_{56}|\Delta\delta \exp\,[-(1/2)C^2k^2R_{56}^2\sigma_\delta^2]$, where $\Delta\delta$ is the amplitude of the initial energy modulation; see Saldin, Schneidmiller, and Yurkov (2002a). Note that the required $R_{56}\Delta\delta$ value for the same final modulation amplitude decreases as the 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 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. In addition, 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 will not occur if the bunch's longitudinal profile's frequency spectrum does not overlap with the structure's frequency spectrum. Thus, imprinting multiperiod modulation is not possible for a Gaussian profile. Both of the previously described experiments used flattop 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 see in the transverse manipulation of 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, owing 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.

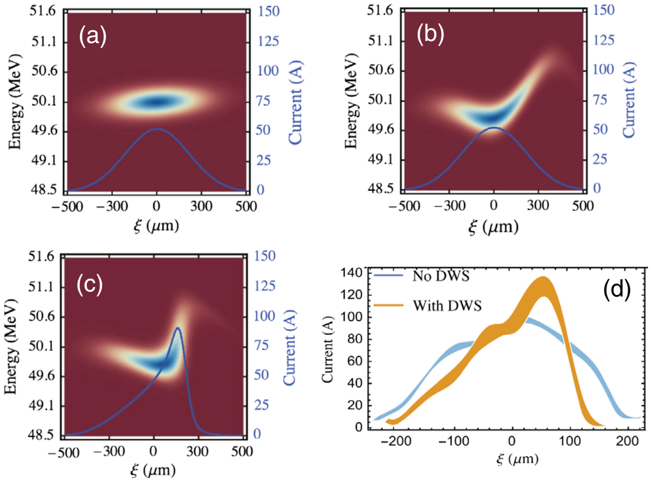

FIG. 39. Experimental demonstration of triangular current-profile generation using a dielectric structure followed by a chicane. (a)–(c) Simulation results. (d) Reconstructed profile from interferometer measurements with and without the structure. (a)–(c) Longitudinal phase space before the structure,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

Andonian *et al.* (2017) experimentally demonstrated this method 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 Fig. 39(d).

The method can be extended to generate arbitrary current profiles when we use a series of structures with appropriate frequencies. As in the work of Piot *et al.* (2012) and Ha, Conde, and Power (2019), 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, Conde, and Power (2019) and can be directly used for the longitudinal phase space by simply replacing transverse wigglers with wakefield structures. Simulations by Mayet, Assmann, and Lemery (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 yet been proposed. However, combining wakefield-based shaping with methods introduced in that paper could provide advantages. For example, one can adopt a wakefield-based deflector [see Novokhatski (2015)] instead of rf deflecting cavities for the method of 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 significant advantage regarding 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).

4. Control over the energy distribution

Antipov *et al.* (2014) and Emma *et al.* (2014) both 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 decreases, a quasilinear wakefield amplitude increases, and it successfully reduces the energy spread.

The low correlated energy spread from a dechirper can benefit all modern accelerator applications. For example, the x-ray FEL oscillator requires an energy spread of $< 1 \times 10^{-4}$ (Kim, Shvyd'ko, and Reiche, 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 (~400 fs) having a relative energy spread of ~2.6×10^{-5} (105 keV rms).

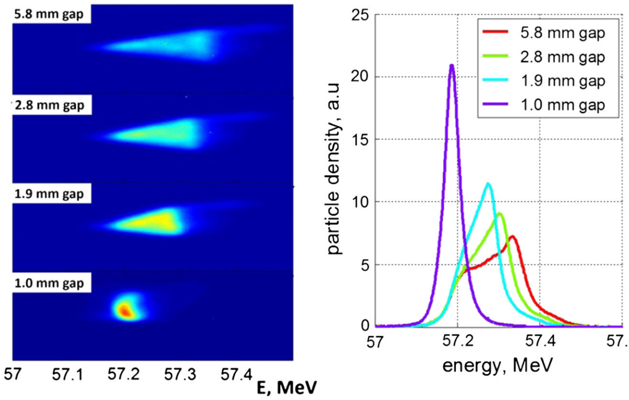

FIG. 40. Experimental demonstration of energy-spread reduction by wakefield from dielectric structure. Left panel: beam images on the spectrometer with different structure gap sizes. Right panel: corresponding profiles. From Antipov *et al.*, 2014.

Although these methods 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 $k\sigma_z \ll 1$, Eqs. (122) and (205) can be simplified to

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

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 have a nonlinear periphery in their phase space.

A simulation by Antipov *et al.* (2014) showed the feasibility of extending the linear region of the wakefield using a multimode structure. However, this approach does not provide any tunability after the fabrication of a structure. Antipov *et al.* (2014) also remarked on the use of two or more dechirpers to control nonlinearity issues. This multistructure approach was recently simulated by Mayet, Assmann, and Lemery (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 noticed 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, which was 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 $E_{\mathrm{rf}} \cos(k_{\mathrm{rf}} z)$, this gain can be approximated as $E_{\mathrm{rf}}[1 - (k_{\mathrm{rf}} z)^2/2]$ for $k_{\mathrm{rf}} z \ll 1$. Similarly, the wakefield from a structure having only a single mode and a uniform profile can be written as $(n_0 W_0/k_w) \times \sin[k_w(z + \Delta z/2)]$, where Δz is the width of the uniform profile and n_0 is the density. Because $k_w t$ should be small and the wavelength of the wakefield should be shorter than the wavelength of the rf field in the linac, we can approximate it as $n_0 W_0[-(\Delta z k_w^2/4)z^2 + (1 - k_w^2 \Delta z^2/8)z + \Delta z/2]$. 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 done by simply 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 the other hand, the 2-mm gap provided a 3.7-nm bandwidth.

This experiment was designed to observe the effect of linearization on the radiation. However, we 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 at a 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 for 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 an arbitrary correlation in the phase space; see Ha, Conde, and Power (2019). This concept was demonstrated in a simulation (Mayet, Assmann, and Lemery, 2020) using a series of dielectric structures. Because the choice of parameter set becomes too complex, Mayet, Assmann, and Lemery (2020) used an optimization algorithm to achieve the best results. Figure 41 shows two of their simulation results. The top panels show the longitudinal phase space with and without linearization by wakefields, and the bottom panels show the corresponding structure geometries. They used a total of ten 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.

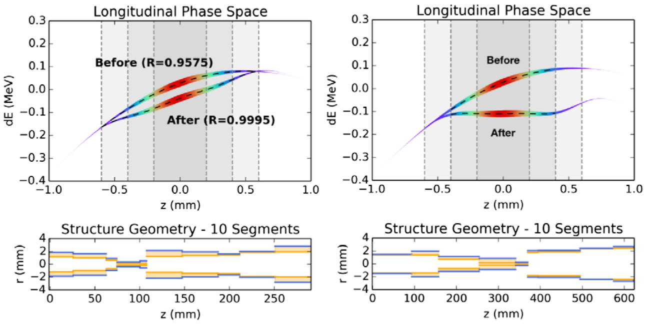

FIG. 41. Longitudinal correlation control via a series of dielectric structures. Top panels: simulated longitudinal phase spaces before and after the correction using ten dielectric structures. Bottom panels: the corresponding structure geometries for the simulations. The left column shows an example of linearization while the right column shows dechirping. Adapted from Mayet, Assmann, and Lemery, 2020.

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 d.o.f.'s. In Sec. 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 d.o.f.'s.

B. Coupling between the 2 transverse degrees of freedom

1. Producing beams with canonical angular momentum

Canonical-angular-momentum- (CAM-) dominated, or "magnetized," beams have important applications in the 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 copropagates 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 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 a 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 $L = (eB_z/2)r^2$ 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 will become coupled.

As discussed in Sec. 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 a random momentum direction; in Fig. 42(b) the beam is dominated by its angular momentum and the electrons shear in a vortex pattern; and in Fig. 42(c) the beam is dominated by space charge, and the electrons repel each other by the Coulomb force and move outward.

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

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

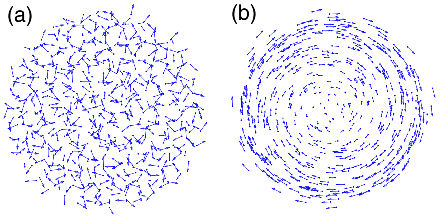

FIG. 42. Motions of the electrons when the beam is dominated by (a) emittance, (b) angular momentum, or (c) space charge. Each dot represents an electron in (x, y) space, and the arrows show the magnitude and direction of the electrons' velocities. From Sun, 2005.

where σ is the transverse rms size, $K_p = 2I/I_0\beta^3\gamma^3$ is the generalized perveance, I is the absolute value of the peak beam current, $I_0 = 4\pi\epsilon_0 mc^3/e \approx 17$ 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 Sec. 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 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 in Eq. (207), where the canonical angular momentum term $\mathcal{L}$ plays the same role as the emittance term in the beam-envelope equation. It can therefore be introduced to an effective emittance $\epsilon_{\mathrm{eff}} = \sqrt{\epsilon_u^2 + \mathcal{L}^2}$, as noted in Sec. II. However, a large magnetic field is required to produce an angular-momentum-dominated beam in order for the correlation between the two transverse degrees of freedom to dominate, i.e., $\mathcal{L} \gg \epsilon_u$.

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. Sun *et al.* (2004) further accelerated the beam using a superconducing cavity, and the MAM was measured at ~15 MeV. Similar investigations were conducted more recently at Jefferson Laboratory on a 300-keV electron beam produced with 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).

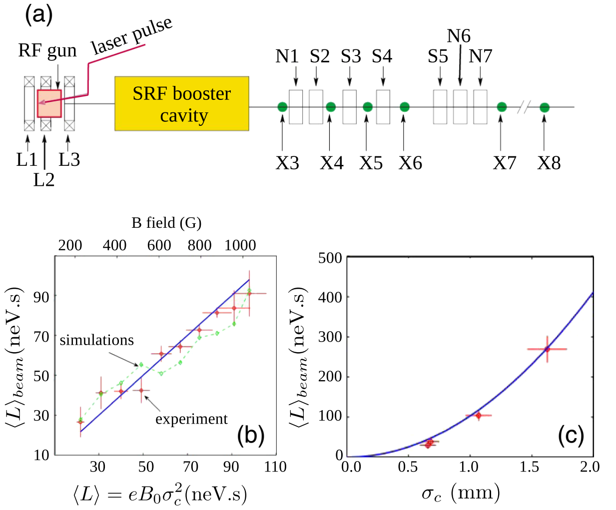

FIG. 43. (a) Experimental setup used to produce a CAM-dominated beam at the Fermilab A0 facility. L, solenoidal magnetic lenses; N and S, normal- and skew-quadrupole magnets; X (which represents "cross"), diagnostic stations. Dimensions are in mm. (b) Example of measured conversion between CAM and MAM. The dashed curve shows the corresponding simulations. (c) Demonstration of the quadratic dependence on laser spot size on the cathode. (b),(c) The circles are measurements and the lines are fits. 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 2 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, Derbenev, and Flöttmann, 2001). The technique consists of manipulating an angular-momentum-dominated beam produced by a photoinjector using the linear transformation described by Derbenev and Skrinsky (1978). A round-to-flat beam transformer, consisting of three skew quadrupoles and drift spaces, was discussed by Burov and Danilov (1998). This technique was proposed as a way to produce a high-aspect-ratio beam to mitigate beamstrahlung in future linear colliders while also circumventing the use of an electron damping ring conventionally used to reduce the vertical emittance (Brinkmann, Derbenev, and Flöttmann, 2001). Likewise, the technique was also adapted for applications of microwave- and terahertz-radiation generation (Carlsten and Bishofberger, 2006; Kim and Kumar, 2007) using the concept of the Smith-Purcell backward oscillator (Andrews *et al.*, 2005).

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

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

We specify the coordinates of a particle in transverse trace space using the following two vectors:

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

The corresponding 4×4 beam matrix is

$$\Sigma = \begin{pmatrix} \langle \mathbf{X}\mathbf{X}^T \rangle & \langle \mathbf{X}\mathbf{Y}^T \rangle \\ \langle \mathbf{Y}\mathbf{X}^T \rangle & \langle \mathbf{Y}\mathbf{Y}^T \rangle \end{pmatrix}. \tag{210}$$

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

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

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

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

From Eq. (212), we obtain

$$\langle \mathbf{X}\mathbf{X}^T \rangle \cos^2\theta + \langle \mathbf{Y}\mathbf{Y}^T \rangle \sin^2\theta + (\langle \mathbf{X}\mathbf{Y}^T \rangle + \langle \mathbf{Y}\mathbf{X}^T \rangle)\sin\theta\cos\theta = \langle \mathbf{X}\mathbf{X}^T \rangle.$$

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

$$\langle \mathbf{X}\mathbf{X}^T \rangle = \langle \mathbf{Y}\mathbf{Y}^T \rangle, \tag{213}$$

$$\langle \mathbf{X}\mathbf{Y}^T \rangle = -\langle \mathbf{Y}\mathbf{X}^T \rangle. \tag{214}$$

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

$$\langle \mathbf{X}\mathbf{Y}^T \rangle^T = -\langle \mathbf{Y}\mathbf{X}^T \rangle^T = -\langle \mathbf{X}\mathbf{Y}^T \rangle, \tag{215}$$

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

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

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

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

$$\langle \mathbf{X}\mathbf{X}^T \rangle = \langle \mathbf{Y}\mathbf{Y}^T \rangle = \varepsilon T_0, \quad \text{with} \quad T_0 = \begin{pmatrix} \beta & -\alpha \\ -\alpha & (1+\alpha^2)/\beta \end{pmatrix}, \tag{218}$$

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

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

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

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

$$I_1 = \varepsilon_{4\mathrm{D}} = \sqrt{\det(\Sigma)}, \tag{221}$$

$$I_2(\Sigma) = -\tfrac{1}{2}\mathrm{Tr}(J_{4\mathrm{D}}\Sigma J_{4\mathrm{D}}\Sigma), \tag{222}$$

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

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

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

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

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

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

$$J_{2\mathrm{D}}T_0 J_{2\mathrm{D}}T_0 = -I, \tag{225}$$

which leads to

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

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

Therefore, from Eq. (222) we have

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

Finally, the two transverse emittances can be derived from Eqs. (224) and (226) as follows:

$$\varepsilon_\pm = \varepsilon \pm \mathcal{L}. \tag{227}$$

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 as follows:

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

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 Figs. 44(a)–44(c). Ultimately, the experiment demonstrated the generation of a flat beam with a transverse-emittance ratio of $\varepsilon_y/\varepsilon_x \simeq 100$ (Piot, Sun, and Kim, 2006); see Fig. 44. The experiment confirmed the underlying physics and was validated against numerical simulation; see Figs. 44(d) and 44(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 of ~2 smaller than the thermal emittance estimated from the laser spot size on the cathode, which was ~1 μm.

Further experiments, carried out at the AWA facility at Argonne, demonstrated the generation of a flat beam using a similar principle but with a 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).

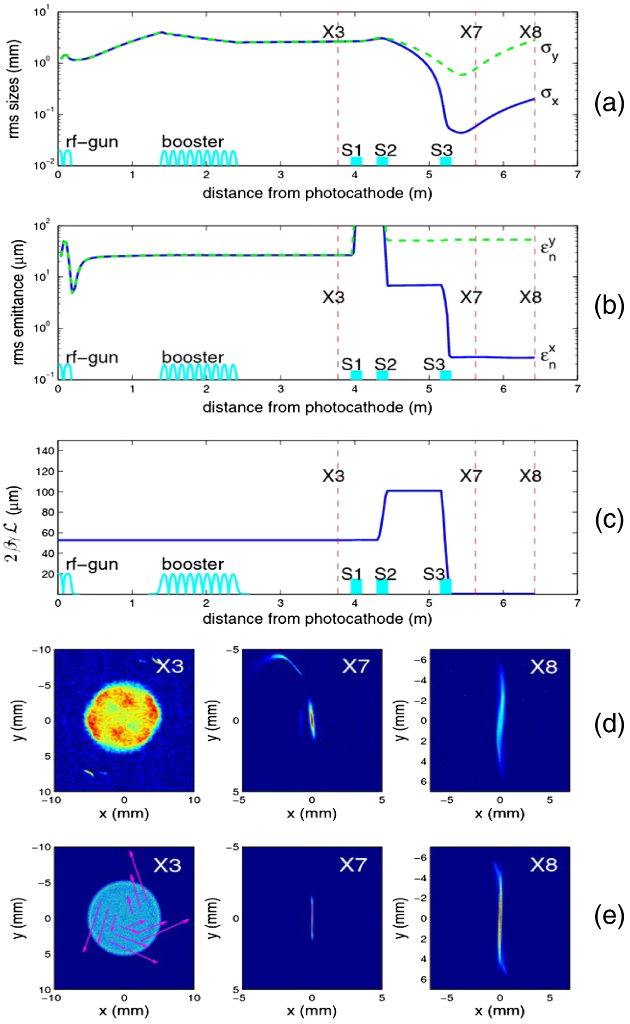

FIG. 44. (a)–(c),(e) Simulations and measurements of flat-beam generation for the Fermilab A0 photoinjector. (a)–(c) ASTRA simulations of, respectively, the beam size, emittance, and magnetization evolution along the beamline showcasing the removal of angular momentum using three skew-quadrupole magnets $S1$, $S2$, and $S3$ to transform the incoming magnetized beam into a flat beam. The images displayed in (d) are measured beam distributions for the $X3$, $X7$, and $X8$ screens, while the distribution shown in (e) shows the corresponding numerical simulations. The labels Xi refer to the diagram appearing in Fig. 43(a). Adapted from Piot, Sun, and Kim, 2006.

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 intrabeam scattering (Aiba, Ehrlichman, and Streun, 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-breakup (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 a symmetric arrangement of five skew quadrupoles as follows:

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

In Eq. (230) Q_{Si} and L_i are the following 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 the calculations, note that a skew quadrupole is obtained through a 45° rotation of a normal quadrupole as follows:

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

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

In Eqs. (232) we use 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

$$\begin{aligned} R_{XY} &= R_{\pi/4}{}^{-1} \begin{pmatrix} C & 0 \\ 0 & \bar{C} \end{pmatrix} R_{\pi/4} \\ &= \frac{1}{2} \begin{pmatrix} C + \bar{C} & C - \bar{C} \\ C - \bar{C} & C + \bar{C} \end{pmatrix}. \end{aligned} \tag{234}$$

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 for which 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 for which 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

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

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 by Kuske and Kramer (2016) and by Armborst (2016) for the BESSY II machine. The purpose of the manipulation was to exchange the larger horizontal emittance for the smaller vertical emittance prior to injection in order to stay within the acceptance of the ring. However, owing to the additional dispersion generated, Kuske and Kramer concluded that a different way of transverse-emittance exchange, i.e., "resonance crossing" (Carli *et al.*, 2002; Aiba and Kallestrup, 2020), was more feasible. Indeed, the emittance exchange via resonance crossing in a booster synchrotron was demonstrated by Kallestrup and Aiba (2020). On the other hand, in the upgrade of the Advanced Photon Source at Argonne National Laboratory, 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 on those of 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 of mitigating 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 to increase the slice energy spread, thereby reducing their sensitivity to the microbunching instability while simultaneously reducing the transverse emittance to improve the FEL performance (such as to reduce the gain length). Utilization of emittance exchange in a compact XFEL facility was proposed by Graves *et al.* (2018). It was also suggested as a way to mitigate BBU

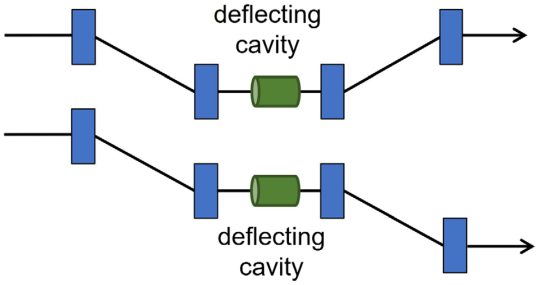

FIG. 45. Top: partial emittance exchange using a deflecting cavity and chicanes. Bottom: complete emittance exchange using a deflecting cavity and double doglegs.

instabilities and improve efficiency in energy-recovery linacs (Piot, 2009). Mathematically, we follow an approach similar to the one described in Sec. VI.B.3 and seek a transformation that provides a matrix of the form

$$R_{XZ} = \begin{pmatrix} 0 & A \\ B & 0 \end{pmatrix},$$

$$\begin{pmatrix} \mathbf{X} \\ \mathbf{Z} \end{pmatrix} \mapsto R_{XZ} \begin{pmatrix} \mathbf{X} \\ \mathbf{Z} \end{pmatrix} = \begin{pmatrix} A\mathbf{Z} \\ B\mathbf{X} \end{pmatrix}. \tag{241}$$

A beamline possibly capable of providing such a transformation was first discussed by 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 Sec. II.B.3, the overall transfer matrix of the system depicted in Fig. 45(a) is

$$R_{\mathrm{EEX}} = R_{-\mathrm{DL}} R_{\mathrm{TDC}} R_{\mathrm{DL}}, \tag{242}$$

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

Cornacchia and Emma (2002) showed 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 antidiagonal blocks of R_{EEX} to vanish for a thin-lens TDC. From this, they showed that the emittance is mapped as

$$\begin{aligned}(\varepsilon_{x,0}, \varepsilon_{z,0}) &\mapsto (\varepsilon_x, \varepsilon_z) \\ &= \left(\sqrt{\varepsilon_{z,0}^2 + \lambda^2 \varepsilon_{x,0}\varepsilon_{z,0}}, \sqrt{\varepsilon_{x,0}^2 + \lambda^2 \varepsilon_{x,0}\varepsilon_{z,0}}\right),\end{aligned} \tag{243}$$

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 $[(\varepsilon_{x,0}, \varepsilon_{z,0}) \overset{R_{\mathrm{EEX}}}{\mapsto} (\varepsilon_{z,0}, \varepsilon_{x,0})]$. Kim and Sessler (2006) subsequently recognized that flipping the second half of the chicane such 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

$$\kappa\eta + 1 = 0, \tag{244}$$

and the resulting transformation takes the form

$$R_{\mathrm{EEX}} = \begin{pmatrix} 0 & \ell_c/4 & -(\ell + \ell_c/4)/\eta & [\eta^2 - \xi(4\ell + \ell_c)/4]/\eta \\ 0 & 0 & -1/\eta & -\xi/\eta \\ -\xi/\eta & (-\ell\xi - \ell_c\xi/4 + \eta^2)/\eta & \ell_c\xi/4\eta^2 & \ell_c\xi^2/4\eta^2 \\ -1/\eta & -(\ell + \ell_c/4)/\eta & \ell_c/4\eta^2 & \ell_c\xi/4\eta^2 \end{pmatrix}, \tag{245}$$

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 antidiagonal. In the most general case, it is possible to exchange the beam emittance with a 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

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

In Eq. (246) $\alpha_{i,0}$ 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 minimized with a proper choice of either the longitudinal or transverse Courant-Snyder (CS) parameters. A possible solution consists of tuning the incoming LPS chirp to fulfill $\alpha_{z,0} = \beta_z/\xi$, corresponding to an incoming longitudinal-phase-space chirp $\mathcal{C} \equiv d\delta/dz|_0 = -1/\xi$ that produces a minimum bunch length at the cavity location (Sun *et al.*, 2007). Such a beamline was numerically investigated by 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 the 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 to provide a simple way of canceling the thick-lens effect of the TDC, which induces beam energy gain, as detailed in Sec. II.B.3. The design also included development of a chicane-based exchanger with quadrupole magnets to control the dispersion and circumvent

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

	Simulated		Measured	
	In	Out	In	Out
$\varepsilon_{\mathrm{x}}^{n}$	2.9	13.2	2.9 ± 0.1	11.3 ± 1.1
$\varepsilon_{\mathrm{y}}^{n}$	2.4	2.4	$2.4 + 0.1$	2.9 ± 0.5
$\varepsilon_{\mathrm{z}}^{n}$	13.1	3.2	13.1 ± 1.3	3.1 ± 0.3

the limitation associated with the early design (Cornacchia and Emma, 2002). A similar design was later discussed by 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 by Nanni and Graves (2015). In addition, in the shaping process collective effects such as CSR can significantly reduce the shaping quality (Carlsten, Bishofberger, Russell, and Yampolsky, 2011; Zholents and Zolotorev, 2011; Ha *et al.*, 2016). Here the exchange introduces more difficulty than other well-studied beamlines such as chicane. Research to control CSR's impact on shaping are under way (Ha, Power, Conde, Doran, and Gai, 2017; Ha, Conde, Doran *et al.*, 2018; Ha, Conde, Power, and Wisniewski, 2018; Malyzhenkov and Scheinker, 2018).

2. Current-profile shaping

An important application of emittance exchange is its potential capability of 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 $(\mathbf{X}, \mathbf{Y})^T = R_{\mathrm{EEX}}(\mathbf{X}_0, \mathbf{Y}_0)^T$. In particular, the electron's final longitudinal coordinates $\mathbf{Z} = (z, \delta)$ are solely functions of its initial transverse coordinates $\mathbf{X}_0 = (x_0, x_0')$:

$$\begin{aligned} z &= -\tfrac{\xi}{\eta} x_0 - \tfrac{\ell\xi - \eta^2}{\eta} x_0', \\ \delta &= -\tfrac{1}{\eta} x_0 - \tfrac{\ell}{\eta} x_0'. \end{aligned} \tag{247}$$

Exploiting the mapping described by Eqs. (247), 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 onto 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 which of the two-dimensional phase spaces the function corresponds to. Consequently, many of the techniques discussed in Sec. IV can be readily applied. We can use a mask to shape the incoming distribution in the $\mathbf{X}$ space and map this distribution onto the $\mathbf{Z}$ plane, as originally proposed by Piot *et al.* (2010) and as further discussed by Jiang *et al.* (2011). In principle, arbitrary bunch current distributions can be achieved using the EEX technique (Piot, Sun, Power, and Rihaoui, 2011), but ultimately the incoming emittance partition limits the shaping resolution, as discussed in Sec. III; see Fig. 46.

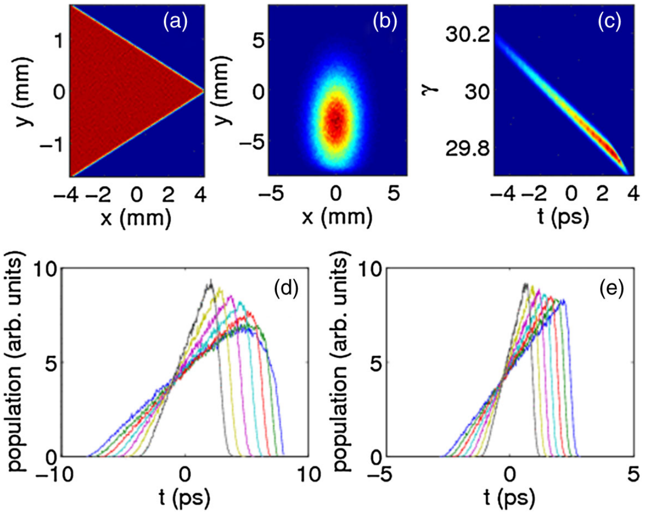

FIG. 46. Example of simulated generation of a linearly ramped current profile from (a) an initial uniform triangular distribution shown with corresponding (b) final transverse distribution and (c) longitudinal phase space downstream of the EEX beamline. (d),(e) Current profiles after EEX are shown for initial horizontal and longitudinal profiles equal to $(10, 1)$ and $(1, 10)$ μm, respectively. From Piot, Sun, Power, and Rihaoui, 2011.

An experimental demonstration of the shaping capability of an EEX beamline was demonstrated by Sun *et al.* (2010b), who transversely sliced the incoming beam into multiple beamlets using a multislit plate and sent it 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 that the beam was temporally modulated when the incoming beam was intercepted by the multislit mask located upstream of the EEX beamline. In addition, the experiment 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 by Sun *et al.* (2010a); see also Figs. 47(b)–47(e). This setup was also employed to investigate the generation of narrowband coherent transition radiation with a tunable central wavelength (Piot, Sun, Maxwell *et al.*, 2011).

LPS shaping was further investigated at the Argonne Wakefield Accelerator (AWA) (Ha *et al.*, 2017) 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

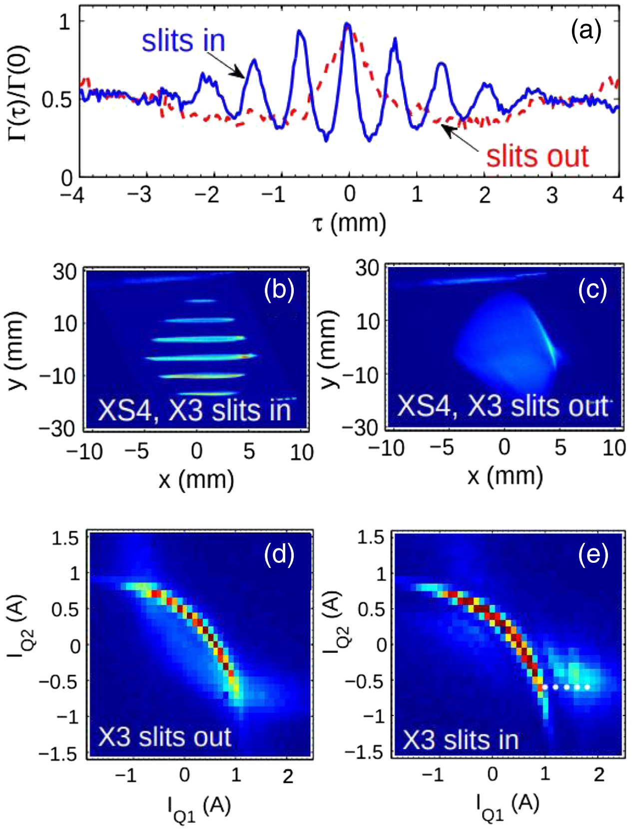

FIG. 47. Experimental demonstration of subpicosecond bunch train generation with an EEX beamline. (a) Normalized autocorrelation function $\Gamma(\tau)/\Gamma(0)$ of a CTR signal recorded with (solid line) and without (dashed line) insertion of a multislit mask upstream of the EEX beamline (τ is the optical path difference). (b),(c) Corresponding beam transverse densities downstream of the EEX. The vertical axes in (b) and (c) are 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. The total normalized CTR energy detected at $X24$ as a function of quadrupole magnet currents I_{Q_1} and I_{Q_2} with $X3$ slits (d) out and (e) in the beamline. The bolometer signal is representative of the inverse of the bunch duration σ_t. The intensity island appearing at $(I_{Q1}, I_{Q2}) \simeq (1.5, -0.5)$ in (e) is indicative of a density-modulated bunch. From Sun et al., 2010a.

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. Figures 48(g)–48(j) display some examples of the final current distribution experimentally that were achieved with four different mask shapes. After these initial experimental demonstrations, AWA's EEX shaping beamline was then used to set new records for transformer ratios. 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 such as a field-emission-array photocathode (Graves et al., 2012) or by impressing the beam modulation in the transverse phase space upstream of the exchanging beamlines, which is then converted in a modulation on the LPS (along either the energy or temporal direction). Nanni, Graves, and Moncton (2018) employed a transmission mask such as that described in Sec. IV.E.2 to produce the initial transverse modulation, while Ha, Conde, Power, Shao, and Wisniewski (2019) explored the use of a transverse wiggler, as shown in Fig. 26.

3. Bunch compression

The final bunch length downstream of the EEX is $\sigma_{z,f} = [R^2_{\mathrm{EEX},33}\sigma_x^2 + R^2_{\mathrm{EEX},44}\sigma_{x'}^2 + 2R_{\mathrm{EEX},33}R_{\mathrm{EEX},44}\langle xx'\rangle]$. Substituting Eq. (245) in, with the assumption that $\ell_c = 0$ for simplicity, and expressing all beam quantities in term of ε_x and the CS parameters (α_x, β_x) associated with the horizontal phase space upstream of the EEX beamline, we obtain

$$\sigma_{z,f} = \sqrt{\frac{\varepsilon_x}{\beta_x}}\left\{\left[\frac{\xi}{\eta}\beta_x - \left(\eta - \frac{\ell\xi}{\eta}\right)\alpha_x\right]^2 + \left(\eta - \frac{\ell\xi}{\eta}\right)^2\right\}^{1/2}. \tag{249}$$

The latter part of Eq. (249) indicates that, upon choosing the proper incoming phase-space correlation $\alpha_x/\beta_x = \xi/(\eta^2 - \ell\xi)$, the final bunch length becomes $\sigma_{z,f} = |\eta - \ell\xi/\eta|\sqrt{\varepsilon_x/\beta_x}$, which depends on the EEX design and can be made small. Based on this observation, Carlsten, Bishofberger, Russell, and Yampolsky (2011) investigated the ability of an EEX beamline to produce short bunches. They specifically discussed design choices that lead to extremely short final bunch duration $\sigma_{z,f}/c$ at the subfemtosecond timescale.

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 (such as timing, energy, and energy spread) will be manifested in the transverse phase space, resulting in beam position and spot size jitter (Ha, Power, Conde, Doran, and Gai, 2017a).

4. Double phase-space exchangers

A limitation of the EEX technique for shaping the beam profile stems from the associated emittance exchange between 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 $(\mathbf{X}_0, \mathbf{Z}_0) \mapsto (\mathbf{Z}_1, \mathbf{X}_1) \mapsto (\mathbf{X}_2, \mathbf{Y}_2)$.

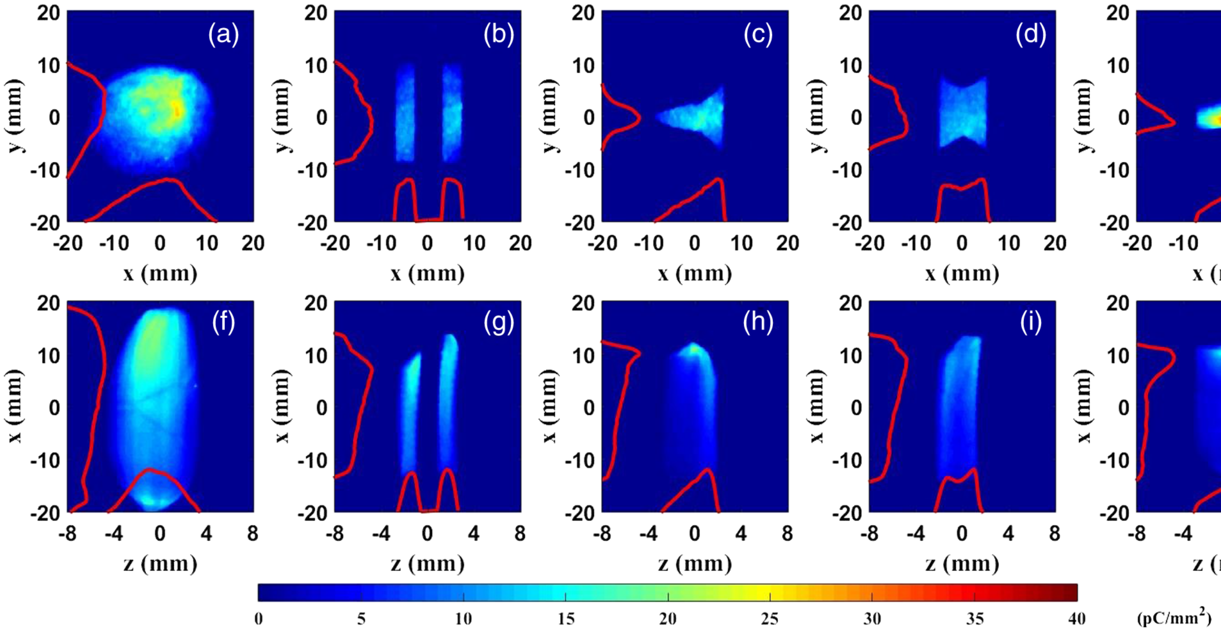

FIG. 48. Experimental demonstration of versatile bunch shaping using an EEX beamline at AWA. (a)–(j) Beam-shaping experiments using different transverse masks at the AWA EEX beamline. (a)–(e) are shot upstream of the EEX, and (f)–(j) are shot downstream of the EEX. From Ha *et al.*, 2016.

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 this limitation, another advantage of the double phase-space exchanger configuration was recognized by 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 representing the EEX using Eq. (245) as follows:

$$R_{\mathrm{DEEX}} = R_{\mathrm{EEX}} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} R_{\mathrm{EEX}}, \tag{250}$$

assuming that $L_c = 0$ such that $R_{\mathrm{EEX}} \simeq \left(\begin{smallmatrix} 0 & M \\ N & 0 \end{smallmatrix}\right)$. Consequently, the total matrix of the double EEX (DEEX) simplifies as

$$R_{\mathrm{DEEX}} = \begin{pmatrix} MBN & 0 \\ 0 & NAM \end{pmatrix} \overset{\gamma \gg 1}{\longrightarrow} \begin{pmatrix} MN & 0 \\ 0 & NAM \end{pmatrix}, \tag{251}$$

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 ultrarelativistic limit $\gamma \gg 1$ comes from the longitudinal dispersion associated with a longitudinal 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 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 the element $A_{11} = 1/A_{22} \equiv a$, we find that

$$NAM = \begin{pmatrix} [\ell\xi + a^2(\ell\xi - \eta^2)]/a\eta^2 & [\xi(a^2+1)(\ell\xi - \eta^2)]/a\eta^2 \\ \ell(a^2+1)/a\eta^2 & (\ell a^2\xi + \ell\xi - \eta^2)/a\eta^2 \end{pmatrix}. \tag{252}$$

As an example, considering the case in which the beamline is designed such that $\ell(a^2+1)/a\eta^2 \ll 1$, Eq. (252) is

$$NAM \simeq \begin{pmatrix} -a & \xi(a + 1/a) \\ 0 & -1/a \end{pmatrix}. \tag{253}$$

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 Fig. 49.

Additionally, the configuration enables control of nonlinear correlation in the longitudinal phase space using nonlinear magnets between the two EEXs. For instance, Seok, Ha *et al.* (2019) discussed various configurations to reduce the beam final energy spread via control of the longitudinal-phase-space nonlinearities with a DEEX configuration. The method

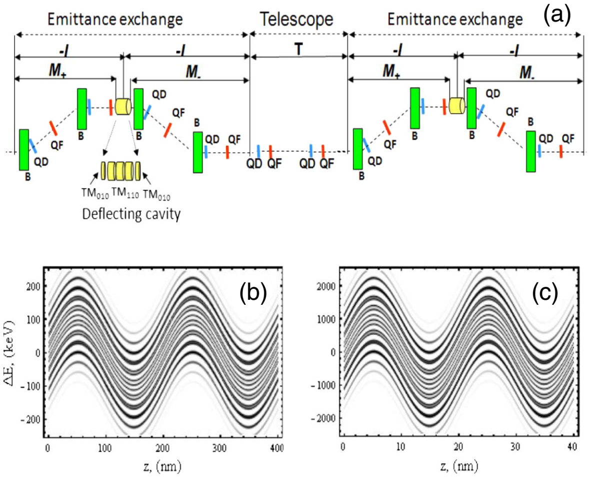

FIG. 49. (a) Schematic of a double-EEX-based bunch compressor with (b),(c) a simulated example of an application that provides a tenfold increase in modulation frequency of an incoming laser-modulated electron beam. From Zholents and Zolotorev, 2011.

closely follows the technique described in Sec. IV.E.1. Here the EEX is first 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 nonlinear transformation is applied to the horizontal phase space,

$$x_1' = f(x_1, y_1), \tag{255}$$

and the second EEX is then 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 depicts a configuration capable of introducing a nonlinear phase-space correlation with a DEEX beamline. In this case, a sextupole magnet, judiciously located downstream of the first EEX, imposes a quadratic correlation. Extension of this technique to control higher-order distortion is straightforward: correction of third-order nonlinearties to suppress current spike “horns” sometimes encountered in bright-beam injectors was demonstrated via numerical simulations (Seok, Chung *et al.*, 2019) as an alternative to the technique described in Sec. IV.B.2; see also Charles *et al.* (2017).

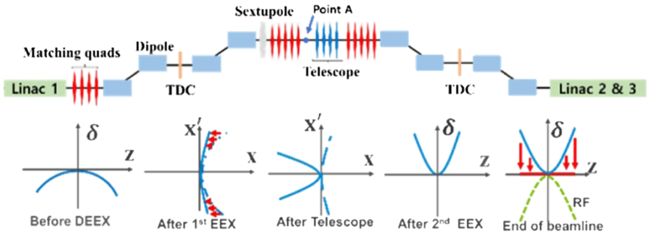

FIG. 50. Overview of a DEEX-based energy-spread reduction via correction on quadratic nonlinearity with a sextupole magnet. From Seok, Ha *et al.*, 2019.

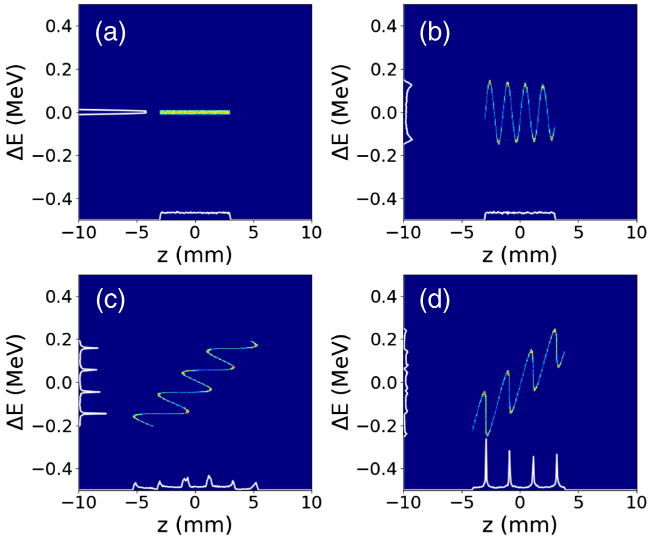

FIG. 51. Longitudinal phase spaces from numerical particle tracking. (a) The beam’s initial phase space. (b) The phase space after a modulator. The final phase space after a double-EEX beamline having (c) spectral or (d) temporal bunching. From Seok *et al.*, 2021.

Another recent development, introduced by Seok *et al.* (2021), is using a DEEX beamline with a phase-space modulator, such as the wakefield structure or transverse wiggler discussed in Secs. IV and V. An approach 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 consider only 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 nonzero, so there is a corresponding h. However, R_{65} is 0 and R_{66} is 1, so no controls are allowed in the energy distribution. On the other hand, a double EEX beamline that provides nonzero 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, Power, Conde, Doran, and Gai (2017b).

Figure 51 displays numerical tracking results. As in the density modulation case, a dielectric structure in front of a double-EEX beamline imprints a sinusoidal modulation on the longitudinal phase space. This modulation is then converted into energy spikes that are 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 multienergy beam with controllable time separation from a single bunch. This may be a useful technique to generate multicolor radiation for various spectroscopy methods or pump-probe experiments.

D. Generalized phase-space repartitioning between the 3 degrees of freedom

1. Flat-beam transformation combined with emittance exchange

Future electron-positron linear colliders (LCs) require unprecedented ultralow transverse emittances. Additionally, the transverse-emittance ratio should be high to mitigate beamstrahlung 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 femtocoulomb to nanocoulomb depends on the LC technology choice. The vertical emittance required to mitigate beamstrahlung 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 an example from the International Linear Collider (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 $\varepsilon_{6\mathrm{D}}^n = \varepsilon_x^n \varepsilon_y^n \varepsilon_z^n \simeq 7150$ $\mu\mathrm{m}^3$ is generally larger than values typically achieved by state-of-the-art photoinjectors (such as 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 (RFBT) and EEX beamlines. Specifically, Xu *et al.* (2021) considered a 3.2-nC bunch produced in an L-band photoinjector and demonstrated 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 a RFBT beamline. They suggested 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

Thus 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 have 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 (x, x') and (z, δ) while following Carlsten, Bishofberger, Duffy *et al.* (2011). They noticed that the matrix of a deflecting cavity in (x, x', z, δ) under the thin-lens approximation [i.e., the one 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 in the case of a magnetized beam [i.e., Eq. (228)] in the (x, x', z, δ) plane combined with several deflecting cavities separated by "drift" [with the matrix described using Eq. (43) with $\eta = 0$] could repartition the emittances between the (x, x') and (z, δ) planes. Carlsten, Bishofberger, Duffy *et al.* (2011) considered the practical case of a tilted-front laser pulse impinging on a photocathode such that the initial correlation is imparted in a (x, x', z, δ) phase space of the form

$$z_c = z + \tau x, \tag{256}$$

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

$$\Sigma_c = \begin{pmatrix} \sigma_x^2 & 0 & \tau\sigma_x^2 & 0 \\ 0 & \sigma_x'^2 & 0 & 0 \\ \tau\sigma_x^2 & 0 & \sigma_z^2 + \tau^2\sigma_x^2 & 0 \\ 0 & 0 & 0 & \sigma_\delta^2 \end{pmatrix}. \tag{257}$$

The corresponding intrinsic emittances are

$$\begin{aligned} \varepsilon_x &\simeq \sigma_x \sigma_\delta \tau, \\ \varepsilon_z &\simeq \frac{\sigma'_x \sigma_z}{\tau}, \end{aligned} \tag{258}$$

where it is assumed that $\tau^2 \gg \sigma'_x \sigma_z / \sigma_x \sigma_\delta$. Carlsten, Bishofberger, Duffy *et al.* (2011) showed 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 [such as the magnetic chicane described using 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 3 degrees of freedom (Yampolsky *et al.*, 2010).

VII. FUTURE DIRECTIONS

The techniques described in this review along with that discussed by Hemsing *et al.* (2014) have opened 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 following list is by no means exhaustive given the vigorous ongoing 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 Sec. 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 past decade, new ultracold electron sources based on trapped atoms have emerged (Claessens *et al.*, 2005; Zolotorev, Commins, and Sannibale, 2007). Although the primary motives of this work were related to the generation of an 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 that discussed in Sec. 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 tool kit for *ab initio* shaping of the beam distribution.

Likewise, novel laser architectures based on coherently combining laser pulses that have been individually controlled (Lemons *et al.*, 2021) could form laser pulses with finer structures to precisely program the photoemission process. However, they are 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 design of electromagnets with unusual properties (such as production of a negative magnetic permeability) that could expand the methods described in Sec. IV. 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. These emerging technologies are expected to be critical to furthering the development of beam-control techniques based on external electromagnetic fields. Finally, the discovery of knotted solutions to Maxwell's equations (Ranada, 1990; Kedia *et al.*, 2013) could also have applications to beam shaping, as they provide faster (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 by Irvine and Bouwmeester (2008) used tightly focused circularly polarized laser pulses.

C. Shaping the beam using collective effects

Along the lines of the discussion in Sec. VII.B, we expect new engineered 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 that these fields are carrier envelope phase 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 past two decades, wakefield structures based on metamaterials 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 to suppress noxious modes in 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 (such as to synthesize the desired longitudinal or transverse force), thereby expanding on techniques presented in Sec. V.

D. Redistributing phase space between planes

Exchanging phase space between 2 and 3 degrees of freedom has facilitated the generation of shaped beams for advanced acceleration concepts such as those discussed in Sec. VI. We note that thus far the implementation of phase-space-exchanging beamlines has often been based on a simple configuration and further optimization, such as 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 by 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 timescales along the lines of the ideas discussed by Graves *et al.* (2019) could have groundbreaking consequences for the development of room-sized coherent x-ray sources.

LIST OF SYMBOLS AND ABBREVIATIONS

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, f, s	cylindrical coordinates
x', y'	transverse angles
δ	relative momentum deviation
$\mathcal{Z}$	6D canonical phase-space variables
ζ_i	ith component of $\mathcal{Z}$
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
$f(\mathcal{Z};s)$	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
$\varepsilon_i^{\mathrm{proj}}$	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	Alfvén current
$\mathcal{K}$	loss factor
$w(z)$	wake function
$W(z)$	wakefield of bunch

ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy (DOE), Office of Science, under Award No. DE-AC02-06CH11357 with Argonne National Laboratory. P. P also acknowledges support from DOE Awards No. DE-SC0018656 and No. DE-SC0022010 to Northern Illinois University and from the Center for Bright Beams via NSF Grant No. PHY-1549132.

REFERENCES

Adolphsen, Chris, Maura Barone, Barry Barrish, Karsten Buesser, and Philip Burrows, 2013, “The International Linear Collider Technical Design Report—Vol. 3.II: Accelerator baseline design,” arXiv:1306.6328.

Aiba, Masamitsu, Michael Ehrlichman, and Andreas Streun, 2015, “Round beam operation in electron storage rings and generalisation of mobius accelerator,” in *Proceedings of the 6th International Particle Accelerator Conference (IPAC 2015), Richmond, VA, 2015*, edited by Stuart Henderson, Todd Satogata, and Volker R. W. Schaa (JACoW, Geneva), p. TUPJE045.

Aiba, Masamitsu, and Jonas Kallestrup, 2020, “Theory of emittance exchange through coupling resonance crossing,” Phys. Rev. Accel. Beams **23**, 044003.

Akre, R., *et al.*, 2008,“Commissioning the linac coherent light source injector,” Phys. Rev. ST Accel. Beams **11**, 030703.

Alvarez, L. W., H. Bradner, J. V. Franck, H. Gordon, J. D. Gow, L. C. Marshall, F. Oppenheimer, W. K. H. Panofsky, C. Richman, and J. R. Woodyard, 1955, “The design of a proton linear accelerator,” Rev. Sci. Instrum. **26**, 111–133.

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., *et al.*, 2020, “Shaped beams from diamond field-emitter array cathodes,” IEEE Trans. Plasma Sci. **48**, 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,” J. Appl. Phys. **104**, 014901.

Antipov, S., L. Spentzouris, W. Liu, W. Gai, and J. G. Power, 2007, “Wakefield generation in metamaterial-loaded waveguides,” J. Appl. Phys. **102**, 034906.

Armborst, Felix, 2016, “Study and optimization of the transverse beam parameters before injection into the storage ring BESSY II,” master’s thesis (Humboldt-Universität zu Berlin), http://edoc.hu-berlin.de/18452/21241.

Arnold, V. I., 1978, *Mathematical Methods of Classical Mechanics* (Springer-Verlag, New York).

Asaka, Takao, *et al.*, 2017, “Low-emittance thermionic-gun-based injector for a compact free-electron laser,” Phys. Rev. Accel. Beams **20**, 080702.

Aschikhin, A., *et al.*, 2016, “The FLASHForward facility at DESY,” Nucl. Instrum. Methods Phys. Res., Sect. A **806**, 175–183.

Bane, K. L. F., P. Chen, and P. B. Wilson, 1985, “On collinear wake field acceleration,” IEEE Trans. Nucl. Sci. **32**, 3524–3526.

Bane, K. L. F., and G. Stupakov, 2012, “Corrugated pipe as a beam dechirper,” Nucl. Instrum. Methods Phys. Res., Sect. A **690**, 106–110.

Bane, Karl, and Gennady Stupakov, 2016, “Dechirper wakefields for short bunches,” Nucl. Instrum. Methods Phys. Res., Sect. A **820**, 156–163.

Baryshev, Sergey V., SergeyAntipov, 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," Appl. Phys. Lett. **105**, 203505.

Batygin, Y. K., 1993, "Beam intensity redistribution in a nonlinear optics channel," Nucl. Instrum. Methods Phys. Res., Sect. B **79**, 770–772.

Benson, S. V., *et al.*, 2018, in "Development of a bunched beam electron cooler based on ERL and circulator ring technology for the Jefferson Lab Electron-Ion Collider," *Proceedings of the 11th Workshop on Beam Cooling and Related Topics (COOL 2017), Bonn, Germany, 2017*, edited by V. R. W. Schaa, M. Marx, and M. Tilmont (JACoW, Geneva), pp. 72–76.

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., *et al.*, 2008, "Multifunctional micro-optical elements for laser beam homogenizing and beam shaping," in *Photon Processing in Microelectronics and Photonics VII*, SPIE Proceedings Vol. 6879, edited by David B. Geohegan, Frank Träger, Jan J. Dubowski, Andrew S. Holmes, Michel Meunier, Craig B. Arnold, and Hiroyuki Niino (SPIE—International Society for Optical Engineering, Bellingham, WA), 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," J. Appl. Phys. **126**, 033102.

Blaskiewicz, Michael, 2014, "Cooling of high-energy hadron beams," Annu. Rev. Nucl. Part. Sci. **64**, 299–317.

Blaskiewicz, Michael, 2019, "3D theory of the plasma cascade instability," arXiv:1910.13300.

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," in *Advances in Particle Physics*, Vol. 1, edited by R. L. Cool and R. E. Marshak (John Wiley & Sons, New York), pp. 71–134.

Brown, Karl L., 1999, "Single element optics," in *Handbook of Accelerator Physics and Engineering*, edited by Alexander Wu Chao and Maury Tigner (World Scientific, Singapore), Chap. 2.2.1.

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 Trans. Nucl. Sci. **22**, 2093–2097.

Burov, A., and V. Danilov, 1998, "An insertion to eliminate horizontal temperature of high energy electron beam," Fermilab Technical Report No. Fermilab-TM-2043.

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, "Calculation of the trajectory of cathode rays in an axisymmetric electromagnetic field," Ann. Phys. (Berlin) **386**, 974–993.

Carey, David C., 1987, *The Optics of Charged Particle Beams*, Accelerators and Storage Rings Vol. 6 (Harwood Academic Publishers, Chur, Switzerland) [ISBN: 978-3718605453].

Carli, C., G. Cyvoct, M. Giovannozzi, E. Metral, G. Métral, and R. Steerenberg, 2002, "Emittance exchange by crossing a coupling resonance," in *Proceedings of the 8th European Particle Accelerator Conference (EPAC 2002), Paris, 2002*, edited by T. Garvey, L. Rivkin, J. Le Duff, C. Petit-Jean-Genaz, P. Le Roux, and J. Poole (European Physical Society, Geneva), pp. 1157–1159.

Carlsten, B. E., 1989, "New photoelectric injector design for the Los Alamos National Laboratory XUV FEL accelerator," Nucl. Instrum. Methods Phys. Res., Sect. A **285**, 313–319.

Carlsten, B. E., 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, 2011, "Arbitrary emittance partitioning between any two dimensions for electron beams," Phys. Rev. ST Accel. Beams **14**, 050706.

Carlsten, Bruce E., Kip A. Bishofberger, Steven J. Russell, and Nikolai A. Yampolsky, 2011, "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, A. W., 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.

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, E. D., 1966, in *Perspectives in Modern Physics: Essays in Honor of Hans A. Bethe on the Occasion of His 60th Birthday, July 1966*, edited by R. E. Marshak (Wiley Interscience, New York), pp. xii + 673.

Courant, E. D., and H. S. Snyder, 1958, "Theory of the alternating-gradient synchrotron," Ann. Phys. (Amsterdam) **3**, 1–48.

Craievich, Paolo, and Alberto A. Lutman, 2017, "Effects of the quadrupole wakefields in a passive streaker," Nucl. Instrum. Methods Phys. Res., Sect. A **865**, 55–59.

Decking, W., *et al.*, 2020,"A MHz-repetition-rate hard x-ray free-electron laser driven by a superconducting linear accelerator," Nat. Photonics **14**, 391–397.

Deng, Haixiao, *et al.*, 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," Fiz. Plasmy **4**, 492–500.

Derbenev, Ya., 1998, "Adapting optics for high energy electron cooling," University of Michigan Report No. UM-HE-98-04.

Derbenev, Ya. S., J. Rossbach, E. L. Saldin, and V. D. Shiltsev, 1995, "Microbunch radiative tail-head interaction," TESLA-FEL Report No. 95-05.

Ding, Y., *et al.*, 2015, "Generating femtosecond x-ray pulses using an emittance-spoiling foil in free-electron lasers," Appl. Phys. Lett. **107**, 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," Nucl. Instrum. Methods Phys. Res., Sect. A **528**, 448–452.

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," U.S. Particle Accelerator School, course materials, University of California, Santa Cruz, January 2008, https://uspas.fnal.gov/materials/08UCSC/UCSC-High-Brightness.shtml.

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.

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" (unpublished), https://www.physics.umd.edu/dsat/docs/TOC12Nov2019.pdf.

Du, Ying-Chao, Wen-Hui Huang, and Chuan-Xiang Tang, 2012, "A new method to generate relativistic comb bunches with tunable subpicosecond spacing," Chin. Phys. C **36**, 151–155.

Duan, Zhaoyun, Xianfeng Tang, Zhanliang Wang, Yabin Zhang, Xiaodong Chen, Min Chen, and Yubin Gong, 2017, "Observation of the reversed Cherenkov radiation," Nat. Commun. **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, edited by Peter W. Hawkes (Elsevier, New York), pp. 1–44.

Duncan, C. J. R., 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 (W. A. Benjamin, New York).

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 *Proceedings of the XX International Linac Conference (Linac 2000), Monterey, CA, 2000*, edited by A. W. Chao, pp. 122–124, https://accelconf.web.cern.ch/l00/papers/MOB13.pdf.

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 *Proceedings of the 2001 IEEE Particle Accelerator Conference (PAC 2001), Chicago, 2001*, edited by P. Lucas and S. Webber (IEEE, New York), pp. 73–75.

Egerton, R. F., 2009, "Electron energy-loss spectroscopy in the TEM," Rep. Prog. Phys. **72**, 016502.

Emma, C., J. Wu, K. Fang, S. Chen, S. Serkez, and C. Pellegrini, 2014, "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., *et al.*, 2014, "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, Upton, NY, 1975*, edited by J. C. Herrera and A. M. Thorndike (Brookhaven National Laboratory Report No. BNL-20550, 1975), p. 486.

Ferrario, M., J. E. Clendenin, D. T. Palmer, J. B. Rosenzweig, and L. Serafini, 2000, "HOMDYN study for the LCLS RF photoinjector," in *Proceedings of the 2nd ICFA Advanced Accelerator Workshop on the Physics of High Brightness Beams*, edited by J. Rosenzweig and L. Serafini (World Scientific, Singapore), pp. 534–563.

Ferrario, M., *et al.*, 2007, "Direct Measurement of the Double Emittance Minimum in the Beam Dynamics of the SPARC High-Brightness Photoinjector," Phys. Rev. Lett. **99**, 234801.

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," Nucl. Instrum. Methods Phys. Res., Sect. A **1025**, 166051.

Flöttmann, K., 1997, "Note on the thermal emittance of electrons emitted by cesium telluride photo cathodes," Report No. TESLA FEL-1997-01.

Flöttmann, K., 2014, "Generation of sub-fs electron beams at few-MeV energies," Nucl. Instrum. Methods Phys. Res., Sect. A **740**, 34–38.

Flöttmann, K., T. Limberg, and P. Piot, 2001, "Generation of ultrashort electron bunches by cancellation of nonlinear distortions in the longitudinal phase space," Report No. TESLA-FEL-2001-06.

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., 2016, "Field electron emission theory," in *Proceedings of Young Researchers in Vacuum Micro/Nano Electronics (VMNE-YR), St. Petersburg, Russia, 2016*, edited by edited by N. V. Egorov and K. A. Nikiforov (IEEE, New York), p.1–8, 10.1109/VMNEYR.2016.7880403 .

Fowler, R. H., and L. Nordheim, 1928, "Electron emission in intense electric fields," Proc. R. Soc. A **119**, 173–181.

Fraser, J. S., R. L. Sheffield, E. R. Gray, and G. W. Rodenz, 1985, "High-brightness photoemitter injector for electron accelerators," IEEE Trans. Nucl. Sci. **32**, 1791–1793.

Fu, Feichao, *et al.*, 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., *et al.*, 2018, “Observation of High Transformer Ratio of Shaped Bunch Generated by an Emittance-Exchange Beam Line,” Phys. Rev. Lett. **120**, 114801.

Garault, Y., 1964, “Study of a class of guided electromagnetic waves: EH waves, applications to deflecting structures for traveling wave separators of relativistic particles,” 10.5170/CERN-1964-043.

Ginzton, E. L., W. W. Hansen, and W. R. Kennedy, 1948, “A linear electron accelerator,” Rev. Sci. Instrum. **19**, 89–108.

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,” Rev. Sci. Instrum. **18**, 347–356.

Goldstein, Herbert, 2002, *Classical Mechanics*, 3rd ed. (Addison-Wesley, Reading, MA).

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., 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, W. S., *et al.*, 2019, “Nanopatterned electron beams for temporal coherence and deterministic phase control of x-ray free-electron lasers,” arXiv:1906.01525.

Graves, William, *et al.*, 2018, “ASU compact XFEL,” in *Proceedings of the 38th International Free-Electron Laser Conference (FEL 2017), Santa Fe, 2017*, edited by Kip Bishofberger, Bruce Carlsten, and Volker R. W. Schaa (JACoW, Geneva), 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. E. Conde, D. S. Doran, W. Gai, and J. G. Power, 2018, “Preparations for installation of the double emittance-exchange beamline at the Argonne Wakefield Accelerator facility,” in *Proceedings of the 38th International Free Electron Laser Conference (FEL 2017), Santa Fe, 2017*, edited by K. Bishofberger, B. Carlsten, and V. R. W. Schaa (JACoW, Geneva), pp. 340–343.

Ha, G., M. E. Conde, and J. G. Power, 2019, “Arbitrary transverse profile shaping using transverse wigglers,” in *Proceedings of the 4th North American Particle Accelerator Conference (NAPAC 2019), Lansing, MI, 2019*, edited by Y. Yamazaki, T. Raubenheimer, A. McCausey, and V. R. W. Schaa (JACoW, Geneva), pp. 403–406.

Ha, G., M. E. Conde, J. G. Power, and E. E. Wisniewski, 2018, “CSR shielding effect in dogleg and EEX beamlines,” in *Proceedings of the 9th International Particle Accelerator Conference (IPAC ’18), Vancouver, British Columbia, Canada, 2018*, edited by S. Koscielniak, T. Satogata, V. R. W. Schaa, and J. Thomson (JACoW, Geneva), pp. 1498–1500.

Ha, G., J. G. Power, M. Conde, D. S. Doran, and W. Gai, 2017a, “Limiting effects in double EEX beamline,” J. Phys. Conf. Ser. **874**, 012061.

Ha, G., J. G. Power, M. Conde, D. S. Doran, and W. Gai, 2017b, “Preliminary simulations on chirpless bunch compression using double-EEX beamline,” in *Proceedings of the 8th International Particle Accelerator Conference (IPAC ’17), Copenhagen, 2017*, edited by G. Arduini, M. Lindroos, J. Pranke, V. R. W. Schaa, and M. Seidel (JACoW, Geneva), pp. 3862–3864.

Ha, G., *et al.*, 2017, “Precision Control of the Electron Longitudinal Bunch Shape Using an Emittance-Exchange Beam Line,” Phys. Rev. Lett. **118**, 104801.

Ha, Gwanghui, Manoel Conde, John Power, Jiahang Shao, and Eric Wisniewski, 2019, “Tunable bunch train generation using emittance exchange beamline with transverse wiggler,” in *Proceedings of the 10th International Particle Accelerator Conference, Melbourne, 2019* edited by Mark Boland, Hitoshi Tanaka, David Button, Rohan Dowd, Volker R. W. Schaa, and Eugene Tan (JACoW, Geneva), 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,” Rev. Sci. Instrum. **34**, 1094–1100.

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 *Proceedings of the 25th International Linear Accelerator Conference (LINAC ’10), Tsukuba, Japan, 2010*, edited by Masanori Ikegami, Volker R. W. Schaa, and Hiroyuki Sako (JACoW, Geneva), 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., *et al.*, 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. Newsl. **38**, 37–51, https://www-public.slac.stanford.edu/sciDoc/docMeta.aspx?slacPubNumber=slac-pub-11597.

Hyun, J., P. Piot, T. Sen, and J. C. Thangaraj, 2019, “Micro-bunching for generating tunable narrow-band THz radiation at the FAST photoinjector,” Nucl. Instrum. Methods Phys. Res., Sect. A **916**, 195–203.

Iogansen, L. V., and M. S. Rabinovich, 1960, “Coherent electron radiation in the synchrotron. II,” Sov. Phys. JETP **37**, 93–87, www.jetp.ras.ru/cgi-bin/dn/e_010_01_0083.pdf.

Irvine, W., and D. Bouwmeester, 2008, “Linked and knotted beams of light,” Nat. Phys. **4**, 716.

Jackson, John David, 1998, *Classical Electrodynamics* (Wiley, New York).

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 the 31st International Free Electron Laser Conference (FEL 2009), Liverpool, England*, edited by S. Waller, V. R. W. Schaa, M. Marx, L. Liljeby, and J. Poole, 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,” J. Appl. Phys. **108**, 094322.

Jensen, Kevin L., 2018, “A tutorial on electron sources,” IEEE Trans. Plasma Sci. **46**, 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,” Chin. Phys. C **39**, 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 *Proceedings of the 2nd International Conference on High-Energy Accelerators and Instrumentation, Geneva, 1959*, edited by L. Kowarski (CERN, Geneva), pp. 274–287.

Kashy, E., and B. Sherrill, 1987, “A method for the uniform charged particle irradiation of large targets,” Nucl. Instrum. Methods Phys. Res., Sect. B **26**, 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).

Khojoyan, M., M. Krasilnikov, and F. Stephan, 2014, “Studies on the application of the 3D ellipsoidal cathode laser pulses at PITZ,” in *Proceedings of the 5th International Particle Accelerator Conference, Dresden, 2014*, edited by C. Petit-Jean-Genaz, G. Arduini, P. Michel, and V. R. W. Schaa (JACoW, Geneva), p. THPRO043.

Kim, K.-J., 2020, “Simple formula to determine the parameters of *XY* emittance exchange,” Argonne National Laboratory Technical Report No. ANL/APS/LS-365.

Kim, K.-J., Z. Huang, and R. Lindberg, 2017, *Synchrotron Radiation and Free-Electron Lasers: Principles of Coherent X-Ray Generation* (Cambridge University Press, Cambridge, England).

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 International Free-Electron Laser Conference (FEL ’11), Shanghai, 2011*, edited by Z. Zhao and D. Wang (Shanghai Institute of Applied Physics, Shanghai), pp. 156–158.

Kim, Kwang-Je, 1989, “rf and space-charge effects in laser-driven rf electron guns,” Nucl. Instrum. Methods Phys. Res., Sect. A **275**, 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, and Andrew Sessler, 2006, “Transverse-longitudinal phase-space manipulations and correlations,” AIP Conf. Proc. **821**, 115–138.

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.

Klimontovich, Yu. L., 1995, *Statistical Theory of Open Systems* (Springer Netherlands, Dordrecht).

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,” J. Lightwave Technol. **34**, 20–35.

Kowalski, S., and H. Enge, 1972, “Beam rotator,” in *Proceedings of the 4th International Conference on Magnet Technology (MT-4), Upton, NY, 1972*, Vol. C720919, edited by Yvonne Winterbottom (Brookhaven National Laboratory, Upton, NY), p. 182.

Krafft, G. A., 1996, “Correcting M56 and T566 to obtain very short bunches at CEBAF,” AIP Conf. Proc. **367**, 46–55.

Kur, E., G. Penn, J. Qiang, M. Venturini, R. Wells, and A. Zholents, 2009, “Accelerator design study for a soft x-ray free electron laser at the Lawrence Berkeley National Laboratory,” Lawrence Berkeley National Laboratory Report No. LBNL-2670, pp. 40 and 41.

Kuske, Peter, and Felix Kramer, 2016, “Transverse emittance exchange for improved injection efficiency,” in *Proceedings of the 7th International Particle Accelerator Conference (IPAC 2016), Busan, Korea, 2016*, edited by Kyung Sook Kim, In Soo Ko, Kyung Ryul Kim, Volker R. W. Schaa, Christine Petit-Jean-Genaz, and Dong Eon Kim (JACoW, Geneva), p. WEOAA01.

Kuzmin, I., *et al.*, 2019, “Shaping triangular picosecond laser pulses for electron photoinjectors,” Laser Phys. Lett. **16**, 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**, 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**, 2678–2686.

Landau, L. D., and E. M. Lifshitz, 1969, *Mechanics, Vol. 1*, 2nd ed. (Pergamon Press, New York).

Laszczyk, K. U., 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., *et al.*, 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,” Sci. Rep. **11**, 796.

Lewellen, J. W., S. Biedron, A. Lumpkin, S. V. Milton, A. Nassiri, S. Pasky, G. Travish, and M. White, 1998, “Operation of the APS rf gun,” in *Proceedings of the 1998 Linear Accelerator Accelerator Conference (LINAC ’98), Chicago, 1998*, edited by C. E. Eyberger, R. C. Pardo, M. M. White, and K. M. Jaje (Argonne National Laboratory, Argonne, IL), pp. 863–865.

Li, R., 2001, “Sensitivity of the CSR self-interaction to the local longitudinal charge concentration of an electron bunch,” Nucl. Instrum. Methods Phys. Res., Sect. A **475**, 498–503.

Li, S., *et al.*, 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**, 1996–1998.

Li, Yuelin, and Kwang-Je Kim, 2008, “Nonrelativistic electron bunch train for coherently enhanced terahertz radiation sources,” Appl. Phys. Lett. **92**, 014101.

Li, Yuelin, and John W. Lewellen, 2008, “Generating a Quasiellipsoidal Electron Beam by 3D Laser-Pulse Shaping,” Phys. Rev. Lett. **100**, 074801.

Liouville, J., 1838, “Note on the theory of variation of arbitrary constants,” J. Math. Pures Appl. **3**, 342–349, https://gallica.bnf.fr/ark:/12148/bpt6k16382m/f351.item.n8.

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.

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.

Litvinenko, Vladimir N., *et al.*, 2021, “Plasma-cascade instability,” Phys. Rev. Accel. Beams **24**, 014402.

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**, 1467–1478.

Loisch, G., *et al.*, 2018, “Photocathode laser based bunch shaping for high transformer ratio plasma wakefield acceleration,” Nucl. Instrum. Methods Phys. Res., Sect. A **909**, 107–110.

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, 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.

Malyzhenkov, Alexander, and Alexander Scheinker, 2018, “Phase space exchange-based bunch compression with reduced CSR effects,” arXiv:1809.05579.

Mamun, M. A., *et al.*, 2018, “Production of magnetized electron beam from a DC high voltage photogun,” in *Proceedings of the 9th International Particle Accelerator Conference (IPAC ’18), Vancouver, British Columbia, Canada*, edited by S. Koscielniak, T. Satogata, V. R. W. Schaa, and J. Thomson (JACoW, Geneva), pp. 4567–4570.

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,” Phys. Plasmas **18**, 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**, 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,” Nat. Phys. **7**, 785.

McDonald, K. T., 1988, “Design of the laser-driven RF electron gun for the BNL accelerator test facility,” IEEE Trans. Electron Devices **35**, 2052–2059.

Meads, Jr., P. F., 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 non-linear tuning of beam expanders,” Nucl. Instrum. Methods Phys. Res., Sect. A **379**, 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*, edited by Loretta Lizama and Joe Chew (IEEE, New York), pp. 219–221.

Miginsky, S. V., 2009, “Emittance compensation of elliptical beams,” Nucl. Instrum. Methods Phys. Res., Sect. A **603**, 32–34.

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,” Appl. Phys. Lett. **107**, 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**, 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, 2010, “Resonant excitation of plasma wakefields,” AIP Conf. Proc. **1299**, 495–499.

Muggli, P., B. Allen, V. E. Yakimenko, J. Park, M. Babzien, K. P. Kusche, and W. D. Kimura, 2010, “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, J. B., S. Krinsky, and R. L. Gluckstern, 1997, "Longitudinal wake field for an electron moving on a circular orbit," Part. Accel. **57**, 9–64, https://cds.cern.ch/record/1120287.

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 nano-tip 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., *et al.*, 2021, "Accelerator and beam physics research goals and opportunities," arXiv:2101.04107.

Nagayama, Kuniaki, 2011, "Another 60 years in electron microscopy: Development of phase-plate electron microscopy and biological applications," J. Electron Microsc. **60**, 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., *et al.*, 2000, "Sustained Kilowatt Lasing in a Free-Electron Laser with Same-Cell Energy Recovery," Phys. Rev. Lett. **84**, 662–665.

Neil, G. R., *et al.*, 2006, "The JLab high power ERL light source," Nucl. Instrum. Methods Phys. Res., Sect. A **557**, 9–15.

Neil, V. Kelvin, and Andrew M. Sessler, 1965, "Longitudinal resistive instabilities of intense coasting beams in particle accelerators," Rev. Sci. Instrum. **36**, 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," J. Appl. Phys. **105**, 053304.

Neumann, J. G., 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," Nucl. Instrum. Methods Phys. Res., Sect. A **507**, 498–501.

Nguyen, D. C., and B. E. Carlsten, 1996, "Amplified coherent emission from electron beams prebunched in a masked chicane," Nucl. Instrum. Methods Phys. Res., Sect. A **375**, 597–601.

Nichols, K. E., *et al.*, 2020, "Demonstration of transport of a patterned electron beam produced by diamond pyramid cathode in an rf gun," Appl. Phys. Lett. **116**, 023502.

Nielsen, C. E., and A. M. Sessler, 1959, "Longitudinal space charge effects in particle accelerators," Rev. Sci. Instrum. **30**, 80–89.

Nielsen, C. E., A. M. Sessler, and K. R. Symon, 1959, "Longitudinal instabilities in intense relativistic beams," in *Proceedings of the 2nd International Conference on High-Energy Accelerators and Instrumentation, Geneva, 1959*, edited by L. Kowarski (CERN, Geneva), pp. 239–252.

Noether, Emmy, 1971, "Invariant variation problems," Transp. Theory Stat. Phys. **1**, 186–207.

Novokhatski, A., 2015, "Wakefield potentials of corrugated structures," Phys. Rev. ST Accel. Beams **18**, 104402.

O'Shea, Brendan, *et al.*, 2011, "Measurement of self-shaped ellipsoidal bunches from a photoinjector with postacceleration," Phys. Rev. ST Accel. Beams **14**, 012801.

Panofsky, W. K. H., and W. A. Wenzel, 1956, "Some considerations concerning the transverse deflection of charged particles in radio-frequency fields," Rev. Sci. Instrum. **27**, 967–967.

Paramonov, Valentin, and Klaus Flöttmann, 2019, "Fundamental characteristics of transverse deflecting field," arXiv:1806.11023.

Parkhomchun, V., and A. Skrinskii, 2000, "Electron cooling: 35 years of development," Phys. Usp. **43**, 433–452.

Pasquali, M., *et al.*, 2019, "Dynamic response of advanced materials impacted by particle beams: The multimat experiment," J. Dyn. Behav. Mater. **5**, 266–295.

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.

Penco, G., *et al.*, 2017, "Passive Linearization of the Magnetic Bunch Compression Using Self-Induced Fields," Phys. Rev. Lett. **119**, 184802.

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 *Proceedings of the 4th North America Particle Accelerator Conference (NAPAC 2019), Lansing, MI, 2019*, edited by Y. Yamazaki, T. Raubenheimer, A. McCausey, and V. R. W. Schaa (JACoW, Geneva), 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 *Proceedings of the 2nd Workshop on High Average Power and High Brightness Beams, Los Angeles, 2009* (unpublished), https://www.osti.gov/biblio/951117.

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., 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," Fermi National Accelerator Laboratory Report No. FERMILAB-CONF-14-142-APC, https://inspirehep.net/literature/1342754.

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 Conf. Proc. **1299**, 592–597.

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, 2011, "Observation of coherently enhanced tunable narrow-band terahertz transition radiation from a relativistic sub-picosecond electron bunch train," Appl. Phys. Lett. **98**, 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, 2011, "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., *et al.*, 2014, "Operation of an ungated diamond field-emission array cathode in a L-band radiofrequency electron source," Appl. Phys. Lett. **104**, 263504.

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 Conf. Proc. **1086**, 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 *Proceedings of the 7th International Particle Accelerator Conference (IPAC 2016), Busan, Korea, 2016*, edited by K. S. Kim, I. S. Ko, K. R. Kim, V. R. W. Schaa, Christine Petit-Jean-Genaz, and Dong Eon Kim (JACoW, Geneva), pp. 3020–3023.

Ranada, A. F., 1990, "Knotted solutions of the Maxwell equations in vacuum," J. Phys. A **23**, L815–L820.

Rand, R. E., and T. I. Smith, 1980, "Beam optical control of beam breakup in a recirculating electron accelerator," Part. Accel. **11**, 1–13, https://inspirehep.net/literature/159457.

Rao, Triveni, and David H. Dowell, 2013, *An Engineering Guide to Photoinjectors* (CreateSpace Independent Publishing Platform, Scotts Valley, CA) [ISBN: 978-1481943222].

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, New York).

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**, 57–61.

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," Nucl. Instrum. Methods Phys. Res., Sect. A **393**, 376–379.

Rosenzweig, J. B., A. M. Cook, R. J. England, M. Dunning, S. G. Anderson, and Massimo Ferrario, 2006, "Emittance compensation with dynamically optimized photoelectron beam profiles," Nucl. Instrum. Methods Phys. Res., Sect. A **557**, 87–93.

Rosenzweig, J. B., *et al.*, 2019, "Next generation high brightness electron beams from ultrahigh field cryogenic rf photocathode sources," Phys. Rev. Accel. Beams **22**, 023403.

Roussel, R., *et al.*, 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, 1996, "Calculation of energy diffusion in an electron beam due to quantum fluctuations of undulator radiation," Nucl. Instrum. Methods Phys. Res., Sect. A **381**, 545–547.

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," Nucl. Instrum. Methods Phys. Res., Sect. A **398**, 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," Nucl. Instrum. Methods Phys. Res., Sect. A **483**, 516–520.

Saldin, E. L., E. A. Schneidmiller, and M. V. Yurkov, 2002b, "Klystron instability of a relativistic electron beam in a bunch compressor," Nucl. Instrum. Methods Phys. Res., Sect. A **490**, 1–8.

Saldin, E. L., E. A. Schneidmiller, and M. V. Yurkov, 2004, "Longitudinal space charge-driven microbunching instability in the TESLA Test Facility linac," Nucl. Instrum. Methods Phys. Res., Sect. A **528**, 355–359.

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, https://slac.stanford.edu/pubs/slacreports/reports02/slac-r-121.pdf.

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, England).

Seeman, John T., 1992, "Observation and cures of wakefield effects in the SLC linac," AIP Conf. Proc. **255**, 339–346.

Seok, J., M. Chung, M. Conde, G. Ha, and J. Power, 2019, "Double-horn suppression in EEX based bunch compression," in *Proceedings of the 4th North America Particle Accelerator Conference (NAPAC 2019), Lansing, MI, 2019*, edited by Y. Yamazaki, T. Raubenheimer, A. McCausey, and V. R. W. Schaa (JACoW, Geneva), 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.

Seok, J., G. Ha, J. Power, M. Conde, and M. Chung, 2019, "Suppression of correlated energy spread using emittance exchange," in *Proceedings of the 10th International Particle Accelerator Conference (IPAC 2019), Melbourne, 2019*, edited by M. Boland, H. Tanaka, D. Button, R. Dowd, V. R. W. Schaa, and E. Tan (JACoW, Geneva), 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 *Proceedings of the 6th International Particle Accelerator Conference (IPAC 2015), Richmond, VA, 2015*, edited by Stuart Henderson, Todd Satogata, and Volker R.W. Schaa (JACoW, Geneva), 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,” Phys. Scr. **94**, 034004.

Shiltsev, Vladimir D., 2016, *Electron Lenses for Super-Colliders*, Particle Acceleration and Detection (Springer, New York).

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 2^n-pulse Michelson interferometer,” Appl. Opt. **37**, 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,” U.S. Particle Accelerator School, course materials, University of Texas at Austin, January 2012, https://uspas.fnal.gov/materials/12UTA/UTA-Cathode.shtml.

Smirnova, E. I., 2004, “Photonic band gap structures for accelerator applications,” AIP Conf. Proc. **737**, 309–319.

Smith, T. I., 1986, “Intense low emittance linac beams for free electron lasers,” in *Proceedings of the 13th International Linear Accelerator Conference, Stanford, CA, 1986*, edited by R. Campisi, L. Cheung, K. Johnston, and S. MacBride, pp. TH2–4, https://accelconf.web.cern.ch/l86/papers/th2-4.pdf.

Sørensen, A. H., 1988, “Liouville’s theorem and emittance,” in *CERN Accelerator School: General Accelerator Physics* (CERN, Geneva), pp. 18–36.

Spampinati, S., *et al.*, 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 Conf. Proc. **592**, 205–230.

Sun, Y., 2005, “Angular-momentum-dominated electron beams and flat-beam generation,” Ph.D. thesis (University of Chicago).

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.

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 *Proceedings of the 22nd IEEE Particle Accelerator Conference (PAC), Albuquerque, 2007*, edited by C. Petit-Jean-Genaz (IEEE, New York), pp. 3441–3443.

Sun, Y.-E., J. W. Lewellen, and D. W. Feldman, 2006, “Photothermal cathode measurements at the advanced photon source,” in *Proceedings of the 23rd Linear Accelerator Accelerator Conference (LINAC ’06), Knoxville, 2006*, edited by S. Henderson, pp. 349–351, https://accelconf.web.cern.ch/l06/PAPERS/TUP045.PDF.

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**, 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 the 15th IEEE International Conference on Particle Accelerators, Washington, DC, 1993*, Vol. 1, edited by S. T. Corneliussen and L. Carlton (IEEE, New York), pp. 426–428.

Tournois, Pierre, 1997, “Acousto-optic programmable dispersive filter for adaptive compensation of group delay time dispersion in laser systems,” Opt. Commun. **140**, 245–249.

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, edited by L. Lizama and J. Chew (IEEE, New York), pp. 1695–1697.

Tsoupas, N., *et al.*, 2007, “Uniform beam distributions at the target of the NASA Space Radiation Laboratory’s beam line,” Phys. Rev. ST Accel. Beams **10**, 024701.

Tsuno, K., and E. Munro, 1997, “Design of an omega filter for a 200 kV electron microscope,” Rev. Sci. Instrum. **68**, 109–115.

Uchida, Masaya, and Akira Tonomura, 2010, “Generation of electron beams carrying orbital angular momentum,” Nature (London) **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,” J. Phys. B **47**, 234009.

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,” J. Appl. Phys. **102**, 093501.

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**, 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**, 1563–1572.

Werner, Douglas H., and Kwon, Do-Hoon, 2014, Eds., *Transformation Electromagnetics and Metamaterials: Fundamental Principles and Applications* (Springer-Verlag, London).

Wideröe, R., 1928, "About a new principle for the production of high voltages," Arch. Elektrotechn. **21**, 387–406.

Wiedemann, H., 1999, *Particle Accelerator Physics I* (Springer, New York).

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**, 14922.

Williamson, John, 1936, "On the algebraic problem concerning the normal forms of linear dynamical systems," Am. J. Math. **58**, 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 Trans. Nucl. Sci. **32**, 3089–3091.

Wilson, P. B., 1989, "Introduction to wakefields and wake potentials," AIP Conf. Proc. **184**, 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 Conf. Proc. **1507**, 956–958.

Wollnik, H., and M. Berz, 1985, "Relations between elements of transfer matrices due to the condition of symplecticity," Nucl. Instrum. Methods Phys. Res., Sect. 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 *Proceedings of the 10th International Particle Accelerator Conference (IPAC 2019), Melbourne, 2019*, edited by M. Boland, H. Tanaka, D. Button, R. Dowd, V. R. W. Schaa, and E. Tan (JACoW, Geneva), pp. 3337–3340.

Xu, T., *et al.*, 2021, "A damping-ring-free electron injector for future linear colliders," in *Proceedings of the 12th International Particle Accelerator Conference (IPAC '21), Campinas, Brazil*, edited by L. Liu, J. Byrd, R. Neuenschwander, R. Picoreti, and V. R. W. Schaa (unpublished).

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.

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," J. Appl. Phys. **92**, 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), pp. 415–445.

Yuri, Yosuke, Mitsuhiro Fukuda, and Takahiro Yuyama, **2019**, "Formation of hollow ion beams of various shapes using multipole magnets," Prog. Theor. Exp. Phys. 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, *et al.*, 2016, "Tunable High-Intensity Electron Bunch Train Production Based on Nonlinear Longitudinal Space Charge Oscillation," Phys. Rev. Lett. **116**, 184801.

Zhao, Lingrong, *et al.*, 2018a, "Few-femtosecond electron beam with terahertz-frequency wakefield-driven compression," Phys. Rev. Accel. Beams **21**, 082801.

Zhao, Lingrong, *et al.*, 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," Nucl. Instrum. Methods Phys. Res., Sect. A **425**, 385–389.

Zholents, A. A., and M. S. Zolotorev, 2011, "A new type of bunch compressor and seeding of short wave length coherent radiation," Argonne National Laboratory Technical Report No. ANL/APS/LS-327.

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**, 8488–8492.

Zolotorev, M., E. D. Commins, and F. Sannibale, 2007, "Proposal for a Quantum-Degenerate Electron Source," Phys. Rev. Lett. **98**, 184801.