Annu. Rev. Nucl. Part. Sci. 2003. 53:387–429
doi: 10.1146/annurev.nucl.53.041002.110456

HIGH-CURRENT ENERGY-RECOVERING ELECTRON LINACS

Lia Merminga, David R. Douglas, and Geoffrey A. Krafft
Center for Advanced Studies of Accelerators, Thomas Jefferson National Accelerator Facility, Newport News, Virginia 23606; email: Merminga@JLab.org, Douglas@JLab.org, Krafft@JLab.org

Key Words energy recovery, superconducting radiofrequency, recirculating accelerator, high average current

■ **Abstract** The use of energy recovery provides a potentially powerful new paradigm for generation of the charged particle beams used in synchrotron radiation sources, high-energy electron cooling devices, electron-ion colliders, and other applications in photon science and nuclear and high-energy physics. Energy-recovering electron linear accelerators (called energy-recovering linacs, or ERLs) share many characteristics with ordinary linacs, as their six-dimensional beam phase space is largely determined by electron source properties. However, in common with classic storage rings, ERLs possess a high average-current-carrying capability enabled by the energy recovery process, and thus promise similar efficiencies. We discuss the concept of energy recovery and its technical challenges and describe the Jefferson Lab (JLab) Infrared Demonstration Free-Electron Laser (IR Demo FEL), originally driven by a 35–48-MeV, 5-mA superconducting radiofrequency (srf) ERL, which provided the most substantial demonstration of energy recovery to date: a beam of 250 kW average power. We present an overview of envisioned ERL applications and a development path to achieving the required performance. We use experimental data obtained at the JLab IR Demo FEL and recent experimental results from CEBAF-ER—a GeV-scale, comparatively low-current energy-recovery demonstration at JLab—to evaluate the feasibility of the new applications of high-current ERLs, as well as ERLs' limitations and ultimate performance.

CONTENTS

1. INTRODUCTION

Recently, there has been a growing interest within the accelerator physics community in accelerators using the idea of beam-energy recovery. This paper reviews the work done on energy recovery to date. At the most fundamental level, beam-energy recovery allows the construction of electron linear accelerators that can accelerate average beam currents similar to those provided by storage rings, but with the superior beam quality typical of linacs. Such an ability to simultaneously provide both high current and high beam quality can be broadly utilized in, for example, high-average-power free-electron laser sources designed to yield unprecedented optical beam power; light sources extending the available photon brilliance beyond the limits imposed by present-day synchrotron light sources; electron cooling devices—which would benefit from both high average current and good beam quality to ensure a high cooling rate of the circulating particles in a storage ring collider; or, possibly, as the electron accelerator in an electron-ion

collider intended to achieve operating luminosity beyond that provided by existing, storage-ring-based colliders.

We compare recirculating linacs to two more common types of accelerators: single-pass linacs and storage rings. We then discuss energy recovery conceptually and review the work to date on this technique. The design and operation of the highest existing average-beam-current ERL—the JLab IR Demo Free-Electron Laser (FEL)—is described in detail. From this characterization of the state of the art, we proceed with a review of ERL projects presently under consideration. In order to meet the performance goals in these proposals, some development of the technology is required. This effort is reviewed in two subsequent sections, the first devoted to the extrapolation of recovered beam energy and the second devoted to extrapolation of recovered beam current. We conclude with a list of outstanding issues in the field.

1.1. Traditional Types of Electron Accelerators

For many years, two traditional types of particle accelerators (Figure 1) have primarily been used for high-energy and nuclear physics research. Among the electron accelerators, the first class of accelerators comprises the high-energy electron linacs, perhaps best exemplified by the linear accelerator (linac) at the Stanford Linear Accelerator Center (SLAC) (1). In such accelerators, the electron beam has a definite beginning in a region known almost universally as the injector and a definite end in an electron beam dump. Usually, the beam propagates along a nearly straight line, and between the injector and dump there is a substantial length of beam-acceleration devices. These are known as radiofrequency (rf) cavities because they oscillate at frequencies in the rf region of the electromagnetic spectrum. These cavities impart energy to the electrons by the electromagnetic $\vec{E} \cdot \vec{v}$ interaction, the electromagnetic field in the cavity being chosen so that the electric field in the cavity, $\vec{E}$, is substantially collinear with and along the electron velocity vector $\vec{v}$ as the electron moves through the cavity.

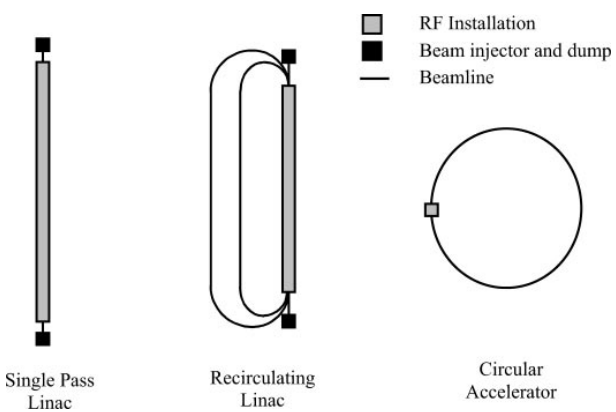

Figure 1 Main accelerator types.

Some main features of an electron linac are as follows: First, an individual beam electron resides in the accelerator only briefly, and particularly it resides in the accelerator for times that are short compared to any relevant radiation-driven emittance buildup times. Second, if a laser-driven photocathode gun is used as the electron source, it is relatively easy to load, or program, the beam current or beam polarization delivered to users by controlling the duration and polarization of the lasers that stimulate electron production at the injector. Third, the emittance, or transverse beam phase space area, of the electrons in a typical beam tends to be set by phenomena in the low-energy electron source region, and this emittance may be well preserved during the acceleration to high energy. Many highly developed procedures have been employed to control emittance at SLAC (2 and references therein), meaning that the emittance at the point of delivery depends mainly on performance in the injector. Fourth, the pulse duration, and more generally the longitudinal phase space distribution, is relatively easily manipulated by using standard beam-rf and electron beam optical techniques. It should be noted that having long distances between the end of the linac and the beam dump, for experiments or other purposes, is easy to arrange in a linear geometry.

The second class of high-energy electron accelerators is the synchrotron-like storage ring (3). Such storage rings have become commonplace, and the highest-luminosity electron colliding beam machines are among them. For the purposes of this review, it is sufficient to explain storage rings in simple terms; one should realize that the design, operation, and performance of storage rings are immensely complicated tasks, not all of the details have even yet been worked out entirely. In an electron storage ring, the electrons are bent on a roughly circular orbit. Because transversely accelerated electrons radiate copious amounts of electromagnetic radiation, to achieve a long-term equilibrium, it is necessary to supply energy to the circulating electrons. Energy is typically supplied, as in linacs, with rf cavities that subtend a small portion of the total machine circumference. After the beam is injected into the ring, the electrons rapidly settle into an equilibrium and the synchrotron radiation losses are made up by the energy transferred from rf cavities to beam. The relevant timescale is of order one radiation damping time $(E/\Delta E)t_{\mathrm{rev}}$, where E is the beam energy, ΔE is the energy loss per turn, and t_{rev} is the time it takes to make one revolution.

The characteristics of the equilibrium are very interesting and point to the main limitations of storage rings. First, the electrons are naturally bunched. The bunches migrate to the phase-stable accelerating phase (there is one such phase for each rf period) and onto a closed orbit within the storage ring. Second, the equilibrium beam emittance, and hence the beam sizes in the storage ring, are set by a competition between the radiation damping, which tends to drive the electrons onto the closed orbit at the correct accelerating phase, and the quantized radiation emission, which tends to excite transverse and longitudinal oscillations. Formulas for the equilibrium size are given elsewhere (3); here it is sufficient to point out that both the emittances and the equilibrium pulse length in an electron storage ring can not be arbitrarily small. It should be noted that much of the increase in the luminosity of recent storage-ring colliders has come from changing the ring

design to allow substantially reduced beam sizes at the collision point, but also from increasing the beam current by filling all useful accelerating phases with the maximum charge per bunch possible.

Third, the lifetime of the beam in a storage-ring collider is set by collective phenomena associated with the beam-beam interactions. Even if there were no beam-beam collisions as in storage-ring light sources, the lifetime would be limited by single intrabeam electron-electron scattering, known as Touschek scattering. Finally, storage rings provide high efficiency. In equilibrium all the energy going into the electron beam is radiated by synchrotron radiation. The ratio of the damping time to the revolution time, a number typically of order 10^{-3}, gives a measure of the efficiency of the energy transfer. High efficiencies have allowed storage rings to operate at much higher average currents than have been possible in linacs.

1.2. Beam Recirculation

Recirculating linacs are accelerators in which, as in linacs, there is a substantial rf system accelerating the beam and the beam has a definite beginning and ending (i.e., there is no closed or equilibrium orbit), but, as in a storage ring, the beam goes through the accelerating cavities more than once (Figure 1). As discussed below, such a hybrid arrangement allows one accelerator to feature some advantages of both of the traditional arrangements.

Early work on beam recirculation started with the development of microtrons. These devices use a resonance between the rf frequency of an accelerating cavity and the relativistic cyclotron frequency in a uniform magnetic field to develop a phase-stable accelerator in which the electron orbit enlarges on each pass through the cavity. To incorporate more substantial acceleration in the recirculation path, and to go to higher energy, racetrack microtrons were developed (4–6). As beam energy increased with time, it was natural to evolve away from the limits imposed by the large end magnets of racetrack microtrons and in a direction where different energy beams had completely different orbits, as in the classical recirculating linac. Jefferson Laboratory's (JLab's) CEBAF machine is the largest recirculating linac, achieving almost 6 GeV beam energy and 200 μA of beam current (7).

The reader has perhaps noted that beam recirculation has been applied mainly to electron accelerators. As discussed in Reference 7, it is essential in a linac, which accelerates with rf fields, to maintain the proper phase relationships between the first beam pass and the higher beam passes. Maintaining the proper phase relationships is particularly easy for high-energy electrons: all of the beam passes are traveling at nearly the velocity of light, largely independent of beam energy. If the proper phase relationship between the various beam passes exists at one cavity in the linac, then it will be proper for all the cavities in the linac. In ion accelerators, where the velocity does depend on the beam energy, more complicated schemes will be needed to recirculate a beam through an rf linac.

Until now, the usual motivation for beam recirculation, extending even to the CEBAF design, has been economics. Because beam recirculation systems tend to

be much cheaper to build than additional rf linac length, it makes sense to reuse the expensive beam-acceleration systems as many times as possible to achieve the highest possible energy from a given rf installation. In several instances (8, 9), front-to-back electron recirculation was used to increase the energy reach of an existing linac.

In the future, it is likely that electron recirculation will be applied to build recirculating linacs because of their superior beam quality. Recirculating linacs share with linacs the ability to accelerate and preserve the very-low-emittance injector beams (Figure 1). Because the transit time is short compared to a typical radiative emittance buildup time, no equilibrium is established as in a storage ring, implying that the emittance delivered to the end user may be smaller. Also, as in linacs, one can manipulate the longitudinal phase space of the electron beam to deliver very short beam pulses to the end user. The minimum pulse length is no longer set by radiative effects but by the ability to generate, and precisely manipulate, the longitudinal phase space of the electron beam, as shown many years ago at CEBAF (7). Such advantages of the recirculating linac might not be so interesting if not for the application of beam-energy recovery. This idea has allowed one to conceive of recirculating linacs with high average currents and efficiencies approaching those in storage rings.

1.3. Beam Energy Recovery

An early suggestion for the use of beam energy recovery, and in fact the earliest use of the words "energy recovery" of which the authors are aware, appears in 1965 (10). In this paper, Tigner explored the possibility of basing a particle physics colliding-beam machine on an energy-recovery linac at a time when the alternative (and ultimately successful) storage-ring colliders were beginning to gather steam. As discussed below, technical realization of accelerators with energy recovery started more than a decade later.

The basic idea of energy recovery is straightforward to those who are accustomed to the notion of recycling. As a familiar example, the recycling of aluminum beverage containers reduces the need to mine and process additional aluminum or to otherwise move aluminum stock into the distribution channel, while supporting beverage distribution in an economy where the demand for aluminum beverage containers is constant. Theoretically, if no aluminum were lost in the recycling program, a constant aluminum stock could support the needs of such beverage distribution. If there are losses, then aluminum must be put into the distribution channel at a level that the losses are recovered to achieve equilibrium.

By analogy, electron-beam energy, which is produced by accelerating electrons in electromagnetic fields of particle accelerators, may also be recycled. The idea of energy recovery in a recirculating rf linac is based on the fact that the rf fields, by proper choice of the time of arrival of the electron bunches in the linac beam, may be used to both accelerate and decelerate the same beam. Consider the simplest case of a single recirculation. A beam is injected into the linac and timed to accelerate

on the first pass through the linac. If the recirculation path is chosen to be precisely an integer plus $^1/_2$ rf wavelengths, then on the second pass through the linac, the beam is actually decelerated by the same rf field that accelerated it on the first pass. For cavities within the recirculation loop, energy is directly transferred, via the rf field, from decelerating beam to accelerating beam. The key point is that these rf power systems do not need to provide the energy to accelerate the first-pass beam. Indeed, the rf power draw becomes almost completely independent of the beam current. An added benefit of the energy-recovery process is that the beam-dump design is much simpler because the beam is being dumped at much lower energy, and under certain circumstances, the dump radioactivation may be reduced.

To quantify the efficiency of energy-recovering linacs (ERLs), we use the concept of "rf to beam multiplication factor," defined as $\kappa \equiv P_{\text{beam}}/P_{\text{RF}}$, the ratio of the beam power at point of use to the power incident to the rf cavities accelerating the beam (G.A. Krafft, L. Merminga, unpublished work). For an electron beam of average current I_b injected into an ERL at injection energy E_{inj} and accelerated to a final energy E_f, in the limit of perfect energy recovery (exact cancellation of the accelerating and decelerating beam vectors), the multiplication factor is equal to

$$\kappa = \frac{P_{\text{beam}}}{P_{\text{RF}}} \simeq \frac{I_\text{b} E_\text{f}}{I_\text{b} E_{\text{inj}} + P_{\text{rf,linac}}}, \qquad 1.$$

where the required rf power consists of two terms: $I_\text{b} E_{\text{inj}}$, the power required to accelerate the beam current I_b in the injector, which is not energy recovered, and $P_{\text{rf,linac}}$, the rf power required to establish the electromagnetic fields in the ERL cavities, which is no longer dependent on the beam current because the positive and negative current are balanced in the ERL. Note that the multiplication factor increases with average beam current, and asymptotically approaches a value that is equal to the ratio of final to injected beam energy, $E_\text{f}/E_{\text{inj}}$. The higher the beam current, the higher is the overall system efficiency.

The introduction of the multiplication factor leads us now to consider another technical aspect of present and future ERLs: the application of superconducting rf technology. Cavities made of superconducting materials are not perfect conductors at rf frequencies. However, due to large efforts extending over many decades, superconducting cavities presently can be manufactured in substantial quantities with accelerating gradients exceeding 20 MV/m and with accelerating mode quality factors approaching 10^{10}, i.e., with minuscule wall losses. Continuous wave (cw) normal-conducting systems have been built with gradients of a few MV/m at quality factors at several tens of thousands. Clearly, the idea of balancing positive and negative beam currents can be applied advantageously to either superconducting or normal-conducting linac systems, and so the recovery factors for both types of systems can be very similar.

What distinguishes the ultimate performance limits of superconducting and normal-conducting ERLs can be seen using the concept of the multiplication factor. In the equation above, consider the term $P_{\text{rf,linac}}$, the incident power needed

to establish rf fields in the linac cavities. If the rf power required to establish the electromagnetic field in the linac cavities in an ERL no longer depends on the beam current, what does it depend on? At minimum, to establish a given accelerating gradient in an rf cavity one must make up the wall losses, and in normal-conducting systems this is the main load. As discussed below, for rf control reasons, it is necessary to reduce the effective, or "loaded," quality factor of the superconducting cavities to several 10^7, meaning the effective cavity losses (actually most of the rf power goes to loads), or the $P_{\mathrm{rf,linac}}$ term, is only about 1000 times better for superconducting cavities than for normal-conducting cavities. This factor of 1000 is reflected directly in the overall multiplication factors of the best ERLs: the best normal-conducting systems have multiplication factors of order 0.1 or less, whereas superconducting ERL proposals tend to reach several hundred. For this reason, most ERL designs are based on the use of superconducting cavities, but there is a strong effort in Russia to apply normal-conducting cavities in a multiturn, energy-recovering accelerator (11).

2. HISTORICAL DEVELOPMENT OF ENERGY-RECOVERING LINACS

2.1. Radiofrequency Superconductivity and Recirculating Linacs

Because they have become such an essential part of many current plans, it is worthwhile to discuss in more detail the development of superconducting rf (srf) systems and their applications to particle accelerators. Two main types of projects have driven the development of srf accelerating systems: applications to superconducting linacs and applications as storage-ring rf cavities. We briefly discuss the two broad categories sequentially, but the reader should recognize that development in general has advanced through contributions from both categories. Many of the efforts proceeded in parallel and benefited from collaboration and healthy competition.

The earliest efforts to build superconducting cavities were tied to projects building high-duty-factor linacs for nuclear physics research. The earliest example of such an accelerator was Stanford University's Superconducting Accelerator (SCA) (12). This 50-MeV linac could accelerate several hundred μA of beam current. The SCA was followed by the University of Illinois's Microtron Using a Superconducting Linac (MUSL) accelerators, which used beam recirculation in a racetrack microtron arrangement (13, 14). This device ultimately achieved 80 MeV and 10 μA of beam current. Next came the S-DALINAC at Technical University Darmstadt (15, 16), followed closely by JLab's CEBAF machine (7). Both S-DALINAC and CEBAF use isochronous beam recirculation: The S-DALINAC runs typically at beam energies of up to 87 MeV in a three-pass configuration and up to 50-μA beam current, and the CEBAF machine, a much larger project, achieves almost

6 GeV at 200 μA in a five-pass machine. The next larger superconducting electron linac is likely to be DESY's TeV-Scale Linear Accelerator (TESLA), being developed as a particle physics linear collider and an X-ray science free-electron-laser source. Many of the ERL projects propose utilizing the TESLA cavity designs as a starting point, for the obvious reason that these superconducting cavities are perhaps closest to providing off-the-shelf availability.

Superconducting cavities were also developed for storage-ring applications, mainly at the behest of the particle physics community. Cornell University's Laboratory for Nuclear Science (now Laboratory for Experimental Particle Physics) has been a leader in the field and produced the first superconducting cavity to actually accelerate the beam in a storage ring, CESR (17). Larger-scale efforts have proceeded at Cornell as the CESR luminosity was progressively upgraded, at CERN as a part of the LEP project, and at KEK's *B* factory. During its final year, the LEP project was the largest single installation of superconducting accelerating structures, with a total installed rf voltage of 2 GV.

The requirements that make superconducting cavities highly desirable are (*a*) cw or other high-duty-factor operation, (*b*) the highly efficient coupling of the energy into the electron beam, (*c*) the high-average-beam current, and (*d*) the reduction in the length of the accelerator. Although we do not discuss this fact in detail, superconducting cavities tend to generate reduced intrabeam collective effects compared to normal-conducting cavities, making them desirable even in low-duty-factor applications where a large charge bunch is to be accelerated.

2.2. Early Work on Energy Recovery

The earliest technical realization of beam-energy recovery was stimulated by work in nuclear medicine. In an effort to obtain compact, high-efficiency, and low-cost electron accelerators for medical applications, reflexotron accelerators were invented, apparently without knowledge of Tigner's work, and developed at Chalk River Nuclear Laboratories (18). In this compact normal-conducting linac, 5–25-MeV beam energy was achieved out of a 13-MeV, 1.6-m rf structure that had the electron beam current bent back and circulated in the opposite direction to the first-pass accelerating current. By changing the circulating path length, the beam energy was varied. At maximum, the beam energy was almost doubled, or at minimum a 5-MeV beam could be extracted. Clearly, when the linac was operated at 5 MeV, 62% of the beam energy was recovered (8 MeV out of 13 MeV). Here, beam-energy recovery was merely a result of the accelerator design when operated at low energy. No detailed rf measurements were made to address energy recovery.

More pointed realizations of energy-recovering accelerators took another decade and were performed by several groups almost simultaneously. In 1985, a group at MIT performed studies of "same-cell" energy recovery (i.e., energy recovery in which both accelerating and decelerating beams traverse the same accelerating cavity) at the Bates linac, which consisted of normal-conducting SLAC-type rf accelerating structures (8). Recirculation loops had been added in this machine

mainly to upgrade the beam energy. The beam was accelerated to about 400 MeV and decelerated to a final energy near 23 MeV, at comparatively high macropulse current.

In July of 1986, same-cell energy recovery was demonstrated at Stanford University's SCA (9). By this time, the SCA was principally used as an FEL driver because of the good quality of the accelerated beam. A single pass through the superconducting linac yielded sufficient energy to produce an IR FEL. In order to produce shorter wavelength FEL radiation, the beam was recirculated through the SCA on a second pass. In this configuration, the superconducting linac took the beam energy from 5 to 50 MeV on the first pass, was recirculated along a path whose length was an integral number of rf wavelengths, and was accelerated from 50 to 95 MeV on the second pass through the linac. One hundred fifty μA of beam current was provided by a stream of 12.5 pC bunches at a repetition rate of 11.8 MHz.

Because the beam-recirculation system allowed the path length to be varied through a full rf wavelength, choosing to shift the path length by half an rf wavelength allowed energy recovery to proceed. Detailed rf power measurements and comparisons between accelerating and energy-recovering modes, and between beam present and absent on a second pass, showed that only 10% of the rf power needed to accelerate a single beam pass was needed to maintain the rf field at the same gradient when the same current was recovered. In speculating about future applications to increase the efficiency of FELs by applying energy recovery, the increased difficulty of recovering a spent FEL beam was noted.

Shortly afterward, a group from Los Alamos (19) used rf means to take energy out of a spent beam from an FEL and recycled this energy by rf means to accelerate fresh beam. Energy recovery was accomplished through deceleration of the beam in rf-excited structures, which were coupled to the accelerating structures through resonant bridge couplers. The rf power generated by the decelerating beam was shared with the accelerators through the couplers. Measurements of the electron transport were performed during decelerations greater than 70% from 21 MeV to 5 MeV.

From 1990 to 1994, a beam-recirculation experiment was carried out on the CEBAF srf injector at JLab (20). This injector was capable of accelerating in excess of 200 μA beam current from 5 to 50 MeV, the beam being a continuous stream of 0.12 pC bunches at 1497 MHz. This experiment was primarily developed to demonstrate beam stability against beam-breakup (BBU) instability; this instability was most severe in energy-recovered recirculation because the average beam energy is lowest. The best performance obtained in this device was 30 μA in energy-recovery mode and between 64 and 215 μA in accelerating mode, the current depending on the beam optics of the recirculator. Up to $45/50 = 90\%$ of the beam energy was recovered, and because the srf systems were not optimized for energy-recovered operation and the beam current recovered was so low, the multiplication factor was only 0.2.

When the design of the CEBAF accelerator changed to a superconducting recirculating linac, it was pointed out that a primary benefit of a recirculating linac is the exceptional beam quality possible at high energy. Superior beam quality

is a primary prerequisite for building FELs. Investigations on the use of an srf accelerator as the electron-beam "driver" for an FEL were completed early in the CEBAF project (21). Considerable interest arose during the construction phase of CEBAF in using the accelerator in this manner (22–24), but attention soon shifted to the concept of a stand-alone srf-based FEL driver accelerator (25). Energy recovery became a standard part of these proposals because the overall system efficiency was higher, because the need to develop new, higher-power rf systems was avoided, and because the beam-dump problem was considerably reduced. In 1997, the Infrared Demonstration Free Electron Laser (IR Demo FEL) was funded, leading to the recent achievement of greater than 2-kW-cw infrared light from an FEL oscillator (26) and providing the most substantial demonstration of energy recovery to date for an average current of 5 mA. This accelerator was the first to explicitly incorporate beam-energy recovery as a fundamental component of machine design. Good energy recovery was essential for obtaining the high average beam (both laser beam and electron beam) power.

3. THE JEFFERSON LABORATORY INFRARED DEMONSTRATION FREE-ELECTRON LASER

Two pivotal technological developments led to a new design paradigm for FELs. Progress toward a successful large-scale implementation of srf in CEBAF allowed the acceleration of a high-repetition-rate cw electron beam. Simultaneously, advances in the design and construction of high-brightness DC electron sources made available cw beams of high-quality micropulses—with small transverse and longitudinal emittance—containing several tens of picocoulombs of charge. The formation of an FEL drive beam with peak currents of several tens of amperes was thus also, in principle, possible. The confluence of these developments led to a new approach for high-average-power FELs. Whereas previous systems had (unsuccessfully) attempted to produce high-average-output powers by using low repetition rates (tens of Hertz) and high peak current (hundreds of amperes) provided by normal-conducting linacs (with or without energy recovery), a new approach, based on a modest micropulse charge in a high-brightness beam (to produce moderate peak currents) with a very high repetition rate (MHz) has successfully driven FEL operation at kW powers. This approach provides promise for scaling performance to tens or even hundreds of kilowatts.

The JLab IR Demo FEL, its 10-kW upgrade, and extrapolation of this system to 100-kW levels are discussed below.

3.1. JLab IR Demo FEL System Design

The JLab IR Demo FEL (Figure 2, see color insert) was decommissioned in 2001 after three years of successful testing and operation (27).

The FEL—a high-repetition-rate, low-extraction-efficiency, optical resonator—produced over 2 kW of tunable light in a 3–6-μm wavelength range. It was driven by a 35–48-MeV, 5-mA–based cw energy-recovering srf electron linac. The driver accelerator design was constrained by the need for low energy, high average current, and a demand for stringent beam control at the wiggler and during energy recovery. These requirements were driven by the need for six-dimensional phase-space management, the existence of potentially deleterious collective phenomena (space charge, wakefields, BBU, and coherent synchrotron radiation), and interactions between the FEL and the accelerator rf system.

The FEL had an 8-m-long optical cavity resonator and used moderate gain and output coupling, low extraction efficiency and micropulse energy, and a high repetition rate to avoid high single-bunch charge while producing high average power. This paradigm ideally combined with srf technology, allowing cw operation, and motivated use of energy recovery to alleviate rf system demands. The system architecture thus imposed two requirements on the driver accelerator: First, the delivery to the wiggler of an electron beam with properties suitable for the FEL interaction, and second, the recovery of the drive-beam power after the FEL.

The first requirement reflects the needs of the FEL system itself. Table 1 gives the design parameters. The nominal FEL extraction efficiency produced with these parameters was greater than 0.5%. The micropulse energy was modest; high output power was achieved through cw operation at the 20th subharmonic of the rf accelerating frequency. The energy-recovery requirement was met by decelerating the beam after the FEL so as to drive the rf cavities. Because the full energy spread after the wiggler exceeded 5%, this created a need for a large acceptance transport system.

The system requirements mentioned above coupled to many phenomena and constraints. Phase-space requirements at the FEL demanded transverse and

TABLE 1 System parameters of the JLab IR Demo FEL

Parameter	**Nominal**	**Achieved**
Beam energy at wiggler	42 MeV	42 MeV
Average beam current	5 mA	5 mA
Bunch charge	60 pC	60–135 pC
Bunch repetition rate	74.85 MHz	18.7–74.85 MHz
Normalized emittance (rms)	13 mm-mrad	5–10 mm-mrad
Bunch length at wiggler (rms)	400 fs	400 fs
Peak current	60 A	60 A
FEL extraction efficiency	$>$0.5%	$>$1%
δp/p before wiggler (rms)	0.5%	$<$0.25%
δp/p after wiggler (full)	5%	6%–8%
cw FEL power	1 kW	2.13 kW

longitudinal phase-space matching during acceleration and transport to the wiggler. Similarly, the machine had to provide adequate transverse beam size control while managing the large longitudinal phase space. Such transport and conditioning of the beam was to be performed in the presence of a number of potential collective effects driven by the high current and low energy. To avoid space-charge-driven beam quality degradation, moderately high injection energy was needed (28). BBU and other impedance-driven instabilities were to be avoided (29). Management of coherent synchrotron radiation (CSR) was needed to preserve beam emittance (30–33). Assurance of rf stability was necessary, particularly in transient regimes such as FEL turn-on and initiation of energy recovery (34). Figure 2 illustrates the system concept, which successfully addressed these issues. The schematic shows the 10-MeV injector, a single eight-cavity JLab cryomodule accelerating by up to ∼40 MeV, an FEL insertion, and energy-recovery transport from wiggler through module to a beam dump. All acceleration was performed using standard CEBAF 1.497-GHz five-cell cavities. A summary of the function and performance of each section follows.

3.1.1. INJECTOR The electron source was a DC photocathode gun nominally producing 60 pC bunches at 320 keV with repetition rates of up to 75 MHz (35). Immediately following the gun, a room-temperature buncher compressed the initial electron pulse, which was then captured by a two-cavity CEBAF cryounit and accelerated to 10 MeV. A four-quadrupole telescope matched beam envelopes to the linac acceptance across a three-bend achromat such that beam position and angle at the exit are independent of the beam energy. RF component phases were adjusted to produce, in concert with the injection line momentum compaction, a parameter which quantifies the variation of path length with energy, a long (∼3 psec rms) bunch with low relative momentum spread (∼0.1% rms) at the entrance of the linac. Injected beam quality depended on gun operating voltage and charge per bunch; typical normalized emittances for 320 kV operation were of order 5–10 mm-mrad (36, 37).

3.1.2. LINAC The linac accelerated the injected beam from 10 MeV to 35–48 MeV using a single high-gradient eight-cavity JLab cryomodule. By accelerating at a phase 8° off crest, a phase/energy correlation was imposed on the longitudinal phase space; this was used downstream for bunch compression. The rf cavities also provided transverse focusing, assisting in beam envelope management. Immediately after the cryomodule, a small dipole was used to separate the accelerated and energy-recovered beams. The low-energy beam was directed to a beam dump; the effect of this bend on the full-energy beam was corrected by a subsequent pair of small bends.

3.1.3. FEL INSERTION The FEL was located immediately beyond the linac. Because this was prior to recirculation bending, it reduced CSR degradation of beam quality and allowed a low-power “straight ahead” operational (non–energy-recovering) mode before the recirculator was fully installed; this remained a

useful diagnostic configuration throughout the machine's operational history. A quadrupole telescope (two triplets) matched beam envelopes from module to wiggler. A four-dipole achromatic chicane between the triplets separated optical-cavity and electron-beam components while compressing the bunch length. The chicane geometry was constrained by the tolerable momentum compaction. Larger chicanes could have provided more space but would have led to higher compactions, with more time-of-flight jitter. To maintain FEL pulse/drive beam synchronism, the chicane transport matrix element M_{56}, which relates the path length deviation Δl to the energy deviation $\Delta p/p$, $\Delta l = M_{56}\Delta p/p$, was restricted to –0.3 m.

The match from module to wiggler, by virtue of rf focusing, depended on linac energy gain. It was therefore adjusted operationally to compensate for gross (several MeV) energy changes. After the wiggler, the electron beam (full momentum spread $>5\%$) was matched to the recirculation transport using a second quadrupole telescope. This avoided beam-envelope mismatch, large spot sizes, aggravated optical aberrations, error sensitivities, and potential beam loss. As in the linac-to-wiggler transport, a dipole chicane embedded in the telescope moved the electron beam off the optical cavity axis; this chicane also lengthened the electron bunch, reducing peak currents and alleviating potential wakefield and CSR effects. Simulations and experience with the machine indicated that space charge effects were not significant above 20–25 MeV; analysis of system performance and operational tuning was therefore possible using single-particle transport models.

3.1.4. RECIRCULATOR/ENERGY-RECOVERY TRANSPORT Following the FEL insertion, the electron beam with greatly increased momentum spread was transported through a recirculation arc to the linac for energy recovery. This recirculator provided both transverse beam confinement and longitudinal phase-space conditioning. Bending was provided by achromatic and nominally isochronous end loops based on an MIT-Bates design (38). Dipole parameters (bend and edge angles) and drift lengths were set to provide $M_{56} = 0$ from wiggler to reinjection point, and, across each end loop, betatron stable motion in the horizontal plane (with a tune of $5/4$) and imaging transport vertically (–I transfer matrix). The end loops were joined by a periodic alternating gradient focusing system with focusing strength selected to ensure that aberrations over the full arc were suppressed.

Beam path length through the recirculator was adjusted using steering dipoles adjacent to the large 180° dipoles and was used to set the phase of the energy-recovered beam with respect to the module rf fields. Each end loop had four quadrupole and four sextupole magnets provided for operational control of dispersion and momentum compaction. A single family each of quadrupoles and sextupoles (adjacent to the 180° bends) was used to modify the linear and quadratic momentum compactions from wiggler to reinjection, so as to compensate the slope and curvature of the rf waveform during energy recovery. This allowed simultaneous recovery of rf power from the electron beam and compression of the beam energy spread at the dump.

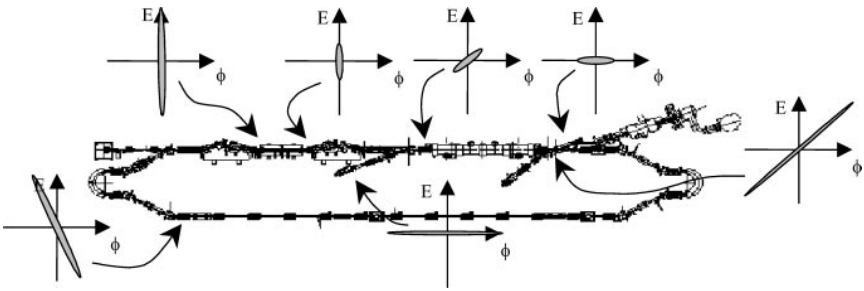

Figure 3 Longitudinal matching scenario in the JLab IR Demo FEL, showing phase versus energy diagrams at critical locations.

The second end loop delivered the longitudinally conditioned beam to the linac axis, where it was betatron matched to the linac acceptance using a four-quadrupole telescope and merged with the injected beam using a small achromatic three-dipole chicane.

3.2. Longitudinal Matching

Key to the operation of this device is the use of bunch-length compression (to create high peak current for FEL gain) and energy recovery (to provide rf power required for acceleration of high average currents) (39, 40). Figure 3 illustrates the longitudinal matching scenario employed in the system. The individual phase-energy plots indicate the orientation of the longitudinal phase space at key locations around the machine.

The injector provided a long, small-momentum spread bunch ($\sim$2.5 psec rms $\times$ 15 keV rms), which was accelerated off-crest in the linac. This imposed a phase-energy correlation, generating $\sim$0.25% momentum spread—about 100 keV at 40 MeV—over an rms bunch length. The momentum compaction of the chicane upstream of the wiggler rotated this slewed phase space upright, providing a short bunch (0.4 psec rms) at the wiggler. The FEL interaction did not affect bunch length but did generate a large full-momentum spread. This is evident in Figure 4, which shows the beam at a dispersed point (a point at which the variation of position with momentum, η, is equal to 0.4 m) in the chicane immediately downstream of the wiggler, without lasing (*right*: full momentum spread $\sim$1%, or 400 keV) and with lasing (*left*: full momentum spread $\sim$5%, or 2 MeV). This is indicative of the rather large acceptance required of the recirculator. The recirculator momentum compaction was used to rotate the bunch so that an appropriate phase-energy correlation occurs at reinjection.

The recirculator path length was adjusted by using the aforementioned dipoles to reinject the recirculated beam 180° out of phase with the accelerated beam. This resulted in a transfer of beam power to the rf structure, with a consequential recovery of the beam power. The phase-energy correlation imposed by the

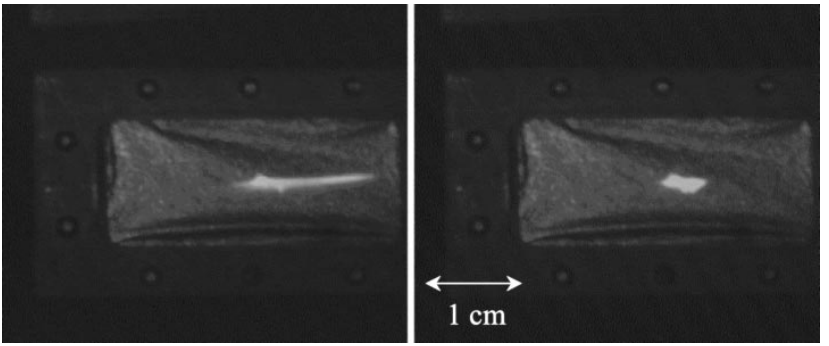

Figure 4 Beam viewer image in chicane downstream of FEL (dispersion of 0.4 m). *Left*: lasing; *right*: no lasing.

recirculator trim quadrupoles was selected to compensate the slope of the decelerating rf waveform.

The 2-MeV full energy spread of the recirculated beam, rather than adiabatically antidamping to a relative energy spread of order 20% during energy recovery to 10 MeV, was compressed to ∼100 keV at the dump, giving a final relative energy spread of 1%. This 20:1 energy compression requires not only the appropriate recirculator M_{56}, but also demands the proper higher-order terms (T_{566}) so as to correct both the lattice quadratic variation of path length with momentum and the curvature of the decelerating rf waveform (40).

Energy recovery proved quite efficient. This is illustrated by Figure 5, which presents the rf system generator power required in each cryomodule cavity with beam off, with 1 mA of beam without energy recovery, and at various currents with energy recovery. Essentially all of the beam power is recovered, inasmuch as no power beyond the zero current value is required.

3.3. System Operation and Performance

The JLab IR Demo project funding started in April 1996; construction and installation were completed in just over two years. Commissioning activities interleaved with construction began in the fall of 1997, with milestones met as indicated in Table 2.

Early in commissioning, the system was limited to ∼30% availability by the electron gun. Effort in this area led to a very reliable electron source with nearly 100% availability. The last installed GaAs wafer provided cathode lifetimes in excess of 600 C and delivered over 5 kC total charge (41).

The driver accelerator and FEL performed flexibly, robustly, and reproducibly. The system restored to full-power lasing in a shift after long shutdowns; during normal operations, lasing was recovered in minutes after a vault access. Operations

TABLE 2 Chronology of the JLab IR Demo FEL

Date	Milestone
October 1997	first electron beam in vault (injector)
December 1997	first electron beam to straight-ahead dump
March 1998	high-current single-pass linac operation (1.1 mA cw to straight-ahead dump)
June 1998	wiggler installed, first light (155 W cw at 5 μm /1.1 mA straight ahead)
July 1998	recirculator construction completed, first energy-recovered beam, first (low-power) lasing with energy recovery
December 1998	high-power lasing with energy recovery ($>$200 W cw at 5 μm/1.4 mA)
March 1999	kW-class 5 μm operation (710 W cw at 3.6 mA, mirror limited)
July 1999	1.72 kW cw at 3 μm/4.4 mA; kW-class tunable light at 3, 5 and 6 μm 5th harmonic (1 μm) lasing
September 1999	detection of Thomson scattered x-rays
August 2001	2 kW IR operation
November 2001	final beam operations, including production of nearly 20 W THz radiation; decommissioning and start of 10-kW upgrade installation

were simplified by a full suite of diagnostics (42), including beam position monitors, optical-transition-radiation-based beam viewers, beam-current monitoring cavities, and an interferometric coherent-transition-radiation-based bunch-length diagnostic built at the University of Georgia. The former pair of diagnostics allowed beam steering and transverse matching; the latter pair supported the longitudinal matching detailed above.

The FEL provided pulsed and cw lasing with variable timing (within limits dictated by the drive-laser fundamental of 75 MHz and the optical cavity length of 8 m) over continuously tunable ranges around 3, 5, and 6 μm (defined by mirror reflectivities). It was used as a source by a growing user community (JLab IR FEL user facility information is available at http://www.jlab.org/FEL/) and for machine studies. The latter included the experimental investigation of topics of rf control-system performance in an ERL configuration, BBU and FEL/RF stability (43), which is described in later sections, as well as investigations of tapered wiggler dynamics (44). The latter study demonstrated (pulsed) lasing with extraction efficiencies approaching 2%.

Also noteworthy were the production of 1-μm light through fifth harmonic lasing (45) and the generation of intense, short x-ray pulses through Thomson scattering (46). The latter holds promise of expanding the scope of the user facility to support pump-probe experimentation.

The phasing and compaction could be varied to perform a variety of gymnastics, such as the production of short bunches for THz radiation generation, with radiation powers of nearly 20 W (47). An MIT-Bates experiment (8), probing the full range of available phasing (from two-pass acceleration to double recirculation with three passes in the linac—including a coasting beam, which effectively doubles the recirculator current—to energy recovery) was duplicated (48, 49).

3.4. 10 kW IR/1 kW UV Upgrade

The U.S. Navy and Air Force have funded an upgrade of this system to 10 kW in the IR and 1 kW in the UV. The upgrade, in commissioning as of spring 2003, entails (*a*) doubling the injected current from 5 to 10 mA by increasing the bunch charge from 60 to 135 pC; (*b*) installing two additional cryomodules to raise the beam energy to ∼160 MeV; (*c*) upgrading the recirculator to accommodate higher beam energy and placing a pair of FEL insertions in the machine backleg; and (*d*) adding a pair of optical cavities to accommodate high-power operation in the IR and UV.

The machine (Figure 6), though almost a completely new installation, is in essence merely an enlarged clone of the IR Demo described above (50, 51). It thus retains the approach used in the earlier machine—that of a low-peak, high-average-power, optical resonator FEL with an energy-recovering srf linac driver operating at a high repetition rate. The 10-kW IR design goal will be achieved via an increase in both drive-beam power (doubled current and quadrupled energy) and FEL extraction efficiency (from 0.5% to 1%). The 1-kW UV FEL takes advantage of the high beam quality available from DC photocathode sources and increased acceleration available from the longer linac to drive an Advanced Photon Source (APS)-style undulator and 32-m optical cavity.

The system design (50, 51) is similar to that of the IR Demo, with design parameters given in Table 3. Following injection at 10 MeV, the beam is accelerated in a linac comprising three srf cryomodules: two exterior modules fabricated during the IR Demo construction from five-cell CEBAF cavities and in the center an entirely new design using the seven-cell, high-gradient srf cavities developed

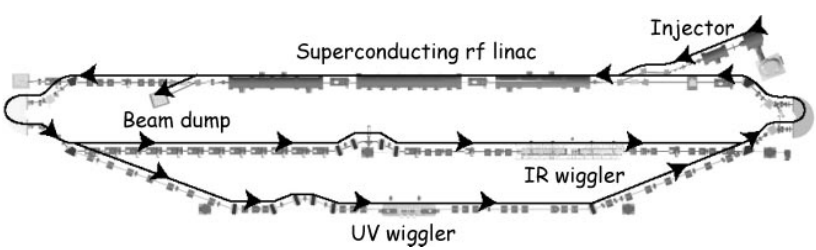

Figure 6 Schematic of JLab 10-kW IR/1-kW UV FEL upgrade configuration. (Courtesy of G.R. Neil.)

TABLE 3 System parameters of the JLab IR and UV FEL upgrade

Parameter	IR FEL Upgrade	UV FEL
Beam energy at wiggler	80–210 MeV	200 MeV
Average beam current	10 mA	5 mA
Bunch charge	135 pC	135 pC
Bunch repetition rate	74.85 MHz	74.85 MHz
Normalized emittance (rms)	13 mm-mrad	5–10 mm-mrad
Bunch length at wiggler (rms)	200 fs	200 fs
Peak current	270 A	270 A
FEL extraction efficiency	1%	0.25%
$\delta p/p$ before wiggler (rms)	0.5%	0.125%
$\delta p/p$ after wiggler (full)	10%	5%
CW FEL power	>10 kW	>1 kW

for use in a 12-GeV CEBAF upgrade. Following transport and matching to a Bates-type recirculation end loop, the beam is directed to either of two FEL insertions, one for IR and the other for UV. The IR, requiring the largest acceptance, lies along the more geometrically symmetric straight-ahead path; the UV is displaced to allow undulator placement in a pit specifically designed for this purpose. Chicanes are included in either insertion for use in bunch-length compression and for the production of THz radiation. After either laser, the recirculation is completed and the beam prepared for energy recovery using a second Bates end loop, whereafter it is matched, reinjected, energy recovered, and dumped.

Beam-dynamics issues are similar to those in the IR Demo, with rf stability and collective effects exacerbated by the higher single-bunch charge and average current. Of particular interest will be the effect of the higher-order mode (HOM) spectrum in the center cryomodule, which is based on a new cavity design.

At this writing, installation of the IR system, with only the two exterior cryomodules, is nearly complete. First beam operation and initial low-power lasing is expected in the spring of 2003, with high-power operation to follow by summer as the system is fully commissioned. Fabrication of the high-gradient center module is under way as of spring 2003, with installation expected in the summer of 2004. In parallel, construction of the UV bypass and optical systems will proceed to allow UV commissioning at that time.

3.5. 100 kW IR Concept

Planning is under way to extend the capability of the JLab 10-kW IR FEL to much higher output powers. Given the apparent scalability of the FEL performance, this can be accomplished by an increase in the microbunch repetition rate from

75 MHz to 750 MHz, with an associated increase in current from 10 mA to 100 mA. This follow-on upgrade path will likely be based on a completely new, 100-mA injector currently under development and an IR wiggler/optical-cavity system using cryogenic sapphire mirrors to accommodate the increased intracavity power. Because the single-bunch charge is fixed in this scenario, CSR and single-bunch wake effects are not enhanced. BBU and multibunch effects, however, are significant challenges and are the focus of ongoing investigation.

4. OVERVIEW OF ENERGY-RECOVERING LINAC PROJECTS AND PROPOSALS

Experience with the JLab IR Demo FEL's driver accelerator has motivated much of the recent interest in high-average-current ERLs. This section reviews many projects and proposals based on recirculating linacs and ERLs. Presently, the applications fall into four categories: high-average-power FELs, ERL-based light sources, high-energy electron-cooling devices, and electron-ion colliders. We emphasize the aspects of the devices that make application of energy recovery attractive.

As mentioned above, ERL-based FELs already exist, and upgrades to still higher average beam powers will continue. Within the next several years, serious efforts to build some form of ERL-based light source are likely to emerge. Brookhaven National Laboratory has been leading explorations of ERLs as electron-cooling devices, with the ultimate goal of increasing the luminosity of the Relativistic Heavy Ion Collider (RHIC). Brookhaven has also been a leader in considering electron-ion colliders beyond HERA, and scientists there have discussed a collider based on a high-average-current 10-GeV electron beam accelerated in an ERL and colliding with the RHIC beam. Similar ideas have been explored at JLab, where the existing CEBAF accelerator would accelerate a high-average-current, energy-recovered electron beam; the collider would be completed by adding a 50–100-GeV/nucleon storage ring designed especially to obtain and manipulate high beam polarization.

4.1. High-Average-Power FELs

Today a number of ERL-based FEL facilities worldwide are at various stages of construction and commissioning. Recently the Japan Atomic Energy Research Institute (JAERI) FEL successfully lased in energy-recovery mode, becoming the second ERL-based FEL (52). The room-temperature Accelerator-Recuperator FEL at the Budker Institute for Nuclear Physics in Novosibirsk (Figure 7) (53), the KAERI FEL in Korea (54), and the JLab 10-kW IR FEL Upgrade (50) are all under construction.

4.2. ERL-Based Light Sources

Present-day synchrotron X-ray sources are based on storage rings to produce the high-current beam required for synchrotron radiation. As discussed above, the

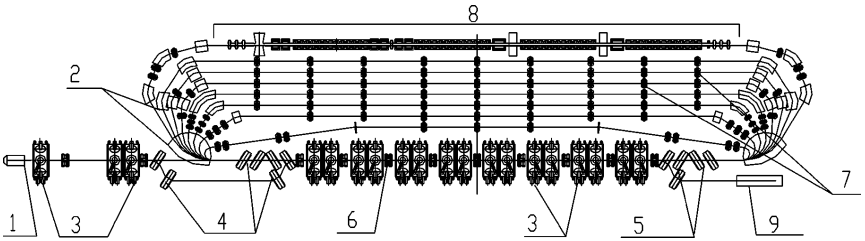

Figure 7 Accelerator-recuperator FEL in Novosibirsk. *1*: electron gun; *2*: bending magnets; *3*: RF resonators; *4,5*: injection and extraction magnets; *6*: focusing quadrupoles; *7*: straight sections with the quadrupole lenses; *8*: FEL magnetic system; *9*: beam dump. (Courtesy of N. Vinokurov.)

performance of these machines is constrained by fundamental processes, which ultimately limit the quality of the electron beam. Producing electron beams with superior characteristics for synchrotron radiation is possible via photoinjector electron sources and high-energy linacs; however, the energy consumption of such machines would be prohibitive. ERLs hold the promise of average current-carrying capability approaching that of present-day storage rings. Further, ERLs have the potential to produce smaller beam emittance and energy spread than is possible in storage rings, which leads to higher photon brilliance and coherence, round sources, and short-pulse-length radiation, while maintaining flexible machine operation. As a consequence, there is much activity in designing such ERL-based light sources. A recent International Committee on Future Accelerators (ICFA) report (55) summarizes this activity.

Recirculating linacs or ERLs are potentially exceptional sources of x-ray pulses in the 100-fsec pulse-width domain (56). The idea that superior beam emittance and energy spread from an ERL lead to higher ultimate photon brilliance has been advocated for several years by a group in the Budker Institute (57). Their project MARS (Multiturn Accelerator-Recuperator Scheme) is designed to reach the energy of 6 GeV and produce radiation both from undulators and from the bends in the recirculating arcs. The MARS scheme, together with the Accelerator-Recuperator FEL, are the only proposed ERLs that are based on normal-conducting rf cavities.

At present, the main srf-based light-source proposals are ERL at Cornell (58, 59), PERL at Brookhaven (60), 4GLS at Daresbury (61), ERLSYN at the University of Erlangen (62), and LUX at Lawrence Berkeley National Laboratory (63). There is also some exploratory effort at KEK in Japan. The Cornell/JLab proposal has 77 pC with a bunch repetition rate of 1.3 GHz yielding 100 mA in its high flux mode. PERL has twice the bunch charge and twice the average current. Both of these projects use an injector that produces very small normalized emittance of less than 2 mm-mrad. According to numerical modeling, this small emittance should be possible given the present understanding of emittance growth in photocathode sources. Daresbury's 4GLS incorporates a superconducting linac first as a FEL

source, then closes the recirculation loop to provide a high-brilliance cw source optimized for deep ultraviolet wavelengths. The ERL of Erlangen's ERLSYN has parameters very similar to those of Cornell, as the second phase of a project whose first phase is to build a third-generation storage-ring source. The Lawrence Berkeley Lab proposal LUX has a much larger charge per bunch, around 1 nC, and a much lower repetition rate of 10 kHz. It maintains the option of an ERL upgrade in case there is demand for increased beam power.

Cornell and JLab have proposed an ERL demonstration prototype as a first step toward a follow-on CEBAF-scale light source. This prototype would demonstrate full current injection with the required beam properties, acceleration to 100 MeV, and high-efficiency energy recovery (64).

4.3. Beam Electron Cooling

In electron cooling, a relatively low-energy electron beam is merged with a relatively high-energy ion beam, the electron-beam energy being chosen so that the average longitudinal velocity of the beams is the same. The electron beam acts as a heat sink, removing thermal energy from the ion beam and allowing collisions with cooled beams at higher luminosity than possible in the same collider without cooling. The cooling rate is proportional to beam average current.

High-energy electron cooling with high cooling rates is deemed possible now that ERLs have demonstrated technical feasibility. Brookhaven, in collaboration with the Budker Institute and JLab, is working on the technical design of an electron-cooling prototype for demonstrating electron cooling of the heavy ions at RHIC. The RHIC cooler will be driven by a 50-MeV, 100-mA ERL (65, 66). Such a device has two main extrapolations from present experience. The first, as discussed above, is the high-average-current source. Such a design may be even more difficult than for the light sources because the bunch repetition rate, to match the RHIC beams, must be reduced to 9 MHz, and to obtain the same average current, the charge per bunch must be increased to $\sim$10 nC. Second, in order to maximize the longitudinal overlap between the RHIC beam and the cooling electron beam and hence maximize the cooling rate, it is advantageous to actually debunch the beam before it enters the cooling channel and rebunch it prior to the second energy-recovery pass through the linac. Such phase-space manipulations have been demonstrated cleanly in ERLs only at much lower bunch charges.

4.4. Electron-Ion Colliders

Finally, ERLs have been suggested for electron-ion colliders for nuclear and/or particle physics research (67–69). An ERL would replace the electron storage ring in these applications. A simple *gedanken* experiment shows why such an arrangement might be advantageous and produce higher luminosity. The luminosity of a

collider is

$$L = \frac{f N_e N_i}{2\pi \left(\sigma_{ex}^2 + \sigma_{ix}^2\right)^{1/2} \left(\sigma_{ey}^2 + \sigma_{iy}^2\right)^{1/2}}, \qquad 2.$$

where f is the collision frequency, N_e is the number of electrons in the colliding bunch, N_i is the number of ions in the bunch, and the rms beam sizes of the two colliding beams (assumed unequal) are given by the σ's. Assume first a stable ring-ring collider design. For the electron ring to be stable, any current-limiting instability growth rate must be slower than one ring damping time. This means that the electrons must be confined up to one damping time, i.e., about 1000 revolutions.

Suppose now one has an ERL collider design with an identical ion-storage ring and an identical set of electron beam parameters to those of the ring-ring collider design above. Because the electron-beam parameters are the same, the ion beam will continue to circulate stably, even if one increases the ion bunch charge and luminosity considerably. Increasing the ion bunch charge will increase the disruption of the electron bunch by the beam-beam effect, but in the ERL it is no longer necessary to confine the electron beam for 1000 turns, only a few turns. Estimates of the emittance increase and the maximum deflection angle generated by a few beam-beam collisions show that there may be room to considerably increase N_i before energy recovery becomes difficult.

Detailed initial parameter lists, covering two collider schemes, have been worked out. The first is eRHIC, an electron-ion collider based at RHIC (67, 70), and the second is ELIC, an electron–light-ion collider based at CEBAF (71). In the eRHIC proposal, one of the RHIC rings is used to contain the ions and a new ERL is built. In ELIC, CEBAF is upgraded to a higher-energy ERL, and a new ion-storage ring is constructed. A recent ICFA report (72) summarizes activities on high-energy cooling and electron-ion colliders.

The parameters required by these ERL proposals are an extrapolation from today's demonstrated performance by one to two orders of magnitude both in beam energy and in average current. Figure 8 shows the location of the ERL applications discussed above in terms of beam energy and average current. ERLs for light sources are designed to operate in the energy range from hundreds of MeV up to a few GeV, with average current that can be as low as 10 mA (in the high-coherence mode) or as high as 100 mA (in the high-average-flux mode). ERLs for colliders are envisioned to operate in the 3–10-GeV energy range, and they require average currents of order 100–200 mA. A number of prototype facilities have been proposed to explore the technical feasibility of future ERLs. Among the proposed prototypes are (*a*) the Cornell/JLab ERL Phase I, which is a 100-MeV, 100-mA ERL designed to demonstrate the technical feasibility of the ERL Phase II machine, (*b*) the Brookhaven electron-cooling prototype designed to demonstrate the technical feasibility of an ERL-based electron-cooling device, and (*c*) JLab's 10-kW FEL Upgrade and its likely successor, the 100-kW IR FEL,

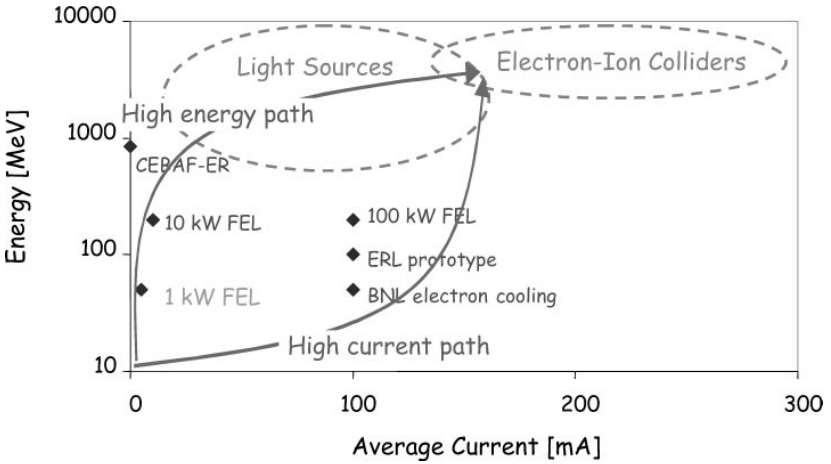

Figure 8 Energy-recovering linacs in terms of energy versus average current: existing, planned, and proposed ERL-based facilities.

which will also be used to ascertain the limits of ERLs in general and ERL-driven FELs in particular. Finally, the CEBAF-based experiment, CEBAF-ER, whose first phase was completed in April 2003, was designed to explore the feasibility of GeV-scale ERLs. The two parallel development paths leading to the envisioned ERL applications, a high-energy path, and a high-current path, are also shown in Figure 8.

We now discuss the technical challenges of the next generation of ERLs, focusing on srf linac-based schemes.

5. SCALING OF ENERGY-RECOVERING LINACS TO HIGHER ENERGIES

Energy recovery allows the efficient generation of high-power, high-quality electron beams. It is consequently desirable to understand any limitations on both the energy and the power that can be achieved in an energy-recovering system. Such machines will, in general, comprise an injector, multiply recirculated superconducting linac(s), and a recirculator. The recirculator will either transport beams at multiple energies in one beam pipe or have individual beam lines transporting monoenergetic beams, using a "spreader" to separate the beams following the linac and a "recombiner" to combine them for reinjection.

The application of energy recovery to high-energy systems must address numerous issues of relevance to large-scale, recirculating, superconducting rf-based accelerators. These include the following:

- the choice of injection energy

- the number of passes through the linac
- general features of the linac topology, such as the use of single or multiple linacs, the use of asymmetric gains in multiple linacs, and the connectivity of the recirculation path
- details of phase-space management, such as the degree of functional modularity and specific schemes for longitudinal and transverse matching
- phase-space preservation throughout the acceleration and energy-recovery cycle
- control of beam halo

Each of these issues is addressed below.

5.1. Injection Energy

The choice of injection energy is a cost/performance issue with significant implications. Given that the injector is nominally not energy recovered, it represents both a primary limitation to the efficiency of machine operation and a significant cost, owing to the high rf power required for the acceleration of large beam currents. Costs and limits imposed by rf windows and couplers are also alleviated by lower injection energy. Low injection energy is therefore economically desirable. However, given that the acceleration and transport systems of the ERL will be common to both low- and high-energy beams, injection energy limits machine performance by introducing regions of low beam rigidity (with commensurate sensitivity to errors and instability), in which the allowable focusing strength is limited.

Higher injection energies in general allow the use of strong focusing at the front end of the ERL. This, in turn, reduces the peak beam-envelope values, with an associated improvement in beam stability and system operability and reduction both in sensitivity to errors and beam loss. This injection-energy issue is coupled to the choice of linac focusing scheme and the available srf gradients. A conventionally focused low-gradient srf linac (such as CEBAF) with a large ratio of full to injected energy typically operates with maximum beam envelopes at or in excess of the linac length—a circumstance for which error sensitivity and halo may prove problematic. Reduction of the energy ratio (for example, through the use of a pre-accelerator, to $\sim$10:1) can provide improvement to about half the linac length (73). However, use of a short linac with very high-gradient rf moves the machine back to a regime in which the peak envelopes are of order the linac length, but in which performance remains robust because of the short linac. The machine concept must balance cost and performance using the best available acceleration technology. Present designs tend to concentrate on injection energies of 5–10 MeV, although some proposals invoke injection energies as low as 1–2 MeV (74). To date, the lowest acceptable limit of injection energy in an ERL has not been unequivocally determined and must be investigated.

5.2. Number of Passes

The compromise between cost and performance is perhaps most clearly illustrated by the selection of the number of passes in a specific machine design (75). The simplest solution for an accelerator is, of course, a linac—a single pass accelerator. However, if multiple beams are present (as is required for energy recovery), the focusing structure required to manage simultaneously beams of different energies will be either very complex—if strong focusing is desired—or essentially ineffective. In contrast, if the beams are in antiparallel motion [as in Tigner's concept (10)], external focusing can be easily provided, but beam-beam collisions can adversely affect performance. Further, the cost of a simple srf-based linac is prohibitive. Recirculation is thus adopted as both a cost-control measure and as a means of optimizing performance. In this approach, a single pass through a large, expensive rf system is traded for multiple passes through a smaller, less expensive system, plus a relatively inexpensive beam-transport system. The optimum cost is driven by the sum of the linac cost (which falls as the length decreases and the number of passes increases) and the recirculator cost (which climbs, initially linearly and then more steeply as the system grows more complex at higher numbers of passes).

Performance is optimized primarily by limiting (through recirculation) the distance over which beams of significantly different energy must be managed in a common structure. In this case, constraints on the allowable focusing, acceleration, and error tolerances—imposed by the lowest energy present—become less profound, and the response of the highest-energy beam to these constraints becomes less severe. Appropriate choice of machine topology—the path of the beam as it is accelerated and energy recovered through a sequence of linac passes—can provide further improvements. The majority of currently proposed ERL designs assume one accelerating pass and one decelerating pass through the linac structure. A thorough discussion of a system with two accelerating and two decelerating passes can be found in Reference 76.

5.3. General Features of Machine Topology

Both the cost and the performance of a recirculating linac can benefit from an appropriate choice of machine topology. As discussed, recirculation reduces srf linac cost; further benefit can be obtained by subdividing the recirculating linac into two (or more) shorter linacs. Both the length of the individual linacs and the return path length can be minimized. The linac length reduction helps control beam envelopes (with performance improvements), and a relatively short return path reduces the cost of the beam-transport system and tunnel. This improvement is achieved at the cost of increased operational complexity: Should individual transport of each energy be chosen, additional spreaders and recombiners will be required to separate and merge the beam at each of the multiple linacs.

As noted above, the selection of injection energy is important because it dictates both the rigidity of the most sensitive beam and the front-end focusing available in the linac. As mentioned, the use of a high-energy injector or a pre-accelerator

can help reduce the impact of injection energy on performance. This is, however, costly and largely unnecessary, since similar functionality and performance can be achieved through the use of split linacs with asymmetric energy gains. Low injection energy places nonrigid beam in the linac structure and limits the focusing strength that can be applied to both the low-energy and higher-energy beams. The use of a short first linac and a partial recirculation, followed by a longer second linac, limits the region over which the focusing is weak, increases the available focusing over the longer accelerating section, and thereby reduces peak beam-envelope functions.

One can obtain further improvements by modifying the machine topology and bisecting the transport system. A conventional split linac re-injects higher-energy beams for further acceleration or energy recovery at the initial (low-energy) injection point. A bisected linac topology recirculates a high-energy beam, re-injecting it into the higher-energy split linac. Although this may increase the complexity and cost of the transport system, it can significantly improve the machine performance by improving the match of the external focusing to the beam energy of the various passes in the common accelerating and extraction structure. A design example of the bisected linac topology has been explored for a JLab light source (77) (Figure 9). This machine design accelerates an injected 10-MeV beam to 10 GeV in two passes, and energy-recovers it to 10 MeV in two subsequent passes, using a bisected linac geometry. A "photon farm" allows implementation of various user-defined synchrotron radiation sources.

5.4. Phase-Space Matching

Acceleration, delivery, utilization, and energy recovery of a high-energy and high-power beam requires careful management of the full six-dimensional beam phase space. The full beam-handling cycle must include the following: transverse and longitudinal matching of injected beam to linac, transport of multiple beams at different energies through a common linac accelerating and focusing structure, phase-space management during recirculation and/or during delivery of beam for

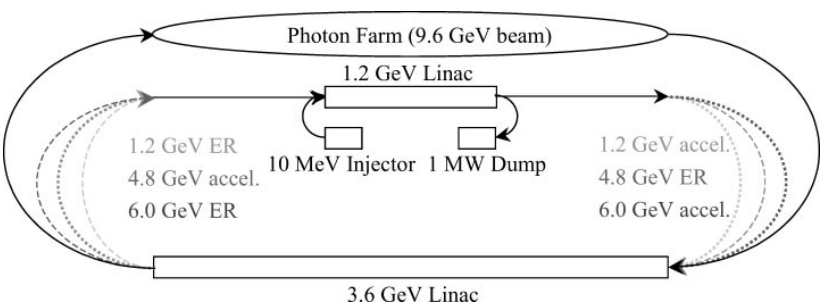

Figure 9 A split-linac topology for ERL-based light source.

users, configuration of phase space during energy recovery, and phase-space management during the energy-recovery cycle.

High-energy ERLs must provide a longitudinal phase-space management scheme that will allow stable transport of high currents, possibly in high-charge-state bunches, through an srf acceleration system. This may be most successful, for example, if the bunch length is kept long during acceleration (to alleviate excitation of higher-order modes in the srf structures) and compressed only when short pulses are required for users. This is essentially the scenario used in the JLab IR Demo FEL driver described above. Such longitudinal matching may utilize microtron-like phase-space management, in which acceleration is off-crest and the transport has nonzero momentum compaction, yielding synchrotron-phase stability. As an alternative (particularly if short bunches are to be accelerated), acceleration can occur on-crest and the recirculation transport can be isochronous (as in CEBAF); this is not phase-stable, but with appropriate feedback, such systems can provide extremely small longitudinal phase space.

Transverse phase-space management schemes must address the issue of focusing multiple beams at multiple energies in the common accelerating structure, as well as addressing the problem of beam recirculation. A variety of solutions (solenoid, quadrupole alternating-gradient singlet, doublet, or triplet focusing) are available for linac optics. The presence of multiple energies provides the primary constraint because the focusing that can be tolerated by the lowest-energy beam without betatron instability limits the stability of higher-energy passes. By matching the system to provide strong focusing to the lowest energy beam—known as graded-gradient focusing—adequate stability can often be provided for all passes. This is particularly true if high accelerating gradients are available; in this case, the pass-to-pass relative focusing strength improves rapidly, as the rigidity of the lowest-energy beam in the structure rapidly increases. Figure 10 shows an example of graded-gradient focusing in a ∼0.5-km linac where betatron envelopes do not exceed 70 m. The accelerating gradient assumed is ∼20 MV/m.

Recirculation transport optics must similarly choose among numerous available options, each of which may be the most appropriate for a specific application. Individual beams can be split off from all others and transported (as in CEBAF), or beams of all energies can be transported through a common beamline structure (as in a microtron or polytron). The former solution provides greater operation

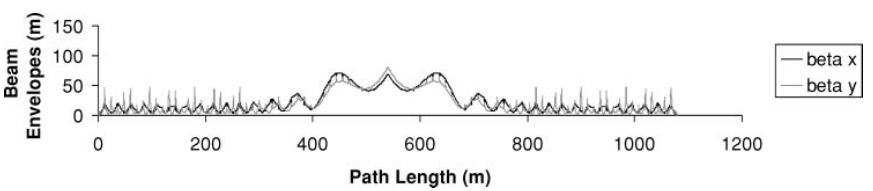

Figure 10 Beam envelopes (m) in a 10-MeV to 10-GeV recirculating, energy-recovering accelerator using graded-gradient focusing.

flexibility, but at higher cost and increased complexity. The choice of method of beam separation (horizontal, vertical, dispersion-suppressed, or dispersive) has further performance and operability implications. Large recirculators will typically bend (primarily) in the horizontal plane; beam separation may therefore be more naturally accommodated in the recirculation bending plane. This, however, may couple the beam splitting and recombination process (and dispersion management therein) to that of matching the transverse phase space from linac to recirculator—an additional operational complexity. Further, the additional bending in the recirculation plane may contribute to quantum-excitation-driven degradation of the phase space, will couple to compaction management (through dispersion handling), and may not be simple to implement mechanically. Vertical beam separation (as used in CEBAF) may therefore provide better performance and simpler implementation.

All such concerns relate to the issue of functional modularity—the extent to which a single beam-optics module is intended to locally accomplish a particular task (such as matching of a particular beam parameter) and/or provision of a single operational "knob" (control parameter) that is associated with a particular beam property. More modular systems are locally simpler to operate but may become so globally complex that they are unreliable and/or prohibitively expensive. Design optimization must address the specific end-user requirements and provide adequate but cost-effective means of meeting those needs.

5.5. Phase-Space Preservation

User requirements for high brightness and the necessity of energy recovery without intolerable beam loss demand that beam phase space be generated, accelerated, delivered, and energy-recovered without undue degradation. Beam quality can be adversely affected by numerous phenomena. Many such effects, such as space charge, BBU, other wakefield-driven degradation or instability, and coherent synchrotron radiation are properly considered collective effects and are current- or charge-dependent; these are discussed in the next section. Others—such as incoherent synchrotron radiation effects (78) and sensitivity to errors in alignment, powering, and electromagnetic component field quality— can be addressed in ERL design and operation as lattice-design issues. Key to successful implementation of energy recovery is the application of standard transport-system design practices such as the following:

- appropriate control of beam envelopes and beam response to perturbative steering (transfer matrix elements)
- provision for operational control of beam parameters ("one parameter, one knob," or "functional modularity" to allow, for example, envelope matching and/or phase-advance control to optimize instability thresholds or beam properties delivered to users)
- avoidance of unduly large beam–optical-aberrations

- implementation of standard methods for suppression of quantum-excitation-driven degradation of the beam phase space
- use of the highest possible accelerating gradient to increase beam energy and reduce the mismatch among the energies of the various beams present and the excitation of the available external focusing.

These principles reduce error sensitivity, improve operability, and lead to more robust, recoverable machines with better beam quality.

ERLs can be performance-limited by effects atypical in more conventional systems. For example, small perturbations imposed by modest errors at high energy can become important when their accumulated response is adiabatically anti-damped during energy recovery. This is illustrated by the fact that magnet field inhomogeneities in transport systems manipulating momentum compaction can lead to large energy spreads after energy recovery (79). A differential (across the beam) magnetic field error at a dispersed lattice point will, through propagation of the resulting differential kick $\delta x'$ by the matrix element M_{52}, produce an rf phase spread in the beam during energy recovery. This in turn leads to unexpected remnant energy spread after energy recovery.

The existence of such phenomena emphasizes the inherently six-dimensional phase-space dynamics of energy-recovered systems and advocates consideration of the influence of novel conditions (such as transverse/longitudinal coupling) when designing and operating ERLs.

5.6. Beam Halo

The control of beam halo is both a phase-space preservation issue and a concern when evolving to high current/high power because it becomes a more severe problem as charge per bunch and current increase. In the JLab IR Demo FEL driver accelerator, several indicators place an upper limit on the amount of beam loss in the recirculator (energy above 10 MeV) of 0.1 μA out of 5 mA (67 pC per bunch at 75-MHz repetition frequency). This amount of beam loss, although extremely small, may be marginally acceptable for some of the proposed ERL designs, since it can potentially give rise to kilowatts of lost beam power. A detailed understanding of the origin of halo, and efforts to control beam loss via careful control of beam envelopes and adequate aperture, will be required for successful high-power ERL operation.

5.7. CEBAF-ER Experiment

CEBAF with Energy Recovery (CEBAF-ER) is an experimental implementation of energy recovery in the GeV-scale CEBAF accelerator (80, 81). CEBAF-ER accelerates a beam through the first pass of the CEBAF linac to $\sim$1 GeV and transports it through a magnetic chicane—thereby introducing a half-rf-wavelength phase delay—prior to recirculation and reinjection for a second pass. With the phase delay provided by the chicane, the second pass is decelerated to the injection energy

and extracted from the machine. This experiment will (*a*) allow investigation of beam-quality preservation throughout the acceleration and energy-recovery cycle, (*b*) allow investigation of the effect of dynamic range (injected to full energy ratio) on system performance, and (*c*) serve as a large-scale technology demonstration.

A future modification to the experiment will reduce the path-length differential of the phase-delay chicane to one-quarter rf wavelength. This will allow machine operation with an acceleration pass, a coasting beam pass, and an energy-recovery pass, effectively doubling the current in the recirculator. Such current doubling has previously been demonstrated on a smaller scale at the MIT-Bates recirculator (8, 38) and in the JLab IR Demo FEL driver (48, 49). This modification will test the viability of the technique as a method for increasing source brightness in high-energy, high-power systems.

When this review was submitted, CEBAF-ER had just engaged in initial operations. During a running period of March 25, 2003–April 2, 2003, this modified CEBAF configuration successfully accelerated and recirculated an injected beam to 1 GeV full energy and recovered it by recirculation and deceleration to the injection energy. Initial energy-recovery operation, and the preponderance of the investigation period, employed a 56-MeV injection energy. With this dynamic range, the system was stable and cw operation at 80 μA was achieved. Beam and rf system properties were characterized; the data are under evaluation. Figure 11 illustrates both passes on a beam viewer at the midpoint of the second CEBAF linac. At this location, the accelerated beam (clipping the hole in the

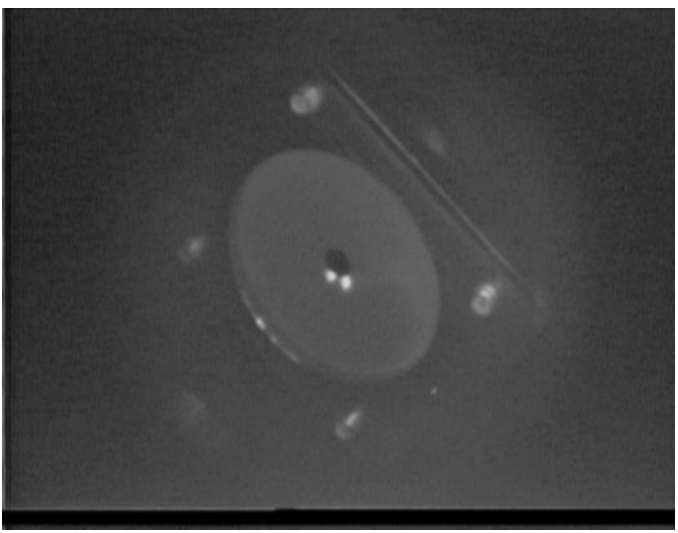

Figure 11 CEBAF-ER experiment. Accelerated (*left*) and recovered (*right*) beams at midpoint of the south linac. This viewer image demonstrates that the decelerating beam remained well-defined and of similar quality to the accelerating beam.

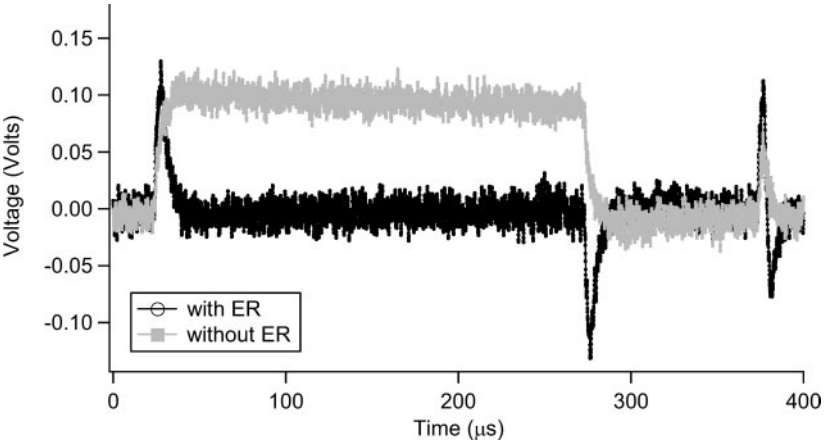

Figure 12 CEBAF-ER experiment. RF system gradient modulator drive signals during pulsed beam operation, with and without energy recovery.

viewer, at the lower left edge) is at ∼750 MeV, whereas the decelerated beam (below and to the right of the hole) is at ∼300 MeV. The scale of beam size and beam quality is set by the view-screen diameter of ∼2.5 cm. Figure 12 illustrates the rf system gradient drive signal during pulsed-beam operation. Without energy recovery, this signal is nonzero when a 250-μsec beam pulse enters the rf cavity, indicating power is drawn from the cavity. This occurs either when the recirculation of beam is completely impeded (as in the long pulse train) or in the period during which the head of the pulse train does not close on the machine circumference (at the leading edge of the long pulse, or during the short pulse, which is a diagnostic pulse of duration shorter than the beam-circulation time). With energy recovery, the signal is zero once the initial transient passage of the leading edge of the pulse is over, indicating no additional power draw is required by the cavity.

The latter portion of the run was devoted to operation with an injection energy of 20 MeV. This configuration tested machine and recovery dynamic range. Although the beam was readily accelerated from 20 MeV to 1 GeV and recovered to 20 MeV, beam stability was not as good as with the higher injection energy, and cw operation was limited by beam loss to ∼10 μA. The beam was, however, well defined, and beam property measurements were completed. Careful beam-halo measurements for both injection energies were recorded.

6. SCALING OF ENERGY-RECOVERING LINACS TO HIGHER CURRENTS

The potential of ERLs is best realized at the highest average beam current that can be transported with acceptable beam-quality degradation. The majority of the proposed ERL-based projects require average currents of order 100 mA, an

extrapolation from today's demonstrated performance by more than one order of magnitude. Below, we discuss the technical issues that must be resolved so that feasibility of these designs can be demonstrated and the ultimate limits of ERLs can be understood.

6.1. Generation and Preservation of Low-Emittance, High-Current Beams

The majority of ERL applications require low-emittance (normalized rms emittance $\sim$1 mm-mrad) and short bunch-length (rms bunch length from $\sim$100 fsec to $\sim$1 psec) beams. In order to take full advantage of the ERL technology, one should both generate and preserve a low-emittance, high-average-current beam. Laser-driven, photoemission guns are considered likely source candidates (82), but technology development is required to demonstrate operation under high-current conditions with adequate lifetime. Once the low-emittance beam is generated, one needs to ensure its preservation first in the low-energy regime, where careful emittance compensation must take place against space-charge effects, and then in the linac and beam lines in the presence of wakefield effects and in the recirculator against coherent synchrotron radiation-induced emittance degradation (30). These single-bunch collective effects are being studied extensively by the designers of present-day high-charge-per-bunch sources, such as Self Amplified Spontaneous Emission (SASE) radiation sources. Here we focus on average-current effects, which can potentially limit the reach of ERLs.

6.2. Multibunch Instabilities

In recirculating linacs, in general, the beam and the rf cavities form a feedback loop, which closes when the beam returns to the same cavity on a subsequent pass. The closure of the feedback loop between beam and cavity can give rise to instabilities at sufficiently high currents, driven predominantly by the high quality factor (Q) superconducting cavities. ERLs, in particular, are more susceptible to these instabilities because they can support currents approaching or exceeding the threshold of the instabilities. Instabilities can result from the interaction of the beam with the fundamental accelerating mode (beam-loading instabilities), from the interaction of the beam with transverse higher-order modes (HOMs) (transverse BBU), and from the interaction of the beam with longitudinal HOMs (longitudinal BBU). The basic mechanism of all three types of multibunch instabilities is fundamentally the same. Next we describe the instability mechanism for each of the three instabilities in physical terms. For simplicity and clarity, we assume a single-cavity, single-HOM, two-pass configuration.

Multibunch, multipass transverse BBU has been observed and understood for a long time. Suppose a beam enters the rf cavity on axis and a previously excited HOM deflects the beam horizontally or vertically. When the beam returns to the same cavity displaced because of the optics of the recirculator, it can exchange

energy with the HOM in a way that excites the HOM and can now further deflect subsequent bunches until they hit the beam pipe.

The mechanism of the longitudinal BBU is as follows (83): If there is an initial excitation of some longitudinal HOM in the linac rf cavity and the bunches enter the cavity on the first pass perfectly spaced, then on exiting the cavity, the energy of the bunches will be modulated by the mode. If the isochronicity of the recirculation optics on the second pass through the cavity is not perfect, the energy modulation will be translated into a spacing modulation. This modulation will generate a side-band current whose magnitude is scaled by the magnitude of the perturbation and whose frequency matches that of the exciting HOM. Thus, on the second pass, the resulting current can enhance the excitation of the HOM that created it. A feedback loop is formed, analogous to that which generates multipass transverse BBU. An important difference, however, is that the induced current can only achieve a value equal to the average beam current, and saturation will occur (83).

The mechanism for the beam-loading instabilities is as follows (84): Suppose ΔE is the electron-beam energy error. This error can shift the beam centroid off its central trajectory and lead to beam scraping on apertures. In addition, an energy error ΔE can couple to the M_{56} of the recirculator and cause phase shifts of the decelerating beam. Furthermore, the time derivative of an energy error coupled to the M_{56} will result in a shift of the bunch arrival frequency at the wiggler, which is equivalent to optical-cavity detuning. This shift changes the FEL gain function, which also changes the laser output power. Changes in the laser power will change the energy of the recirculating beam, potentially leading to additional beam loss on apertures and phase shift of the decelerated beam. All three effects—beam loss, phase shifts, and laser-power variations—change the beam-induced voltage in the cavities through the recirculating beam, hence the term beam-loading instabilities. If the rf feedback lacks sufficient gain and bandwidth, the change in the beam-induced voltage will further change the cavity voltage in a way that amplifies the energy error of the electron beam and drives the loop unstable. For cw accelerators, the beam-loss instability is of no practical concern because losses can never be high enough to induce instability before beam loss itself interrupts operation.

6.2.1. SINGLE-CAVITY MODEL For all three instabilities, there is a well-defined threshold current that occurs when the power fed into the mode equals the mode power dissipation. In the simple case of a single cavity, single mode and single recirculation, an analytical expression for the threshold current applicable to all three instabilities can be derived and is given by

$$I_{th} = \frac{-2p_r c}{e(R/Q)_m Q_m k_m M_{ij} \sin(\omega_m t_r + l\pi/2)e}, \qquad 3.$$

where (r/Q) and Q are the shunt impedance and quality factor of the mode m with frequency ω_m, M_{ij} is the (i, j) transfer-matrix element of the recirculator, $k = \omega/c$ is the wave number of the mode, t_r is the recirculation time, and p_r is the

momentum of the recirculating beam. The integer l is equal to 1 when m denotes a longitudinal HOM, and it is equal to 0 otherwise. The above equation is valid only for $M_{ij}\sin(\omega t) < 0$. Further discussion of the sign of the equation can be found elsewhere (85, 86). When $i, j = 1$, 2 or 3, 4 and m denotes a transverse HOM, this expression gives the threshold current of the transverse BBU. When $i, j = 5$, 6 and m denotes a longitudinal HOM, this expression gives the threshold of the longitudinal BBU, and when $i, j = 5, 6$ and m denotes the fundamental accelerating mode, this expression gives the threshold of the beam-loading instabilities. This approximate expression is useful for understanding the parametric dependence of the threshold current on accelerator and beam parameters, and, under certain conditions, it may also be useful for obtaining estimates of the threshold of these instabilities. In general, however, numerical codes that take into account the details of a given configuration and the possible interaction among several modes should be used to calculate the threshold current.

In the following, we discuss recent analytical, numerical, and experimental results for the beam-loading instabilities, which can potentially limit high-average-power FELs, and the transverse BBU, which appears to be the limiting stability mechanism in ERLs in general (43), particularly when operating at higher rf frequencies, 1.3–1.5 GHz. Longitudinal BBU appears to have the highest threshold, because typical values of M_{56} are an order of magnitude smaller than M_{12} or M_{34}—which drive the transverse BBU—whereas typical damping of the strongest longitudinal HOMs is at the 10^4–10^5 level, similar to the transverse HOMs.

6.2.2. BEAM-LOADING INSTABILITIES A theoretical model of the beam-loading instability that describes the coupled system has been developed. The model includes the beam–rf interaction, a precise representation of the rf control system and the electron–photon interaction in the optical cavity (84), and it has been solved analytically and numerically. Further, an experiment has been carried out in the JLab IR Demo FEL to verify the model. The experimental data agree with the theoretical predictions quantitatively when the FEL is off and qualitatively when the FEL is turned on (34). Further analytical and experimental work is planned in the IR FEL Upgrade to reconcile remaining uncertainties.

6.2.3. TRANSVERSE BEAM BREAKUP The theory of transverse BBU is quite mature (87). The most recent highlights include an analysis of the effect for an arbitrary number of cavities and recirculations based on the impulse approximation (88). A generalization of the theory to include operation at a subharmonic of the accelerating frequency was obtained in 1991 (89). For M recirculations and N cavities, the final solution is obtained by solving for the eigenvalues of an M-dimensional matrix, which describes the transverse phase space evolution of the entire system. In the case of subharmonic bunching, the dimensionality increases to $N \times M - 1$.

In 1987, a two-dimensional simulation code, called TDBBU, was written to predict the threshold of the transverse BBU instability for arbitrary recirculating linac configurations (90). Plots of the bunch transverse location as a function of bunch

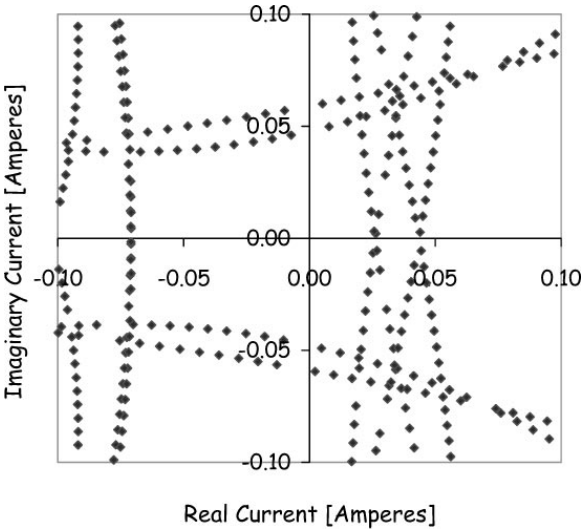

Figure 13 Beam breakup stability plot for the JLab IR Demo FEL.

number are used to determine the instability threshold. A new code, MATBBU, has recently been developed as a complementary numerical tool (91). Plots of complex current eigenvalues are used to determine the instability thresholds in MATBBU. As the coherent frequency is swept in real frequency with an arbitrarily small imaginary part corresponding to growth, families of complex current eigenvalues are determined. The actual threshold current corresponds to the smallest positive real value obtained. Figure 13 shows the stability plot for the JLab IR Demo FEL calculated with MATBBU. All the rf input parameters, including Q values of HOMs, shunt impedances, and frequencies, as well as the recirculation time, have measured values. The optics input parameters are calculated. The calculated threshold current is $\sim$26 mA, in excellent agreement with TDBBU's prediction of 27 mA (92).

Neither code had been benchmarked against experimental data, despite previous attempts (93) in the injector of the CEBAF accelerator. A series of experiments was carried out at the JLab IR Demo FEL in order to (*a*) attempt to induce the BBU instability and (*b*) measure beam-transfer functions in the recirculation mode. The experiment aimed toward inducing the BBU instability consisted of both changing the optics of the recirculator, so that larger beta functions in the cavity locations were obtained, and lowering the injection energy into the linac to 5 MeV and the final energy to 20 MeV. Under these conditions, the predicted threshold was just under 5 mA. However, during the execution of the experiment, the beam quality was sufficiently poor that the beam tripped at 3.5 mA, and the instability was not observed.

The second experiment consisted of beam-transfer function measurements in the recirculating mode. Although these measurements were performed at beam currents below the threshold current, they led to clear estimates of the instability threshold. A broadband beam-position monitor, rewired to serve as an rf kicker, was used to impart transverse momentum to the beam with the modulating frequency of the HOM under study. A network analyzer was driving a broadband amplifier at the proper frequency, sweeping the frequency across the HOM frequency. The signal from the cavities was fed back to the network analyzer's input port to complete the S21 measurement. Figure 14 displays the cavity's response to a frequency scan of the HOM frequency at 1887 MHz at various beam currents from 0 to 4 mA. Data were recorded by exciting different HOMs at several different cavities, with different r/Q and Q values, at two different beam energies, and for several transport optics settings. The threshold current was derived from nonlinear least-square fits to the data (86). Under the various accelerator configurations, the threshold current was determined to vary between 7 and 32 mA. For the nominal IR Demo FEL configuration, the threshold was between 16 and 21 mA. Compared with the theoretical prediction of 27 mA, the resulting agreement is at the 40% level. The observed dependence of the threshold current on the recirculator optics has not been quantified yet. Further experiments in the JLab IR FEL Upgrade and extension of the analysis tools are planned.

Transverse BBU can be controlled by damping the HOMs to lower Q values and/or by active feedback systems. For HOM damping to be effective, it is necessary to extract the HOM power from suitable ports near the cavity cells without removing power at the fundamental accelerating mode. This is done with "notch" filters that reject the fundamental mode. Several mitigating measures can be invoked to raise the instability threshold:

- cavities with fewer cells have fewer modes to be damped and generally are less susceptible to "mode trapping" (modes can be extracted and damped more easily)
- cavities with special shapes, which favor either the extraction or the propagation of modes along the beam pipe, can be utilized
- operation at lower rf frequencies, 350 MHz to 805 MHz, with typically fewer cells per cavity that allow easier access and extraction of the HOMs, can be considered
- Bunch-by-bunch transverse feedback, similar to that used in B factories, is expected to raise the instability threshold to $\sim$1 A

6.3. Superconducting RF Issues and HOM Power Dissipation

Although energy recovery works well with pulsed beam, its potential is truly realized with cw beam (high average current). As a consequence, all the ERL applications proposed to date require cw rf fields. Superconducting rf parameter optimization for ERLs in the multi-GeV energy range, which minimizes linac

length and cryogenic power consumption, points toward gradients of ∼20 MV/m and intrinsic quality factors of the rf cavities, Q_0, of order 1×10^{10}. This level of srf performance has not been demonstrated in cw, high-average-current operating conditions. In addition to the stronger damping of HOMs (both longitudinal and transverse) required, research and development toward increasing the quality factor Q_0 of the cavities would directly reduce the ERL operating costs and increase the overall ERL efficiency.

Finally, efficient extraction of HOMs generated by subpicosecond short bunches must be ensured. High-average-current and short-bunch-length beams in superconducting cavities can excite HOMs, which, in addition to beam stability consequences, could result in increased cryogenic load due to power dissipation in the cavity walls. The power in HOMs, primarily longitudinal, depends on the product of bunch charge, q, and average current, I_{ave}, and it is equal to $2qk_{\parallel}I_{\text{ave}}$, where $k_{\parallel}$ is the loss factor of the superconducting cavity and the factor of 2 accounts for the two beams in the cavity (accelerating and decelerating). The total power depends on the bunch length through the loss factor. At high currents and short bunches, the amount of dissipated power can be quite high. For example, for average current of 100 mA, bunch charge equal to 0.5 nC, and $k_{\parallel} = 10$ V/pC, the HOM power is approximately equal to 1 kW per cavity. Part of this power is expected to be extracted by HOM couplers and be absorbed in room-temperature loads, and part of it is expected to be absorbed by cooled photon absorbers placed between cavities or cryomodules. The excitation of high-frequency HOMs by the short bunches can, in principle, degrade the cavity's quality factor, according to Bardeen, Cooper & Schrieffer (BCS) theory, and increase power dissipation in the cryogenic environment (94).

Experimental measurements of the power dissipation under varying beam parameters were pursued at the JLab IR Demo FEL. The amount of HOM power transferred to the loads was measured and compared with calculations. Temperature diodes were placed on the two HOM loads of one of the linac cavities and temperature data were recorded for values of the charge per bunch ranging from 0 to 80 pC, in steps of 20 pC and three values of the bunch-repetition frequency, 18.7, 37.5, and 75 MHz. Figure 15 displays the measured HOM power versus bunch charge in one of the two HOM loads per cavity, as well as least-square fits to the data constrained to a single value of the loss factor. The data are consistent with the calculated fraction of the HOM power absorbed by the loads, approximately 30% of the total power. At the present time, no statement can be made about the amount of power dissipated in the cryogenic environment because no instrumentation was in place to measure it. Detailed measurements will be needed in the proposed ERL prototypes to demonstrate adequate efficiency of the power extraction schemes.

6.4. RF Coupling Optimization and RF Control

In ERLs, the multiplication factor κ increases as a function of the loaded quality factor Q_L of the superconducting cavities, resulting in higher overall ERL efficiency at higher Q_L. An important question, therefore, is how high Q_L can be.

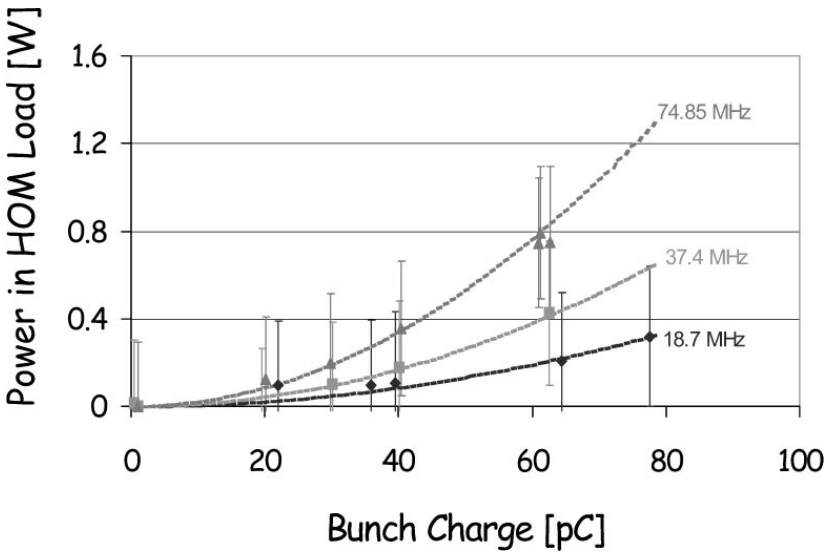

Figure 15 Measured higher-order mode (HOM) power dissipated in one of two HOM loads per linac cavity versus bunch charge at three bunch repetition rates.

A high Q_L implies a narrow resonance of the superconducting cavity; therefore, microphonic vibrations can cause large phase and amplitude fluctuations that need to be corrected if a certain value of the energy spread is to be maintained at the exit of the linac. Furthermore, for high-gradient and high-Q_L cavities, the radiation pressure during gradient turn-on can shift the resonant frequency of the cavity by several bandwidths of the cavity resonance, resulting in operational difficulty and, under certain conditions, unstable behavior (95).

In superconducting cavities, in the absence of beam loading, the coupling optimization is dominated by the amplitude of microphonic noise (96). For example, in the Cornell/JLab ERL, the optimum Q_{ext} is 2.6×10^7 assuming 25 Hz of microphonic noise. With this coupling, the required rf power is 8 kW per cavity. Several rf control-system concepts have been proposed, including the self-excited loop, the generator-driven system, and a hybrid of the two (97). Ideas for active suppression of microphonic noise and Lorentz-force–detuning using piezo elements are also being explored (98).

7. RESEARCH AND DEVELOPMENT PATH

ERLs are emerging as a powerful application of rf superconductivity and have led to novel accelerator designs worldwide for a variety of applications. The physics of these accelerators is theoretically well understood, and experimental verification of

simulation codes and models has started. Proposed ERL prototypes are expected to elucidate the ultimate limitations of ERLs in the multidimensional space of average current, energy, bunch charge, bunch length, and other fundamental accelerator and beam parameters. The following research and development topics will need to be addressed to ensure technical feasibility of future high-current, high-power ERLs:

- development of high-average-current, low-emittance guns and injectors
- effects of coherent synchrotron radiation on beam quality
- beam-halo formation and control of beam loss
- demonstration of level of srf performance required in cw, high-average-current environment
- adequate damping of HOM *Q*s
- increased quality factor Q_0 of the superconducting cavities
- rf control and stability under maximum practical Q_L
- efficient extraction of HOM power
- development of multibunch BBU feedback systems

8. CONCLUSIONS

The confluence of high-performance srf and the realization of energy recovery offer prospects for accelerator systems providing linac-quality beam at storage-ring efficiencies. Initial implementations of ERL technology have demonstrated the viability of energy recovery at energies of tens to hundreds of MeV and at beam powers of hundreds of kilowatts. The next generation of test systems (the JLab FEL Upgrade and CEBAF-ER) will examine the GeV energy scale and manage megawatt levels of beam power.

As such systems evolve (the Cornell ERL prototype and the JLab 100-kW FEL driver), beam powers are expected to rise to tens of megawatts, allowing detailed investigations of beam stability and dynamics in regions of parameter space interesting to broad classes of users in the nuclear physics, high-energy physics, and photon science communities. Successful operation of these next-generation ERLs will set the stage for high-energy machines at the gigawatt scale, providing intense, high-quality beams for investigation of fundamental processes as well as the generation of photon beams at wavelengths spanning large portions of the electromagnetic spectrum.

ACKNOWLEDGMENTS

The authors thank generous colleagues for information and for significant improvements to the manuscript. I. Ben-Zvi (Brookhaven), W. Funk (JLab), P. O'Shea (University of Maryland), N. Vinokurov (Budker Institute of Nuclear Physics,

Novosibirsk), S. Schriber (Michigan State University), and T. Smith (Stanford University) provided scientific and technical input and useful references. S. Benson, S. Chattopadhyay, C. Leemann, and G. Neil carefully read and commented on the manuscript. S. Corneliussen's editorial advice was invaluable. This work was supported by U.S. Department of Energy contract DE-AC05-84-40150, the Office of Naval Research, Commonwealth of Virginia and the Laser Processing Consortium.

LITERATURE CITED

1. Neal RB, ed. *The Stanford Two Mile Accelerator.* New York: W.A. Benjamin (1968)
2. Seeman JT. *Annu. Rev. Nucl. Part. Sci.* 41:389–428 (1991)
3. Sands M. SLAC-0121 (1970)
4. Axel P, et al. *IEEE Trans. Nucl. Sci.* NS-24:1133 (1977)
5. Herminghaus H, et al. *IEEE Trans. Nucl. Sci.* NS-30:3274 (1983)
6. Penner S, et al. *IEEE Trans. Nucl. Sci.* NS-32:2669 (1985)
7. Leemann CW, Douglas DR, Krafft GA. *Annu. Rev. Nucl. Part. Sci.* 51:413–50 (2001)
8. Flanz JB, Sargent CP. *IEEE Trans. Nucl. Sci.* NS-32:3213 (1985)
9. Smith TI, et al. *Nucl. Instrum. Methods A* 259:1 (1987)
10. Tigner M. *Nuovo Cim.* 37:1228 (1965)
11. Kulypanov G, Skrinsky A, Vinokurov N. *Nucl. Instrum. Methods A* 467:16 (2001)
12. McAshan M, et al. *Proc. IX Int. Conf. High Energy Accel.*, p. 123 (1974)
13. Hanson AO, et al. *Proc. IX Int. Conf. High Energy Accel.*, p. 151 (1974)
14. Axel P, et al. *IEEE Trans. Nucl. Sci.* NS-24:1133 (1977)
15. Graef H-D, Richter A. *Proc. 1988 Linear Acc. Conf.*, p. 231
16. Genz H, et al. *Proc. Third Eur. Part. Acc. Conf.*, p. 49 (1992)
17. Sundelin R. *IEEE Trans. Nucl. Sci.* NS-32:3570 (1985)
18. Schriber SO, et al. *IEEE Trans. Nucl. Sci.* NS-24:1061 (1977)
19. Feldman DW, et al. *Proc. 1987 Part. Accel. Conf.*, p. 221
20. Sereno NS, et al. *Proc. 1993 Part. Accel. Conf.*, p. 3246
21. Krafft GA, Bisognano JJ. *Proc. 1989 Part. Accel. Conf.*, p. 1256
22. Neil G, Bisognano J, Dylla H, Krafft G. *Proc. 1991 Part. Accel. Conf.*, p. 2745
23. Bisognano JJ, et al. *Nucl. Instrum. Methods A* 318:216 (1992)
24. Neil GR, et al. *Nucl. Instrum. Methods* A 318:212 (1992)
25. Douglas DR, et al. JLab TN-91–071 (1991)
26. Neil GR, et al. *Phys. Rev. Lett.* 84:662 (2000)
27. Shinn M, et al. *Proc. SPIE* 3902:355 (2000)
28. Liu H, et al. *Nucl. Instrum. Methods A* 358:475 (1995)
29. Merminga L, et al. *Proc. 1999 Part. Accel. Conf.*, p. 1177
30. Li R. *Proc. 1999 Part. Accel. Conf.*, p. 118
31. Li R. *Proc. 1998 Eur. Part. Accel. Conf.*, p. 1230
32. Li R. *Nucl. Instrum. Methods A* 475:498 (2001)
33. Li R. *Proc. 2002 Eur. Part. Accel. Conf.*, p. 1365
34. Merminga L, et al. *Proc. Free Electron Laser Conf.*, pp. II–3 (2000)
35. Engwall D, et al. *Proc. 1997 Part. Accel. Conf.*, p. 2693

36. Piot P, et al. *Proc. 1998 Eur. Part. Accel. Conf.*, p. 1447
37. Piot P, et al. *Proc. 2000 Eur. Part. Accel. Conf.*, p. 1546
38. Flanz JB, Sargent CP. *Nucl. Instrum. Methods A* 241:325 (1985)
39. Piot P, Douglas DR, Krafft GA. *Proc. 2000 Eur. Part. Accel. Conf.*, p. 1543
40. Piot P, Douglas DR, Krafft GA. *Phys. Rev. ST-AB* 6:030702 (2003)
41. Siggins T, et al. *Nucl. Inst. Methods A* 475:549 (2001)
42. Piot P, et al. *1999 Part. Accel. Conf.*, p. 2229
43. Merminga L. *Nucl. Inst. Methods A* 483:107 (2002)
44. Benson S, et al. *Nucl. Inst. Methods A* 475:276 (2001)
45. Benson S, et al. *Proc. 1999 FEL Conf.*, p. II-1
46. Boyce JR, et al. *Proc. Conf. Appl. Accel. Research Industry.* In press (2003)
47. Carr GL, et al. *Nature* 420:153 (2002)
48. Douglas DR, Tennant C. JLab TN-01–043 (2001)
49. Douglas DR. JLab TN-01–048 (2001)
50. Douglas DR, et al. *Proc. 2000 Linear Accel. Conf.*, p. 857
51. Douglas DR, et al. *Proc. 2001 Part. Accel. Conf.*, p. 249
52. Nishimori N, et al. *Proc. 2002 Eur. Part. Accel. Conf.*, p. 822
53. Antokhin EI, et al. *J. Synchrotron Radiat.* In press (2003)
54. Lee BC, et al. *Proc. 2002 Linear Accel. Conf.*, p. 857
55. Zhang Y, Krafft GA, eds. *ICFA Beam Dyn. Newsl. No. 26* (2001)
56. Krafft GA. *Proc. 1999 Part. Accel. Conf.*, p. 2448
57. Kulypanov G, Skrinsky A, Vinokurov N. *J. Synchrotron Radiat.* 5:176 (1998)
58. Bazarov I, et al. CHESS Tech. Memo 01–003, JLAB-ACT-01–04 (2001)
59. Gruner S, et al. *Rev. Sci. Inst.* 73:1402 (2002)
60. Ben-Zvi I, et al. *Proc. 2001 Part. Accel. Conf.*, p. 350
61. Poole MW, Clarke JA, Seddon EA. *Proc. 2002 Euro. Part. Accel. Conf.*, p. 733
62. Steffens E, Schindler U, eds. *Proc. ERL-SYN 2002 Workshop* (2002)
63. Padmore HA, Schoenlein RW, Zholents AA. *Synchrotron Radiat. News* 14:26 (2001)
64. Bazarov I, et al. *Proc. 2002 Eur. Part. Accel. Conf.*, p. 644
65. MacKay W, Ben-Zvi I, et al. *Proc. 2001 Part. Accel. Conf.*, p. 3126
66. Parkhomchuk V, Ben-Zvi I. BNL C-A/AP/47 (2001)
67. Ben-Zvi I, Kewisch J, Murphy J, Peggs S. *Nucl. Instrum. Methods A* 463:94 (2001)
68. Merminga L, Krafft GA, Lebedev V. *Proc. 18th Int. Conf. High Energy Accel.* (2001)
69. Merminga L, Krafft GA, Lebedev VA, Ben-Zvi I. *A.I.P. Conf. Proc.* 588:204 (2001)
70. Peggs S, Ben-Zvi I, Kewisch J, Murphy J. *Proc. 2001 Part. Accel. Conf.*, p. 37
71. Merminga L, et al. *Proc. 2002 Eur. Part. Accel. Conf.*, p. 203
72. Wei J, Merminga L, eds. *ICFA Beam Dyn. Newsl. No. 30* (2003)
73. Douglas DR. *Proc. 1993 Part. Accel. Conf.*, p. 584
74. Kulipanov G, Skrinsky A, Vinokurov N. *Proc. ERLSYN 2002 Workshop*, p. 1 (2002)
75. Douglas DR, JLAB TN-98–040 (1998)
76. Free Electron Lasers for Industry, Volume 2: UV Demo Conceptual Design, JLab publication (1995)
77. Douglas DR. JLab TN-00–027 (2000)
78. Douglas DR. JLab TN-97–038 (1997)
79. Douglas DR. JLab TN-01–053 (2001); JLab TN-02–002 (2002)
80. Douglas DR. JLab TN-01–018 (2001)
81. Douglas DR. JLab TN-01–045 (2001)
82. Sinclair CK. *Proc. 1999 Part. Accel. Conf.*, p. 65
83. Bisognano JJ, Fripp ML. *Proc. 1988 Linear Accel. Conf.*, p. 388
84. Merminga L, et al. *Nucl. Inst. Methods A* 429:58 (1999)
85. Krafft GA, Bisognano J, Laubach S. JLab TN-01–011 (2001)

86. Merminga L, et al. *Proc. 2001 Part. Accel. Conf.*, p. 173
87. Rand RE. 1984. *Recirculating Electron Accel.*, Vol. 3, Accel. Storage Rings. New York: Harwood Acad., 236 pp.
88. Bisognano JJ, Gluckstern RL. *Proc. 1987 Part. Accel. Conf.*, p. 1078
89. Yunn BC. *Proc. 1991 Part. Accel. Conf.*, p. 1785
90. Krafft GA, Bisognano JJ. *Proc. 1987 Part. Accel. Conf.*, p. 1356
91. Beard KB, Merminga L, Yunn BC. *Proc. 2003 Part. Accel. Conf.* In press
92. Merminga L, Campisi IE. *Proc. 1998 Linear Accel. Conf.*, p. 460
93. Sereno NS. *Experimental studies of multipass beam breakup and energy recovery using the CEBAF injector linac*. PhD thesis. Univ. Ill., Urbana-Champaign (1994)
94. Merminga L, et al. *Proc. 2000 Linear Accel. Conf.*, p. 860
95. Delayen JR. *Phase and amplitude stabilization of superconducting resonators*. PhD thesis. Calif. Inst. Tech. (1978)
96. Merminga L, Delayen JR. JLab TN-96–022 (1996)
97. Hovater C, et al. *Proc. 2002 Linear Accel. Conf.* In press
98. Liepe M, Moeller WD, Simrock SN. *Proc. 2001 Part. Accel. Conf.*, p. 1074

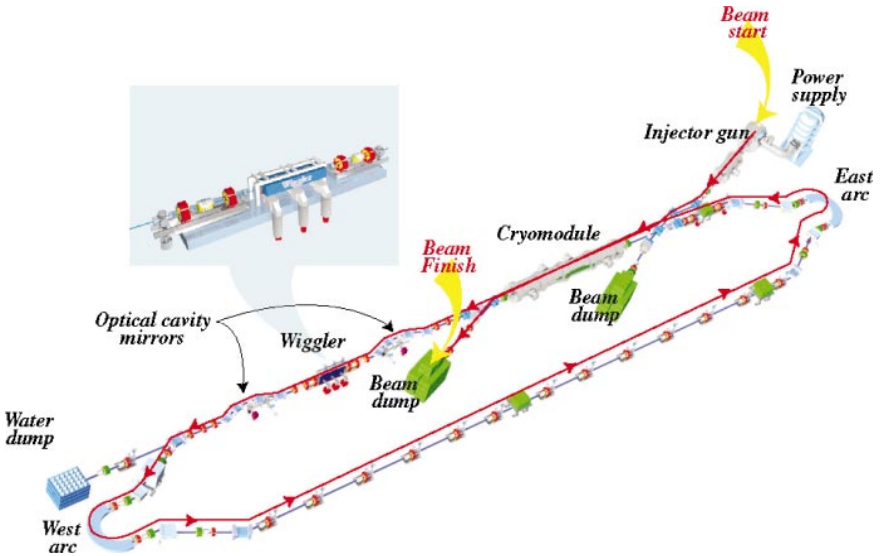

Figure 2 The Jefferson Laboratory Infrared Demonstration Free-Electron Laser.

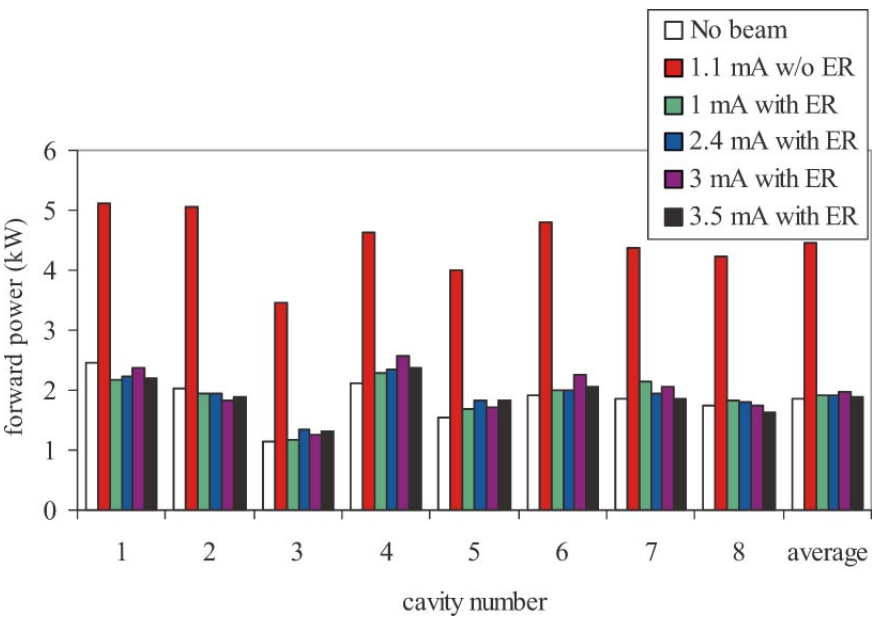

Figure 5 RF system generator power for each linac cavity without beam, without and with energy recovery at various current levels.

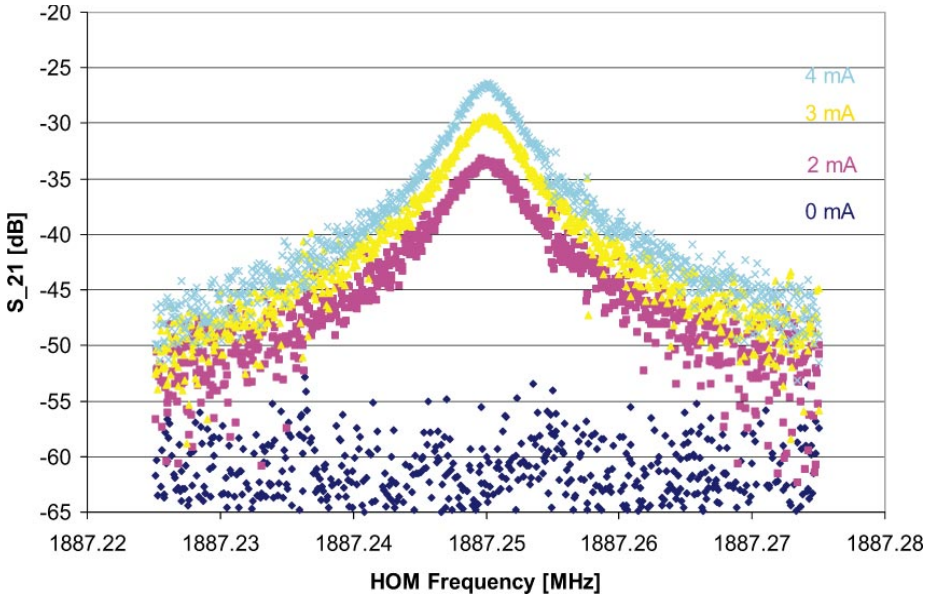

Figure 14 RF cavity response to beam excitation at the higher-order mode frequency of 1887 MHz at various beam currents from 0 to 4 mA.