FOURTH GENERATION LIGHT SOURCES*
Herman Winick, Stanford Linear Accelerator Center
Stanford, CA 94309, USA

1 INTRODUCTION
Since the first indirect observation of synchrotron radiation (SR) in 1945[1] there has been a rapid growth in its scientific use. Starting in the 1950’s cyclic electron synchrotrons were used, yielding to the superior properties of electron storage rings starting in 1968. Storage ring sources have evolved through three generations. First generation rings are those built for high energy physics research. The second generation are those built from the start as light sources. Third generation rings coming online since 1992 have many straight sections for insertion devices and lower electron beam emittance. Undulators on third generation rings provide ~10⁴ higher brightness than bend magnet sources of earlier rings. There are now about 40 operational rings of all generations used as SR sources in 14 countries, 10 of which are third generation sources.

As remarkable as this performance improvement has been, even higher brightness and laser-like coherence appear achievable and are needed scientifically, particularly at soft and hard X-ray wavelengths. Reaching higher performance levels is the goal of fourth generation sources which we may define as sources which exceed the performance of previous sources by one or more orders of magnitude in an important parameter such as brightness, coherence, or shortness of pulse duration. The most promising directions for fourth generation sources in the wavelength range from the VUV to hard X-rays are storage rings with even lower emittance than third generation rings, and short wavelength free-electron lasers (FELs) which offer sub-picosecond pulses with full transverse coherence.

The extraordinary properties of SR stem largely from the fact that the copious emission by relativistic electrons curving in magnetic fields is concentrated into an instantaneous forward cone with opening angle given by \( \gamma \) \({\gamma}^{-1} = mc^2/E \), the electron’s rest mass energy divided by its total energy. For example, this angle is only 0.1 mrad at 5 GeV. This small natural emission angle is key to understanding the properties of SR and the characteristics of the different source generations and types of insertion devices.

2 EARLY SOURCES
From the early 1950’s to the early 1970’s cycling electron synchrotrons, developed for high energy physics research, were used as SR sources. These are the zeroth generation. Although their SR is intense, cycle-to-cycle fluctuations and spectral, intensity, and source position changes within a cycle pose limitations. With the development of high energy physics storage rings, SR became available with constant spectrum and source position, and long stored-beam lifetime. These are the first generation SR sources. The superior radiation from these rings led to a rapid growth in SR programs, and their evolution from a parasitic effort to partly dedicated, and often fully dedicated, use of the ring.

Radiation from the bend magnets of first generation rings provided about 10⁵ times more tunable, continuum radiation than conventional sources, including rotating-anode X-ray tubes. The immediate successful use of this radiation, even in parasitic operation, resulted in an explosion of scientific interest[2]. The demand for SR in the mid-1970’s led Europe, Japan and the US to construct second generation SR sources; rings fully dedicated to SR research. When designs of the first round of these were finalized there was no experience using SR from wiggler and undulator insertion devices. Thus these rings were designed for many bend magnet beam lines, and a few locations in which insertion devices could be added later.

3 WIGGLERS AND UNDULATORS
Starting in 1978 wiggler and undulator insertion devices (periodic magnets placed between the bending magnets of a ring) were tested in first generation rings[3], offering higher flux, brightness, and spectral range than bend magnet sources. Although wigglers and undulators are both periodic magnetic structures, they produce different spectra due to the different angular deflection in each pole. For a wiggler this deflection is larger than the natural emission angle of synchrotron radiation \((\gamma^{-1} = mc^2/E)\). For an undulator it is typically \(\gamma^{-1} \), so undulators provide more concentrated radiation than wigglers or bend magnets. Furthermore, the small deflection in each undulator pole means that poles can be short, so that more can be accommodated in a given length. Permanent magnet technology has had a major impact on insertion device design, since the absence of coils allows for even more poles.

As an electron traverses an undulator, interference in the radiation at each of the collinear source points enhances intensity at certain wavelengths, resulting in a quasi-monochromatic spectrum rather than the broad continuum of bend magnet and wiggler sources. Peaks occur at wavelengths given by \(\lambda = \lambda_0[1 + K^2/2 + \gamma^2]/(2\gamma^2)\) and

* Work supported by the US Dept. of Energy, Office of Basic Energy Sciences, Contract DE-AC03-76SF00515
its harmonics. $\lambda_u$ is the undulator period, $\theta$ is the observation angle, and $K=0.934B[T]\lambda_u[cm]$ is the angular deflection in each pole in units of $\gamma^{-1}$. Peaks are tuned by varying the electron energy or the undulator field.

4 BRIGHTNESS, EMITTANCE, AND THIRD GENERATION SOURCES

The concentration of the radiation is called the brightness, measured in photons/(s,mm$^2$,mrad$^2$,0.1%bandwidth). The brightness produced by a beam of electrons depends on the electron beam transverse size and divergence, the product of which is called the emittance. Horizontal emittance ($\varepsilon_x=\sigma_x\sigma_y$) is determined by the electron energy and the ring design. The vertical ($\varepsilon_y=\sigma_x\sigma_y$) depends primarily on coupling to the horizontal and can be as small as $\approx0.5\%$ of the horizontal emittance. Second generation sources were originally designed with horizontal emittances of one hundred to several hundred nm-rad, resulting in undulator beam brightness of up to $10^{16}$.

Since further reduction of electron beam emittance would result in even higher brightness, in the mid-1980’s efforts began to design and construct a new round of storage rings, the third generation sources. These have many straight sections for insertion devices and electron beam emittance of about 5-20 nm-rad. These rings began operation in the early 1990’s, reaching undulator brightness as high as $10^{20}$, opening new opportunities for research. A brightness of $10^{20}$ is about $10^{13}$ times higher than that provided by rotating-anode X-ray tubes. Although spectacular, this brightness is far from fundamental limits. Further reduction of electron beam emittance would result in increased photon beam brightness, particularly at X-ray wavelengths. Achieving this is one of the most important objectives of fourth generation sources.

5 DIFFRACTION LIMITS

Diffraction sets an ultimate limit, on the geometric properties of photon beams. Because of diffraction the lower limit on the photon beam emittance is given approximately by the wavelength, $\lambda$. Using standard deviation values for Gaussian distributions, this diffraction-limited photon beam emittance is given by $\lambda/4\pi$. For light produced by electron beams, photon beam brightness increases as electron beam emittance decreases until the electron beam emittance reaches a value of $\approx \lambda/4\pi$. Thus third generation rings with an emittance of 5 nm-rad can produce diffraction-limited light at wavelengths longer than $\approx 60$ nm, (photon energies below $20$ eV). An emittance three orders of magnitude lower, 5 pm-rad, would be needed to reach the diffraction limit at 0.06 nm (20 KeV).

6 FOURTH GENERATION SOURCES

As mentioned earlier, we consider a light source to be fourth generation if it exceeds the performance of previous sources by an order of magnitude or more in an important parameter such as brightness, coherence, or pulse duration. The main directions that have emerged for fourth generation light sources in the wavelength range from the VUV to hard X-rays are lower emittance rings and short wavelength FELs using both rings and linacs as drivers. Linac-based FELs offer sub-picosecond pulses compared to 10-50 ps for present storage rings. However, it may be possible to operate existing rings or design new rings with low momentum compaction factor[4] to produce a pulse duration of $\approx 1$ ps, albeit with relatively low current.

6.1 Lower Emittance Storage Rings

The relative ease with which third generation light sources have reached, and indeed exceeded, design goals indicates that fourth generation storage rings can reach even lower electron beam emittance, producing higher photon beam brightness and diffraction-limited light at shorter wavelengths. The challenges that must be met to accomplish this have been considered at workshops on fourth generation light sources[5,6]. Of concern are the various aspects of stability of the electron beam; position stability, reproducibility, single and multi-bunch instabilities, etc. A variety of countermeasures (such as feedback systems, Landau cavities, high-harmonic cavities) have been successfully developed to deal with these problems in presently operating rings. These will need to be pushed to higher performance levels to meet the stability demands of fourth generation rings. A major obstacle is the reduced lifetime and increased emittance due to intrabeam scattering (the Touschek effect) as bunch density increases.

6.1.1 Lattice design and dynamic aperture

A formidable challenge in the design of fourth generation storage ring sources is to develop a very low emittance magnet lattice with sufficient dynamic aperture to accommodate stable orbits with large amplitude oscillations resulting from, for example, Coulomb scattering of electrons on the residual gas and off-axis injection. The former results in the continuous population of a halo much larger than the core. As the large amplitude betatron oscillations of particles in this halo are damped, they coalesce with the core. If the aperture (dynamic or physical) is too small, particles in the halo are lost before being damped, reducing lifetime. It is difficult to maintain a large dynamic aperture in a low-emittance lattice because of the chromatic effects of the strong quadrupoles. This chromaticity, the energy dependence of the betatron tune, is corrected by sextupole magnets, whose non-linear fields reduce the dynamic aperture.

A possible countermeasure is the “modified sextupole”[7], which provides a magnetic field with a quadratic dependence over the core of the beam, but which then levels off or rises much less rapidly with distance from the axis, thereby lowering the non-linear fields experienced by particles with large amplitude oscillations. Dynamic aperture might also be enlarged by alternately rotating the lattice cells by $+/- 45^\circ$, so that sextupoles can be placed at
locations of maximum dispersion in each plane for efficient chromaticity correction[8].

The large dynamic aperture needed for injection in present rings is due to the fact that stored beam is accumulated with off-axis injection of many low intensity "shots", each of which executes large amplitude betatron oscillations until they coalesce with the already stored beam due to radiation damping. The aperture requirement can be reduced with single-shot, on-axis injection from another ring, in which a high intensity beam has been accumulated with multi-shot, off-axis injection. Injection into synchrotron phase space is another possibility.

The horizontal emittance in an electron storage ring scales as the square of the electron energy and the third power of the bend magnet length. Thus, lower emittance fourth generation rings would have many bend magnets separated by quadrupoles, and many straight sections for insertion devices, leading to larger circumference at a given energy than third generation rings. For example, the lattice working group at the Grenoble Workshop[6] presented a “straw-man” design for a 2-3 GeV fourth generation ring with ~0.3 nm-rad emittance and a circumference the same as the 6 GeV ESRF machine, ~850 m. LBNL is studying a 2 GeV ring[9] also with ~0.5 nm-rad emittance and a circumference of about 350 m. Such rings might achieve a brightness at soft X-ray wavelengths of about 5x10^23, more than 3 orders of magnitude greater than third generation VUV sources. Note that 0.3 nm-rad is the diffraction limit for light at 3.6 nm, or 0.34 keV.

Fourth generation rings for hard X-rays (below ~2 Å) would require higher electron energy and larger circumference. They would cost much more than the lower energy rings discussed above, even to reach an emittance of about 0.3 nm-rad, which is much larger than the diffraction limit for hard X-rays. Limited use has been made of undulators on the large circumference PEP[10] and TRISTAN[11] rings as third generation sources. However both are now being converted to B-Factories. Although the PETRA ring is part of the HERA injection system, an undulator has been installed in PETRA for use between HERA injections. Operating at 12 GeV, this undulator provides third generation brightness extending to very high photon energy. In the future fourth generation, hard X-ray rings may be installed in these tunnels.

6.1.2 Beam lifetime - Touschek effect

Very low emittance fourth generation rings will have very high bunch charge density, leading to short lifetime due to the collisions of electrons within a bunch; the Touschek effect. This is particularly severe at low energy and is already a problem in third generation 1-2 GeV rings. To achieve lifetime of the order of 10 hours, the bunch density must be reduced in several third generation rings. This is usually done by increasing the vertical emittance above its minimum value, trading brightness for lifetime. Multiple Touschek scattering also enlarges emittance. If all beam dimensions and the charge per bunch are kept constant, Touschek lifetime increases quadratically with electron energy and with the cube of the energy acceptance of the rf system. Thus lower emittance VUV/soft X-ray rings are designed with higher electron energy [the Swiss Light Source (2.1 GeV), Soleil in France(2.15 GeV) and the Shanghai Synchrotron Radiation Facility (2.2-2.5 GeV)] and with large rf overvoltage to increase the energy acceptance. The Swiss Light Source plans to use a high-Q superconducting passive rf cavity, tuned several bandwidths away from the main rf system, to increase the rf overvoltage and energy acceptance[12].

Short lifetime can also be compensated with frequent, or “top-up”, injection. The nearly constant stored current also keeps a constant heat load on beam line optical elements and compensates for lifetime reduction if small gap, short period undulators, which extend the spectral range, are used. “Top-up” injection is planned for the 7 GeV APS facility at Argonne National Laboratory[13].

6.1.3 Other considerations

Reducing the energy at which a given ring is operated can be used to reduce the emittance, taking advantage of the quadratic dependence of emittance on electron energy, as has been done at PEP[10] and TRISTAN[11]. However damping time constants increase and instability threshold currents decrease as energy is reduced, limiting the effectiveness of this approach. To some extent this can be compensated by making more radiation with damping wigglers, which also reduce the emittance[14].

Improving undulator field quality by shimming[15] extends their spectral range beyond the 5th harmonic, previously the highest that could be used in practice. It also opens the possibility of designing future rings to produce hard X-ray brightness comparable to that of third generation hard X-ray sources with lower electron energy than 6-8 GeV. For example, high harmonics of undulators in a 3.5-4 GeV ring with a circumference of about 300-400 m and an emittance of ~10 nm-rad could produce a brightness of ~10^18 or greater at photon energies up to ~15 keV. Brightness of high harmonics is determined by emittance, undulator errors, and electron energy spread. A comprehensive code taking all these into account has been developed at APS[16]. As emittance is reduced and undulators are made more perfect, the energy spread ultimately determines high harmonic brightness.

6.2 FELs Based on Storage Rings

FELs produce extremely high brightness, transversely coherent radiation by inducing a bunch-density modulation of the electron beam at the optical wavelength. This is achieved by the interaction of a bright electron beam with an intense optical field in the spatially periodic magnetic field of an undulator. When electrons are bunched within an optical wavelength, the power radiated varies as the number of electrons squared, rather than linearly as for an unbunched beam. FELs have operated at wavelengths
from the IR to the UV for many years, using storage rings such as ACO, TERA-(UV), UVSOR, VEPP-3, Super ACO, and others. Several storage rings have been designed with long straight sections to accommodate long FEL undulators. These include NIIH-IV in Japan and the new rings at Duke and Dortmund Universities. Long straight sections are also included in several proposed rings such as the Swiss Light Source, the Shanghai Synchrotron Radiation Facility, Soleil (France), and Diamond (U.K.).

Storage ring-based FELs provide light with very high brightness and coherence and may already be considered to be fourth generation sources in the wavelength range in which they now operate. Reviews of operating storage ring-based FELs and the prospects for future development, particularly the prospects for extending their operation to shorter wavelength, have been given[17]. Present storage ring FELs operate in the oscillator mode, using optical cavities to build up the radiation from many passes of the electron beam until the optical field is strong enough to induce a density modulation of the electron bunch at the optical wavelength, resulting in coherent, stimulated emission of radiation at that wavelength. It is difficult to make optical cavities at wavelengths below ~200 nm due to the lack of good reflectors. To overcome this, grazing incidence reflection, with higher reflectivity at shorter wavelength, can be used in a multiple-mirror, ring cavity configuration. Also, harmonics have been used to reach shorter wavelength. Using these approaches some groups are aiming for the 20-50 nm range.

An alternative approach is to eliminate the cavity and to achieve lasing in a single pass of a very bright electron beam through a long undulator, either by amplifying an input signal or with no input, in a process called self-amplified spontaneous emission (SASE)[18]. A design for such a single-pass, high-gain FEL amplifier operating down to 40 nanometers in a bypass of a 750 MeV storage ring was proposed at BNL[19] and was considered for PEP[20].

6.3 FELs Based on LINACs

FELs using low energy linacs have operated for several years, providing coherent infra-red radiation at several user facilities. These use optical cavities in oscillator configurations, as do the storage ring-based FELs. Recent developments open the possibility to construct much shorter wavelength FELs, using bright electron beams from high energy linacs to achieve lasing in a single pass through a long undulator. With no optical cavity the lack of good short wavelength reflectors is no longer a limitation. However, the demands on the electron beam and undulator quality are severe, particularly to reach Ångstrom wavelengths. The developments opening the path to single-pass FELs operating at such short wavelengths are:

1. Photocathode rf electron guns[21], which provide short (5-10 ps), 1 nC pulses with normalized emittance (geometric emittance times $\gamma$) approaching 1 mm-mrad.

2. Control over emittance degradation during acceleration and compression, as demonstrated in the SLAC SLC project. Based on this and subsequent studies[22] it appears possible to accelerate and compress the beam from the gun to produce multi-GeV, kiloampere beams with emittance approaching the diffraction limit at wavelengths down to a few Angstroms. Note that geometric emittance varies as $\gamma^{-1}$ in a linac, as opposed to $\gamma^2$ in a storage ring.

3. Precision undulators as have been built at many SR sources. These must be extended to 50-100 m lengths, while including distributed focusing and maintaining tight tolerances on magnetic properties and alignment.

Designs are being developed for single-pass FELs operating from the VUV to the Ångstrom range, using photocathode rf guns and bunch length compressors to achieve high peak current in sub-picosecond pulses. BNL will use an existing 230 MeV linac to reach ~75 nm in a deep UV FEL by harmonic generation and single-pass amplification[23]. At DESY the TESLA Test Facility (TTF) superconducting linac will be used to drive a single-pass FEL[24] for SASE tests at ~250 Å. The linac will then be extended to ~1 GeV for an FEL user facility operating down to ~60 Å. The DESY group also proposes[25] to include several SASE-based FELs, operating down to about 1 Å, as an integral part of a proposed 250 GeV-per-beam linear collider. Energies up to about 30 GeV will be used for the FEL. Two approaches are being considered; a 1.3 GHz superconducting linac operating for the FEL at 5 Hz with 11,300 microbunches in each macro pulse and a 3 GHz linac operating at 50 Hz with 125 microbunches per macro pulse. The goal is average brightness of $10^{24-25}$ and peak brightness of $10^{26.1-27}$.

The SLAC group[26] proposes to use the last third of the 3 km linac (the first 2 km will be used for injection to the B-Factory now in construction) to generate a 5-15 GeV beam for a 1.5-15 Å SASE FEL with an average brightness up to ~$10^{23}$ and a peak brightness up to ~$10^{24}$. A design report for the project, called the Linac Coherent Light Source (LCLS), is in preparation. With an available 15 GeV linac, SLAC provides an opportunity to study the SASE process at very short wavelengths and to start using the remarkable brightness, coherence and short pulse duration of an X-ray FEL.

The calculated LCLS beam properties at 1.5 Å are: bandwidth = 0.1%, pulse duration (FWHM) = 280 fs, peak coherent power = 10 GW, coherent ph/pulse = $2 \times 10^{12}$, coherent ph/s = $2 \times 10^{14}$ (120Hz), average coherent power = 0.3 W, transverse beam size (FWHM) = 70 microns, divergence (FWHM) = $10^7$ radians. In addition, use will also be made of the broad spectrum of spontaneous undulator radiation with the same pulse duration, several times higher peak power and a larger opening angle.
Many laboratories (ANL, BNL, DESY, LANL, SLAC, TJNAF, UCLA) are pursuing single-pass FEL r&d including: SASE studies at micron wavelengths of startup from spontaneous radiation, exponential gain\[27\], and saturation; studies of the effects of space charge and coherent SR in bunch-length compressors; undulator design and alignment; photocathode rf gun design and characterization; electron and photon diagnostics; and X-ray optics.

The projected characteristics of linac-based short wavelength FELs, particularly their short pulse duration, peak brightness, and coherence, are extraordinary. The peak brightness of the X-ray FELs proposed at SLAC and DESY is \(\sim 10^{10}\) times higher than that of third generation storage ring sources, with \(\sim 100\) times shorter pulses. These properties are likely to open entirely new opportunities in imaging, non-linear physics, and pump-probe experiments. There is increasing confidence in the accelerator community that linac-based, short wavelength FELs can be built. There is also increasing realization that their properties will open new science in the 21st century.

ACKNOWLEDGMENTS
It is a pleasure to acknowledge discussions with and comments on the manuscript by J. Barger, A. Bienenstock, R. Carr, J. Corbett, M. Cornacchia, J. Galayda, R. Gould, K. Halbach, A. Hofmann, A. Jackson, E. Johnson, C. Limborg, B. McSwain, H.-D. Nuhn, C. Pellegrini, J. Schmerge, R. Tatchyn, H. Wiedemann.

REFERENCES