Recent advances in ultrafast X-ray sources

Robert Schoenlein$^1$, Thomas Elsaesser$^2$, Karsten Holldack$^3$, Zhirong Huang$^1$, Henry Kapteyn$^4$, Margaret Murnane$^4$ and Michael Woerner$^2$

$^1$SLAC National Accelerator Laboratory, 2575 Sand Hill Road, Menlo Park, CA 94025, USA
$^2$Max-Born-Institut für Nichtlineare Optik und Kurzzeitspektroskopie, 12489 Berlin, Germany
$^3$Helmholtz-Zentrum Berlin für Materialien und Energie GmbH, Albert-Einstein-Strasse 15, 12489 Berlin, Germany
$^4$Department of Physics and JILA, University of Colorado, Boulder, CO 80309-0440, USA

Over more than a century, X-rays have transformed our understanding of the fundamental structure of matter and have been an indispensable tool for chemistry, physics, biology, materials science and related fields. Recent advances in ultrafast X-ray sources operating in the femtosecond to attosecond regimes have opened an important new frontier in X-ray science. These advances now enable: (i) sensitive probing of structural dynamics in matter on the fundamental timescales of atomic motion, (ii) element-specific probing of electronic structure and charge dynamics on fundamental timescales of electronic motion, and (iii) powerful new approaches for unravelling the coupling between electronic and atomic structural dynamics that underpin the properties and function of matter. Most notable is the recent realization of X-ray free-electron lasers (XFELs) with numerous new XFEL facilities in operation or under development worldwide. Advances in XFELs are complemented by advances in synchrotron-based and table-top laser-plasma X-ray sources now operating in the femtosecond regime, and laser-based high-order harmonic XUV sources operating in the attosecond regime.

This article is part of the theme issue ‘Measurement of ultrafast electronic and structural dynamics with X-rays’.
1. Introduction

Since their discovery by Röntgen in 1895, X-rays have revolutionized our fundamental understanding of matter, and thereby redefined chemistry, physics, biology and many related fields of science and technology. The broad scientific importance of X-rays is reflected by the 25 Nobel Prizes in physics, chemistry and medicine that have been awarded for research involving X-rays. This importance of X-rays for science stems from three key attributes: the short wavelengths of X-rays suitable for probing the structure of matter at the atomic scale via scattering; the chemical specificity of X-ray spectroscopy for characterizing the local bonding and structural environment of selected elements in complex matter; and the penetrating capabilities of X-rays for revealing the hidden interior of complex matter via X-ray imaging.

The development of synchrotron-based X-ray sources over the past 40 years has ushered in a modern age of X-ray science (leading to six Nobel Prizes since 1997) by harnessing high-energy electron accelerator technology to provide X-ray beams that are intense, highly directional and tunable over a wide wavelength range. Synchrotron sources have developed into large regional scientific centres, servicing thousands of experiments per year, and X-rays are being applied in disparate science areas such as environmental science, astrophysics and art history. While X-rays have qualitatively advanced our understanding of the electronic and atomic structure of matter, a significant historical limitation has been an inability to access the fundamental timescales of electronic and atomic motion—the structural dynamics that determine the functional properties of matter.

In parallel with the development of accelerator-based (long-pulse) X-ray sources, the field of ultrafast science has burgeoned in both scope and impact over the past 40 years, driven primarily by dramatic advances in table-top laser technology. One important hallmark of this development has been the continued push of ultrafast table-top laser capabilities towards shorter wavelengths. The two most successful approaches in this effort are (i) laser-based plasma X-ray sources, which generate incoherent hard X-ray pulses in the femtosecond regime via strong-field laser interaction, typically with solid-density targets [1], and (ii) high-order harmonic generation (HHG) [2–4], which creates coherent XUV pulses in the femtosecond to attosecond regimes via strong-field laser interaction, typically with gas-phase atomic targets. In addition, the combination of advanced femtosecond lasers with conventional synchrotron-based X-ray sources has enabled the generation of broadly tuneable femtosecond X-rays at these facilities via laser manipulation of the relativistic electron beam [5–7].

One of the most important recent advances in ultrafast X-ray sources has been the emergence of X-ray free-electron lasers (XFELs), which have sparked another revolution in X-ray science that promises to transform the field for the twenty-first century [8,9]. XFELs generate tuneable coherent X-ray beams, many orders of magnitude brighter than any previously available X-ray source, and have been enabled by advances in electron accelerator technology. Current XFEL pulse durations are typically in the 10–100 fs range, but subfemtosecond pulses have recently been demonstrated [10]. The pulsed nature of the X-rays invite their use as a probe of sample dynamics on a fast timescale appropriate for observing atomic motions. The very high pulse intensity can create a significant dataset from a single X-ray pulse. Whereas previous X-ray sources, including synchrotron sources, have primarily engaged in studies of static structures, XFELs are by their nature suited for studying dynamic systems at the time and length scales of atomic interactions.

This paper provides an overview of the most recent advances in accelerator-based and table-top ultrafast X-ray sources that are currently in operation and able to probe electronic and atomic structural dynamics on fundamental timescales, thereby opening important new scientific frontiers.

2. X-ray free-electron lasers

Free-electron lasers (FELs), invented by Madey [11], and subsequently demonstrated experimentally by his group at Stanford University in the 1970s [12], use relativistic electrons
propagating through a periodic magnetic structure (undulator) to generate and amplify coherent electromagnetic radiation. FELs have the widest frequency tuning range of any type of laser, and can generate very high peak and average power light beams. From their initial demonstration in the infrared regime [12], there has been continued effort to extend FEL operation towards the EUV and X-ray regimes, where fundamental principles inhibit the efficient operation of conventional (quantum) lasers [13]. To circumvent the need for mirrors or coherent seed sources at short wavelengths, the initial random radiation spontaneously emitted by relativistic electrons in an FEL undulator may be further amplified in a medium comprised a bright electron beam propagating through a long undulator, thereby creating intense, quasi-coherent radiation [14,15]. A high-gain XFEL operated in this self-amplified spontaneous emission (SASE) mode can generate multi-gigawatt (GW) and femtosecond (fs) coherent pulses in the X-ray wavelength range (from a few nanometres to less than 1 Å). The extremely high power together with the excellent transverse coherence of such XFELs provide a dramatic enhancement in peak brightness (approx. $10^9$–$10^{10}$-fold) beyond that offered by the current synchrotron radiation sources based on electron storage rings. Thus, XFELs are very powerful probes for both the ultra-small and the ultrafast worlds.

Consider a planar magnetic undulator with a sinusoidal vertical magnetic field. The peak field strength is $B_0$, and the undulator period is $\lambda_u$. An electron beam with the energy $\gamma mc^2$ executes a nearly sinusoidal trajectory in the undulator and emits fundamental undulator radiation at the wavelength

$$\lambda_r = \frac{\lambda_u}{2\gamma^2} \left( 1 + \frac{K^2}{2} \right) = \frac{2\pi c}{\omega_r},$$  

(2.1)

where $K = eB_0\lambda_u/(2\pi mc)$ is the undulator strength parameter (i.e. $K^2$ is the electron ponderomotive energy normalized to the rest energy), $e$ is the charge of the electron, $c$ is the speed of light in vacuum, $m$ is the electron mass, and $\gamma$ is the electron energy in units of the rest energy $mc^2$. Higher harmonic radiation, especially odd harmonics, are also generated at reduced intensity [16].

For a sufficiently bright electron beam and a sufficiently long undulator, the resonant interaction leads to an exponential growth of the fundamental radiation intensity along the undulator. Such a high-gain XFEL does not require an optical cavity or external seed, and the initial spontaneous undulator radiation is amplified through the SASE process. The scaling behaviour of a high-gain XFEL amplifier in the one-dimensional (1D) limit can be well characterized by the so-called FEL Pierce parameter [15]

$$\rho = \left[ \frac{1}{64\pi^2} \frac{I_p}{I_A} \frac{\lambda_u}{\gamma x^2} \frac{K^2 [JJ]^2}{\xi^3} \right]^{1/3},$$  

(2.2)

where the Bessel function factor $[JJ] = [J_0(\xi) - J_1(\xi)]$ with $\xi = K^2/(4 + 2K^2)$ for a planar undulator, $I_p$ is the electron peak current, $I_A \approx 17$ kA is the Alfvén current, and $\sigma_x$ is the rms transverse size of the electron beam. The power grows exponentially with undulator distance $z$: $P(z) \propto \exp(z/L_G)$ with the power gain length

$$L_G \approx \frac{\lambda_u}{4\pi \sqrt{3\rho}},$$  

(2.3)

At saturation, XFEL peak power is given approximately by

$$P_{sat} \approx \frac{\rho \gamma mc^2 I_p}{e}.$$  

(2.4)

Here, $\rho$ (typically approx. $10^{-3}$) characterizes the efficiency of the XFEL in terms of the electron beam power.

SASE XFELs have excellent transverse coherence as the high-gain process selects a dominant fundamental mode [17]. However, due to the large gain bandwidth of the SASE process, and
initiation from shot noise, SASE XFEL radiation has limited temporal coherence and exhibits shot-to-shot fluctuations in intensity. The normalized rms bandwidth and coherence length of SASE radiation at saturation \([18,19]\) are given by

\[
\frac{\sigma_\omega}{\omega_r} \approx \rho, \quad (2.5)
\]

and

\[
L_{\text{coh}} \approx \sqrt{\frac{\pi c}{\sigma_\omega}} = \frac{\lambda_r}{2\sqrt{\pi\rho}}. \quad (2.6)
\]

For a ‘flat-top’ electron bunch of length \(L_b > L_{\text{coh}}\), the average number of temporal spikes in a SASE pulse is \(M = L_b / L_{\text{coh}}\). For typical parameters of operation of the LCLS, the output of roughly 1000 micro e-bunches results in an approximately 1 fs coherent spike. A typical FEL pulse of \(10^{12}\) photons will be composed of a few tens to hundreds of coherent spikes with no fixed phase relation to each other. In the exponential gain regime, the relative rms fluctuation of the SASE intensity is proportional to \(1/\sqrt{M}\) because these spikes are independent radiation sources.

Extremely bright electron beams and very high-quality undulator arrays are required to drive XFELs. A high-energy linear accelerator or linac, with an appropriate injector and electron transport optics, can deliver electron beam brightness at requisite levels for successful XFEL operations. While other options are possible, an electron beam from a linac source is typically more than 1000-fold brighter than standard storage rings, and is a natural choice for initial X-ray FEL studies. Once such a high-brightness beam is created in a linac, it must be temporally compressed to achieve a high-peak current and then accelerated to the undulator entrance with a small transverse cross-section. Both higher peak current and smaller transverse cross-section increase the FEL Pierce parameter and reduce the one-dimensional gain length (cf. equations (2.2) and (2.3)) while the electron beam energy spread and angular spread, due to finite emittance, increases the overall gain length from the one-dimensional limit. The FEL design optimization is therefore multi-dimensional and is well beyond the scope of this paper, but the typical requirements on electron beams are

\[
I_p \geq 1 \text{ kA}, \quad \frac{\sigma_\gamma}{\gamma} \leq \frac{\rho}{2\gamma} \frac{\varepsilon_N}{4\pi} \sim \frac{\lambda_r}{4\pi}. \quad (2.7)
\]

where \(\varepsilon_N\) is the normalized emittance in the transverse direction. Note that these requirements apply to the time sliced beam qualities defined on the scale of the coherence length (cf. equation (2.6)), rather than the beam qualities projected over the entire bunch length. This adds additional flexibilities to temporal manipulations of the electron bunch in order to shape the temporal profiles of the X-ray pulses.

(a) X-ray free-electron laser facilities: present status

A new era in X-ray science has been launched by the development of XFELs. The FLASH facility at DESY in Hamburg (the first EUV XFEL—2005) [8], the LCLS facility at SLAC National Accelerator Laboratory (the first hard X-ray XFEL—2009), the SACLA facility in Japan [20] and the FERMI facility in Trieste (first seeded EUV XFEL) [21] represent the first generation of XFEL sources, and they have already demonstrated the tremendous scientific potential and impact across broad areas of science. XFEL facilities are now proliferating around the world, with the Pohang accelerator laboratory (PAL)-XFEL in Pohang [22], the SwissFEL in Switzerland [23] and the European XFEL (EuXFEL) in Hamburg [24] all beginning operations in 2016–2017 [25].

Table 1 provides a summary of characteristic capabilities of the operating XFEL facilities worldwide, and the projected capabilities of the LCLS-II and LCLS-II-HE upgrades, now underway, to the LCLS facility in the USA. A new XFEL project in Shanghai, China, has recently been initiated (and is not included in table 1). The Shanghai XFEL will based on a new 8 GeV continuous-wave superconducting RF accelerator (CW-SCRF) that will drive three undulator beamlines (at a repetition rate up to 1 MHz) to cover a spectral range comparable to LCLS-II-HE.
Table 1. The characteristic parameters of the various XFEL facilities worldwide. These includes the supporting country, name of the facility, electron and photon beam energy ranges, pulse energy and length, repetition rate and the year of start of operation.

<table>
<thead>
<tr>
<th>country</th>
<th>name</th>
<th>electron (keV)</th>
<th>photon (keV)</th>
<th>X-ray pulse energy (μJ)</th>
<th>X-ray pulse length (fs)</th>
<th>rep rate (Hz)</th>
<th>start of operation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Japan</td>
<td>SACLA BL2,3</td>
<td>6–8</td>
<td>4–20</td>
<td>0.1–1</td>
<td>2–10</td>
<td>60</td>
<td>2011</td>
</tr>
<tr>
<td></td>
<td>SACLA BL1</td>
<td>0.8</td>
<td>0.04–0.15</td>
<td>0.1</td>
<td>60</td>
<td>60</td>
<td>2015</td>
</tr>
<tr>
<td>Italy</td>
<td>FERMI-FEL-1</td>
<td>0.9–1.5</td>
<td>0.01–0.06</td>
<td>0.08–0.2</td>
<td>40–90</td>
<td>10 (50)</td>
<td>2010</td>
</tr>
<tr>
<td></td>
<td>FERMI-FEL-2</td>
<td>0.9–1.5</td>
<td>0.06–0.3</td>
<td>0.01–0.1</td>
<td>20–50</td>
<td>10 (50)</td>
<td>2012</td>
</tr>
<tr>
<td>Germany</td>
<td>FLASH1</td>
<td>0.4–1.25</td>
<td>0.02–0.3</td>
<td>0.01–0.5</td>
<td>30–200</td>
<td>(1–800) × 10^5</td>
<td>2005</td>
</tr>
<tr>
<td></td>
<td>FLASH2</td>
<td>0.5–1.25</td>
<td>0.01–0.3</td>
<td>0.01–1</td>
<td>10–200</td>
<td>(1–800) × 10</td>
<td>2016</td>
</tr>
<tr>
<td>Korea</td>
<td>PAL-XFEL</td>
<td>4–11</td>
<td>2.5–15</td>
<td>0.8–1.5</td>
<td>5–50</td>
<td>60</td>
<td>2016</td>
</tr>
<tr>
<td></td>
<td>SwissFEL</td>
<td>2–5.8</td>
<td>1.8–12.4</td>
<td>1</td>
<td>10–70</td>
<td>100</td>
<td>2017</td>
</tr>
<tr>
<td>Switzerland</td>
<td></td>
<td>3</td>
<td>0.25–1.2</td>
<td>0.2</td>
<td>5–50</td>
<td>60</td>
<td>2016</td>
</tr>
<tr>
<td>Europe</td>
<td>XFEL-SASE1,2</td>
<td>8.5–17.5</td>
<td>3–25</td>
<td>2</td>
<td>10–100</td>
<td>2700 × 10^4</td>
<td>2017</td>
</tr>
<tr>
<td></td>
<td>XFEL-SASE3</td>
<td>8.5–17.5</td>
<td>0.2–3</td>
<td>2</td>
<td>10–100</td>
<td>2700 × 10</td>
<td>2017</td>
</tr>
<tr>
<td>USA</td>
<td>LCLS</td>
<td>3–15</td>
<td>0.3–12</td>
<td>2–4</td>
<td>2–500</td>
<td>120</td>
<td>2009</td>
</tr>
<tr>
<td></td>
<td>LCLS-II</td>
<td>3–15</td>
<td>1–25</td>
<td>2–4</td>
<td>10–100</td>
<td>120</td>
<td>2021</td>
</tr>
<tr>
<td></td>
<td>LCLS-II</td>
<td>4</td>
<td>0.2–5</td>
<td>0.02–1</td>
<td>10–200</td>
<td>10^6</td>
<td>2021</td>
</tr>
<tr>
<td></td>
<td>LCLS-II-HE</td>
<td>4–8</td>
<td>0.2–13</td>
<td>0.02–1</td>
<td>10–200</td>
<td>10^6</td>
<td>2026^e</td>
</tr>
</tbody>
</table>

^aRough estimates of pulse energy, particularly for projects that are currently under construction.
^bEstimated FWHM of X-ray pulse length based on electron bunch length measurement or designed range.
^cBurst mode operation at 10 Hz, with each macropulse providing up to 800 bunches at 1 MHz.
^dPulsed mode operation at 10 Hz, with each macropulse providing up to 2700 bunches at 5 MHz.
^eProjected project completion date for LCLS-II-HE.

The first generation of XFEL facilities has been largely based on GeV-scale electron beams generated by conventional (normal conducting) pulsed-RF accelerator technology operating at repetition rates in the 10–120 Hz range. Two notable exceptions are the FLASH and European XFEL facilities that are based on pulsed superconducting accelerator technology. This supports a burst-mode of operation providing macro-pulses at 10 Hz, with each macropulse comprised up to 800 (FLASH) or 2700 (EuxFEL) micro-pulses. A new generation of XFELs is now under development that will exploit continuous-wave superconducting RF accelerator technology (CW-SCRF) to provide ultrafast X-ray pulses at high repetition rate (in a uniform or programmable time distribution). This development is driven by important new science opportunities that have been identified and advanced over the past decade through scientific workshops around the world. The LCLS upgrade project (LCLS-II), as described below, will be the first of this new generation of XFELs providing soft and tender X-rays at repetition rates up to 1 MHz [26].

(i) Linac coherent light source

The Linac Coherent Light Source (LCLS) at the SLAC National Accelerator Laboratory began operations in 2009, and was the first XFEL facility to operate in the hard X-ray range, at 1.5 Å [9,27]. LCLS uses the last third of the historic SLAC Cu-linac which accelerates electrons up to 17 GeV at 120 Hz. X-rays in the energy range from 280 to 12.8 keV (fundamental emission) are generated via the SASE process from a single fixed-gap undulator array with 30 mm period and 132 m total length. The X-ray pulse duration from LCLS is adjustable from approximately 2–500 fs,
with energies up to $6 \text{ mJ pulse}^{-1}$ (depending on the pulse duration) [28,29]. Recently, less than 1 fs pulses have been generated from LCLS using nonlinear e-bunch compression techniques as described below [10]. LCLS achieves some control over the longitudinal coherence via self-seeding based on a grating monochromator in the soft X-ray range between 500 eV and 1 keV [30] and based on a diamond Bragg crystal transmission monochromator in the hard X-ray range between 5.5 and 9.5 keV [31]. As described in further detail below, self-seeding increases the spectral brightness of LCLS up to a five-fold compared to SASE operation. The incorporation of a Delta-undulator after-burner in LCLS has enabled the generation of circularly polarized X-ray pulses in the soft X-ray range between 500 and 1500 eV with energies of a few hundred $\mu\text{J pulse}^{-1}$ [32]. Finally, a number of schemes have been developed at LCLS for generation of multiple X-ray pulses at different photon energies [33–36], these are discussed further below.

Two major upgrades of LCLS are presently underway as indicated in table 1. The LCLS-II upgrade [37] (projected first light in 2021) will provide ultrafast X-rays in the 0.25–5 keV range at repetition rates up to 1 MHz with two independent XFELs based on two new adjustable-gap undulator arrays: the soft X-ray undulator (SXU) spans the range from 0.25 to 1.6 keV, and the hard X-ray undulator (HXU) spans the range from 1 to 5 keV [37]. LCLS-II will be based on a new 4 GeV CW-SCRF linac installed in the first approximately 0.75 km section of the SLAC linac tunnel. The existing 120 Hz Cu-linac will also be able to generate X-rays from either the SXU or from the HXU (extending the energy reach to 25 keV). A second upgrade, LCLS-II-HE [38], is presently in the preliminary design phase, and will double the CW-SCRF electron beam energy to 8 GeV and thus increase the spectral reach of the HXU at high repetition rate to more than 12 keV by installing additional cryomodules in the final 250 m of the refurbished SLAC linac tunnel. Anticipated improvements in electron beam emittance will extend the energy reach beyond 18 keV at a high repetition rate.

(ii) SACLA

The SACLA XFEL in Japan began user operation in 2012, following first lasing in 2011, and is based on a compact pulsed-RF C-band Cu-linac (normal conducting) with a maximum energy of 8.5 GeV and a maximum repetition rate of 60 Hz [20,39]. The SACLA XFEL has two independent hard X-ray sources (BL2 and BL3), based on novel variable-gap in-vacuum undulator technology, spanning the energy range from 4 to 20 keV with X-ray pulses of approximately $0.5 \text{ mJ pulse}^{-1}$ and durations between 2 and 10 fs [39]. A unique design of an electron-beam injector system comprised of a thermionic cathode and multi-stage bunch compressors has enabled the generation of short pulse durations with greater than 50 GW peak power under normal operating conditions. In 2016, SACLA began user operation of a dedicated soft X-ray FEL (BL1), with a maximum photon energy of approximately 150 eV, based on a separate 800 MeV Cu-linac [40]. A unique capability of the SACLA-SPring-8 experimental facility is the capability for combining ultrafast X-ray pulses from SACLA with synchrotron pulses from SPring-8 in a common experimental station.

(iii) Pohang accelerator laboratory X-ray free-electron laser

The third hard X-ray FEL in the world, PAL-XFEL in Korea, began user operations in 2017, following first lasing in 2016. Similar to LCLS and SACLA, PAL-XFEL is based on a pulsed-RF Cu-linac (10 GeV) with a maximum repetition rate of 60 Hz [41]. PAL-XFEL has independent hard X-ray and soft X-ray sources (variable-gap planar undulators), operable at 30 Hz each from the common Cu-linac). The hard X-ray FEL spans the energy range from 2.5 to 15 keV with energies up to $1.5 \text{ mJ pulse}^{-1}$ (depending on the photon energy) [41], while the soft X-ray FEL spans the energy range from 0.25 to 1.2 keV with approximately $0.2 \text{ mJ pulse}^{-1}$ [42]. The projected pulse duration from PAL-XFEL is between 5 and 50 fs. A distinguishing feature of PAL-XFEL is the unprecedented temporal stability (for a Cu-linac), with reported rms timing jitter of approximately 20 fs between X-ray pulses and optical pulses from a synchronized laser system [41].
(iv) SwissFEL

The SwissFEL XFEL in Switzerland began pilot experiments in 2018, following first lasing in the X-ray range in 2017 [43]. Similar to the first three hard X-ray FELs described above, SwissFEL is based on a pulsed-RF Cu-linac (5.8 GeV) operating at a repetition rate of 100 Hz, with a goal to generate X-rays at 1 Å with the lowest e-beam energy necessary for operation at that wavelength [41]. SwissFEL consists of two independent X-ray sources: the Athos beamline operates in the 0.25–2 keV range, and the Aramis beamline operates in the 1.8–12.4 keV range. Both beamline use using variable-gap in-vacuum undulators, with Aramis using planar undulators, and Athos using APPLE-II type. The projected pulse duration from SwissFEL is between approximately 2 and 20 fs. Two specialized operating modes are planned: (i) generation of large (incoherent) bandwidth, up to 3.5% FWHM, for EXAFS and related applications and (ii) ultrashort pulses, approximately 60 as [23].

(v) European X-ray free-electron laser

The European XFEL (EuXFEL) facility in Germany began operation in 2017, and is the first hard X-ray XFEL to be based on superconducting accelerator technology (17.5 GeV linac energy) [44–46]. In addition to the high peak X-ray brightness, the EuXFEL is designed to achieve high average brightness by operating in a burst-mode at 10 Hz, with a maximum of 2700 pulses per burst (at an intra-burst repetition rate of 4.5 MHz), for an effective maximum average repetition rate of 27 kHz. The superconducting linac feeds two electron beamlines, with a flexible e-beam distribution system allowing for nearly simultaneous operation. X-rays are generated from three planar variable-gap undulators: SASE1, SASE2 and SASE3, with additional space for more undulators and instruments (SASE4, SASE5). SASE1 and SASE2 span the X-ray energy range from 3 to 25 keV in the fundamental, while SASE3 spans the soft X-ray energy range from 0.26 to 3 keV in the fundamental, augmented by an after-burner for polarization control. EuXFEL exploits several e-beam energy operating points (8.5, 12, 14 and 17.5 GeV) to enable continuous coverage of the X-ray energy ranges indicated above. The X-ray pulse duration ranges from a few femtoseconds at the lowest e-bunch charge up to approximately 100 fs. The projected peak power is several tens of gigawatt at saturation, depending strongly on the electron energy and charge, and X-ray wavelength. With current SASE operation, longitudinal coherence properties of EuXFEL are expected to have a large variation, depending on the electron bunch duration and the operation wavelength. Hard X-ray self-seeding (HXRSS) is presently under development for SASE2 [47]. Two-colour operation with variable delay between pulses is under development for SASE3 [47].

(vi) Soft X-ray free-electron lasers

Two leading XFEL facilities in the soft X-ray range are the FLASH facility at DESY in Hamburg, Germany, and the FERMI facility in Trieste, Italy.

(vii) FLASH

The FLASH facility emerged from the Tesla Test Facility (TTF) and is supported by a 1.25 GeV linac based on TESLA superconducting accelerator technology (later adapted for EuXFEL). FLASH began user operations in 2005 as the world’s first XFEL operating in the EUV range [48]. FLASH is further distinguished by the high repetition rate supported by burst-mode operation at 10 Hz, with a maximum of 800 pulses per burst (for an effective maximum repetition rate of 8 kHz), and an adjustable inter-pulse spacing between 1 and 25 μs. The FLASH accelerator currently serves two SXUs: FLASH1 and FLASH2. FLASH1 is a fixed-gap undulator, and spans the photon energy range from approximately 24 eV to approximately 295 eV (in the fundamental, by tuning the electron energy) with single pulse energies up to 0.5 mJ. FLASH2 is based on a variable-gap undulator and spans the energy range from approximately 14 eV to approximately 310 eV in the fundamental, with single pulse energies up to 1 mJ. For external
seeding development (sFLASH), FLASH incorporates two modulator sections, two magnetic chicanes and a variable-gap radiator section. Direct seeding using an HHG laser source at 38 nm was first demonstrated in 2013 [49]. Current seeding efforts at FLASH focus on high-gain harmonic generation (HGHG) as described below.

(viii) FERMI

The FERMI XFEL is the first user facility to exploit external seeding via HGHG (as described below) as the primary mode of operation [21,50–52], and is further distinguished by the high level of synchronization (approx. 2 fs rms) between the X-ray pulses and the ultrafast external seed laser [53,54], a portion of which is transported to the experimental endstations for pump–probe experiments. FERMI is based on a pulsed-RF Cu-linac (1.5 GeV) operating at a repetition rate between 10 and 50 Hz, and consists of two independent X-ray sources, FEL-1 and FEL-2, both of which are based on adjustable-gap APPLE-II type undulators providing full polarization control [55], and are seeded from an ultrafast laser source continuously tuneable in the VUV. FEL-1 operates in the 14 to 61 eV range, with EUV pulse energies from 80 to 200 µJ. FEL-2 operates in the 60–300 eV range, with EUV pulse energies from 10 to 100 µJ. The external HGHG seeding process results in narrow bandwidth pulses ($\Delta E/E < 10^{-3}$) with typical duration of less than 100 fs (near the Fourier transform limit) [54,56], very stable operation (approx. $10^{-4}$ relative wavelength stability, and approx. 10% rms pulse energy stability) and laser-like statistical radiation properties [57]. A special two-colour operating mode of FEL-1 uses two seed pulses (at different wavelengths) and two sections of the radiator tuned to the harmonics of the two seeds [58]. The high longitudinal coherence of FERMI also allows the generation of phase-stable two-colour pulses at different harmonics [59,60].

(b) X-ray free-electron laser recent advances

The dramatic early success of free-electron-based X-ray lasers has triggered intense development of XFEL facilities around the world. At the same time, it is widely recognized that such sources still have significant potential for improvement. Close interaction between the X-ray scientific community and FEL physicists has driven the development of new modes of XFEL operation and improved capabilities. Major directions for improvements include: ultrashort X-ray pulse generation and characterization, control over the temporal coherence (near the Fourier transform limit) via seeding and related schemes, and versatile operating modes such as two-colour operation, multi-pulse sequences, and large coherent bandwidth.

(i) Ultrashort X-ray pulse generation

Among the most compelling scientific capabilities of XFELs are ultrashort X-ray pulses which open entirely new domains of ultrafast science. Owing to the exceptional electron beam quality and feedback control, the electron bunch length can easily be varied during FEL operation. To accommodate user requests, LCLS has developed two operating modes to deliver pulses with durations in the few femtosecond range: (i) a low-charge operating mode [28] and (ii) a slotted-foil method [61]. Both ultrashort pulse modes are delivered in routine operations. The operation of LCLS in the low charge mode (20 pC) substantially improves the transverse emittance of the e-beam from the injector and further mitigates collective effects in the accelerator, allowing for extreme e-bunch compression. The compressed e-bunch length is estimated to be less than 5 fs FWHM. Stable saturated XFEL operation with estimated X-ray peak power levels comparable to the nominal charge (150–250 pC) are routinely achieved over the entire LCLS wavelength range, while the total energy per pulse is lower by approximately 10-fold, approximately in proportion to the pulse duration. In this low charge mode, the FEL pulse consists of only one or two coherent spikes of radiation. The second method for ultrashort pulse generation exploits a slotted foil to spoil the e-beam emittance in selected regions (effectively clipping the leading and trailing edges of the energy-dispersed electron pulse) leaving a very short time slice of low-emittance
electrons to produce femtosecond X-rays [61]. With this approach, the variable pulse duration and separation is achieved with an array of slotted foils in different arrangements. A single slot with tapered width can control the X-ray pulse duration from 50 to 6 fs, while V-shaped double slots with different separation can generate two short X-ray pulses with variable delay between approximately 10 and 80 fs for X-ray pump/probe studies [62].

The generation of subfemtosecond (attosecond) pulses from XFELs represents an important area of development with significant potential impact in many areas of science. LCLS has recently generated the first attosecond pulses from an XFEL. Single pulses of approximately 200 as duration have been generated in the hard X-ray regime (5.6 and 9 keV) using nonlinear e-bunch compression techniques (figure 1) [10]. An attosecond X-ray pulse energy of approximately 10 µJ is achieved (from a 20 pC e-bunch) enabling new opportunities in attosecond and nonlinear X-ray science. Ongoing R&D on laser manipulation of the electron beam [63] (LCLS XLEAP project), along with new precision attosecond diagnostics [64], is expected to significantly expand on the attosecond capabilities of LCLS and other XFELs.

(ii) Control over temporal coherence

While the process SASE has been an effective starting point for XFEL operation (and remains the most common XFEL operating mode), the fluctuating spectrum and coherence properties pose limitations for many areas of science. To overcome this, several effective approaches for controlling the temporal coherence (i.e. spectral amplitude and phase) of XFELs have been developed, and R&D efforts are ongoing to develop more effective control extending over the entire operating energy range of XFELs.

Self-seeding

The concept of self-seeding (i.e. using X-rays from the first half of an XFEL undulator array, with spectral filtering, to seed the second half, as shown in figure 2) is a straightforward extension to SASE XFEL operation, and was originally proposed at DESY for soft X-rays [65] and later for hard X-rays [66]. HXRSS has been demonstrated at LCLS using an in-line thin diamond crystal, which transmits most of the SASE X-ray pulse but also generates a time-delayed monochromatic seed pulse. The chicane provides a delay of the e-bunch to temporally overlap with, and amplify, the seed pulse in the second part of the undulator array. Self-seeding at the approximately 1 Å scale, with approximately 40-fold bandwidth reduction (approx. 0.5 eV bandwidth), was demonstrated at LCLS as illustrated in figure 2 [31]. Currently, HXRSS at LCLS provides X-ray from 5.5 to 9.5 keV with up to a four-fold increase in photons/pulse compared to the combination of SASE and a post-XFEL monochromator.
Figure 2. (a) Schematic of XFEL self-seeding concept, from [30]. The seeding monochromator, consisting of a grating (G), three mirrors (M1; M2; M3) and an adjustable slit, selects a small bandwidth (green line), while the electron chicane directs the electron bunch around the monochromator and resets the electron beam to shot noise. Finally, the overlap diagnostics (BOD10 and BOD13) co-align electrons and monochromatic X-rays in the second half of the FEL. Soft X-ray self-seeding spectrum (b, from [30]) and HXRSS spectrum (c, from [31]) results from LCLS.

More recently, a compact SXRSS system, based on a grating monochromator, has been implemented at LCLS covering the photon energy range from 0.5 to 1 keV. To date, a resolving power of 2000–5000 and wavelength stability of $10^{-4}$ have been demonstrated, with an increase in peak brightness by a factor of two to five across the photon energy range. By eliminating the need for a post-XFEL monochromator (and associated optics losses), SXRSS can deliver as much as 50-fold higher brightness to users [30]. Further SXRSS development is underway to reduce the spectral ‘pedestal’ and improve the achievable resolution.

External seeding

External XFEL seeding with a conventional laser seed source has been pioneered at the FERMI XFEL in Trieste, employing the HGHG scheme [67–69]. In HGHG, a laser at long wavelength (e.g. visible or UV) modulates the energy of an e-beam in a resonantly tuned undulator. A subsequent dispersive magnetic chicane translates this modulation into longitudinal charge-density modulation (electron bunching), at the seed laser wavelength, and at higher harmonics. These harmonic modulations serve as the input to a subsequent XFEL radiator tuned to be in resonance with the harmonic. FERMI FEL-1 employs a single-stage HGHG scheme spanning a harmonic range from $h$ approximately 3–13 of the initial seed photon energy (approx. 4.7 eV). FERMI FEL-2 employs a two-stage HGHG scheme with combined harmonic range from $h \sim 12$ to 65 of the initial seed photon energy (approx. 4.7 eV) [21,50,51]. Owing to the sensitivity of HGHG to the stochastic energy spread in the e-beam, harmonic conversion is typically limited to $h < 15$ in a single stage. Cascaded HGHG is limited by stochastic phase errors induced by the electron beam phase-space modulation due to the microbunching instability occurring in the linac. R&D is now underway at FERMI, LCLS and other XFEL facilities to explore possible routes to extend external seeding schemes towards 1 nm. In addition to the control of the XFEL bandwidth (and longitudinal first-order and higher-order coherence), external seeding provides...
Figure 3. Fresh-slice multi-pulse scheme. The wakefield in the dechirper creates a strong transverse head–tail kick in the e-bunch, and the oscillating orbit (combined with fixed-magnet corrections) insures that the tail of the bunch (orange) lases in the first undulator section (at energy $E_1$) and the head of the bunch (blue) lases in the second undulator section (at energy $E_2$). The current LCLS layout allows for up to three pulses with controlled photon energies and pulse delays, from [71].

for exquisite inherent optical/X-ray synchronization [53] and multi-pulse multi-colour operation with close locking between pulses as described above [58,59].

(iii) Two-colour X-ray free-electron laser pulses

The simultaneous generation of two XFEL pulses, with independently tunable energies, and adjustable relative arrival time, promises to open important new areas of ultrafast X-ray science. The two-colour XFEL methods developed to date are categorized by classes. One class relies on generating two X-ray colours by splitting the XFEL undulator array into two sections, each tuned to distinct K values (see equation (2.1)) with a quasi mono-energetic electron beam. Electrons in each undulator section generate an XFEL pulse at a distinct photon energy (with an energy separation limited only by the tuning range of the undulators), and a chicane between the sections can be used to control the relative time delay (typically up to approx. 1 ps). This approach has been demonstrated at LCLS [33] and at SACLA [70]. Recently, a novel fresh-slice technique for multi-colour XFEL pulse production has been demonstrated at LCLS. In this approach, different temporal slices of an e-bunch lase at different photon energies [71]. As shown in figure 3, a passive wakefield device (dechirper) controls the trajectory of the e-bunch to provide multi-colour pulses. The fresh-slice scheme enhances the achievable peak X-ray power, and has been extended generate variably polarized two-colour pulses and three-colour pulses.

The second approach for two-colour generation relies on using two e-bunches at different energies to generate two XFEL pulses of different colours simultaneously in one undulator array. This method has been demonstrated at LCLS [36], and requires generation, acceleration and compression of double e-bunches within a single RF cycle of the accelerator structure. One advantage of this approach is that each X-ray pulse can reach the full saturation power, improving the two-colour intensity by approximately 10-fold in comparison with the split-undulator approach described above. This ‘twin-bunch’ approach can also be combined with HXRSS using appropriate crystal orientations to generated two seeded hard X-ray colours [35]. The time delay between the two pulses can be adjusted using an energy-dispersive chicane from close to zero to about approximately 130 fs, and the achievable energy separation is typically approximately 3% (1%) in the soft (hard) X-ray range [36].

3. Laser-based high-order harmonic generation

Laser-driven coherent X-ray sources represent an important laboratory-based alternative to large-scale XFEL facilities for ultrafast pulses in the XUV to soft X-ray range. The HHG process coherently up-converts long-wavelength pulses from an ultrafast laser to much shorter wavelengths, extending to the soft X-ray regime. Thus, table-top-scale X-ray lasers can be implemented, in the same sense that many green, blue and ultraviolet lasers are often based on infrared lasers that are up-converted in frequency, while retaining the coherence of the fundamental light. The ability to up-convert from the near-IR to the EUV and beyond, in a
Figure 4. Examples of high-order harmonic spectra emitted by low-pressure argon gas. (a) The first example of HHG driven with a few-cycle (25 fs) laser pulse, showing a cut-off harmonic order of 61 (13 nm), from [72]. (b) Spectrum when the flux is fully optimized, illustrating much brighter emission but with a lower energy phase-matching cut-off. The dashed line corresponds to the transmissivity of the argon gas, illustrating that the spectrum coincides with a window in the EUV absorption, allowing for higher pressure (and thus higher flux), from [73].

single step, not only allows for compact implementation of many analytical, microscopy and metrology methods first developed at synchrotrons, but also enables new studies of ultras fast dynamics at extreme timescales. The combination of spatial coherence, temporal coherence and very broad spectral bandwidths generated in the HHG process has spawned new areas of research. Furthermore, the coherent nature of the HHG process enables unprecedented control of the generated XUV light, through control of the radiation field (wavelength, chirp, polarization, orbital angular momentum) of the laser that drives the process. The HHG process is the first complex attosecond dynamic process as-yet investigated, and remains an active area of fundamental research in intense laser–matter interactions. As a light source, HHG enables investigations of the very fastest phenomena in the natural world, with important applications in atomic, molecular, chemical and materials science.

(a) Basic physics and history of high-order harmonic generation

In HHG, an intense laser pulse is focused into a gas, where the strong-field ionization process up-converts a (small) fraction of the incident light to much shorter wavelengths. High-harmonic emission in the EUV (illustrated in figure 4) was first observed in 1987 [2], preceded by related observations of high harmonics driven by a CO₂ laser [74]. Subsequent work clearly demonstrated the non-perturbative nature of the laser–atom interaction, distinguished from conventional nonlinear optics in that a large number of nonlinear orders behave similarly (forming a ‘plateau’ in the emission), followed by a relatively abrupt cut-off photon energy [3]. The basic mechanism for HHG was first understood through quantum numeric simulations [75], in a process of discovery that identified the need to use a simulation grid much larger than the extent of the atomic wave function, and associated the cut-off energy with the ionized electron
trajectory according to an approximate relation of \( h\nu_{\text{cut-off}} \sim 3U_p \), where \( U_p = e^2 I / 2c\varepsilon_0 m\omega^2 \) is the ponderomotive energy of a free electron (of mass \( m \)) in a laser field in vacuum (\( \varepsilon_0 \) is the vacuum permittivity) [76]. These simulations established the emission as resulting from the coherent interaction of the ionized electron with its parent ion—a ‘re-collision’ process. The cut-off relation was soon after elegantly derived using a classical trajectory model that refined the cut-off relationship to \( h\nu_{\text{cut-off}} = I_p + 3.17U_p \) [77], and followed by a full semiclassical theory in the strong-field approximation, which has been broadly applied and expanded-upon [78]. The HHG process is essentially a coherent version of the Roentgen X-ray tube, generating high-energy photons as a result of a high-energy electron-atom collision, and relying on the fact that during the ionization process, the electron–ion ensemble remains an isolated, coherent quantum system. The strong electric field of a laser ionizes an electron, pulls it away from its parent ion and then slams it back. The ionizing electron wavefunction reacts deterministically to the applied field, so that the high-frequency dipole emission; i.e. the high-harmonic emission—is also deterministic, with the emission from a large number of phased emitters resulting in a coherent beam. As such, the HHG process is an excellent example of the utility of relatively complex dynamics in a quantum-entangled system—though in this case, the electron–ion separation is small—larger than an atom, but still microscopic.

Furthermore, the HHG process is the first example of non-trivial attosecond time-scale dynamics. Although certainly many elementary-particles with lifetimes in the attosecond or even shorter timescales (the Higgs boson has a lifetime of approx. \( 10^{-22} \) s) are known, these lifetimes are inferred from the linewidths of a simple lifetime-broadened resonances. By contrast, the HHG spectrum is typically complex, reflecting complex dynamics. The harmonic spectral peaks are mutually coherent, so that spectral bandwidth of the entire HHG spectrum is reflective of dynamics on a characteristic timescale that relates to the time-energy uncertainty principle: \( \Delta E\Delta t \sim \hbar \rightarrow \Delta E_{\text{FWHM}}(\text{eV})\Delta t_{\text{FWHM}}(\text{fs}) \sim 1.8 \). With an energy bandwidth of 10s to 100s of electronvolt, HHG dynamics are subfemtosecond.

The most pivotal experimental enabler for both experimentally corroborating the physics of HHG and for practical application has been the development of lasers capable of generating intense few-cycle pulses. Ti:sapphire was developed as a laser medium in the 1980s [79], and has an extraordinarily large spectral gain bandwidth of nearly 0.5 eV, which means it is capable of supporting near single-cycle pulses in the near-IR [80]. This capability opens a new regime of light–matter interactions. Applied to HHG, it allows for attosecond dynamics to be clearly observed and verified in experiment, and further enables the control of interactions on attosecond timescales. Figure 5 shows examples of this: the HHG emission is a result of a periodic re-collision process as the laser field oscillates. With a few-cycle pulse, the relative phase and timing resulting from the re-collision process varies from cycle-to-cycle, as the electron trajectory is being driven by a fast-changing field that also affects the phase of the emitted harmonics. Small changes in the ‘chirp’ of the drive pulse result in changes in the emission spectrum through both intensity and phase components, in a way that can be understood in terms of the re-collision model [81]. Further, control of the pulse shape through a learning algorithm can be used to control the HHG spectrum [82]. These experiments are the first examples of attosecond science, where strong-field control of matter on a cycle-by-cycle basis is controlling sub-optical-cycle electron dynamics, revealed through changes in the HHG spectrum.

The development of lasers for generating intense few-cycle pulses is based on a series of rapid interrelated advances: mode-locked Ti:sapphire oscillators [83], chirped-pulse amplification [84] and the ability to control the group-velocity dispersion, with high precision, across these complex systems [80,85–87]. The first HHG experiments used picosecond [3] or large-aperture femtosecond lasers, where plasma and ionization depletion effects complicated the interaction, making quantitative comparisons with theory inaccurate [88]. By contrast, the application of intense few-cycle pulses for HHG not only generated dramatically higher harmonic orders and higher conversion efficiency, this also immediately resulted in quantitative verification of the \( I_p + 3.17U_p \) cut-off relationship [89,90].
Figure 5. The attosecond dynamics of HHG was first clearly revealed through the effect of the cycle-by-cycle optical phase of a few-cycle driving laser pulse on the spectrum of high-harmonic generation. (a) The effect of simple chirp on the spectrum. The fact that the harmonic spectrum sharpens into narrow spectral peaks with a positive chirp on the driver is the result of the re-collision spectral phase, from [81]. (b) Through control of the driving pulse shape, the intensity of a single harmonic order can be enhanced, demonstrating a useful control over attosecond dynamics that directly confirms that the full spectrum of HHG is mutually coherent. This optimization is effected by applying a nonlinear chirp to the pulse that compensates for chirp of the HHG emission resulting from the re-collision process, from [82].

(b) Phase-matching

Importantly, the use of few-cycle pulses for HHG, combined with novel waveguide methods to control phase-matching in the HHG process, also revealed how HHG is phase-matched. The understanding of HHG as a discipline within nonlinear optics, including both propagation/phase-matching and the (single-atom) nonlinear susceptibility, has proven to be a much tougher problem than the basic mechanism of emission, and has progressed systematically over the past 2 decades. In ‘conventional’ nonlinear optical processes, the interaction can be described as a simple, reversible instantaneous response, with phase-matching treated in the frequency domain, through engineering of the index of refraction. In contrast for HHG, generation and phase-matching are dynamic, non-instantaneous and interlinked between the atomic dynamics and the phase-matching. The interlinked nature was first discussed in work showing that the phase of emission of the HHG light with respect to the phase of the driving laser should vary in such a way as to affect the divergence characteristics of the HHG emission [91]. However, the key realization for HHG as nonlinear optics is that in the few-cycle HHG regime, the process is intrinsically dynamic, and can in-general be phase-matched. The phase-matching occurs transiently; i.e. at some time during the ionization/HHG generation process when the medium is only partially ionized [73,92]. This represents a conceptual shift with profound implications compared with conventional NLO—because of these dynamics, the physics inherently pushes the HHG process to emit shorter pulses than the driving laser. In conventional NLO processes such as second harmonic generation, phase-matching narrows the spectral bandwidth of the emission and therefore generates a longer harmonic pulse. In HHG, phase-matching is relatively insensitive to the wavelength of the unconverted light, while the temporal gating is strong. With the demonstrated generation of coherent bandwidths approaching 1 keV using mid-IR laser-driven HHG as discussed below, we now have a feasible path for generating pulses of just a few attoseconds duration [93].

HHG phase-matching is most straightforwardly observed in a weakly ionized gas [92]. Before interacting with the laser pulse, the gas medium is neutral, with an index of refraction for the driving laser that is incrementally higher than 1; i.e. $n_0 = 1 + \delta(\omega)$. However, the index of
refraction for the high-harmonic light generated (at harmonic number \(q\)) is incrementally less than 1; i.e. \(n_q\omega = 1 - \delta_q\omega\). This is the result of the fact that all the atomic resonance lines that contribute to the index of refraction are at frequencies higher than the driving laser, and lower than the generated harmonics. **However**, when the strong laser field starts to ionize the gas, previously bound electrons become free, and provide a negative contribution to the index of refraction of the driving laser. At a critical ionization fraction, the indices of refraction of the driving laser and the harmonic light match [92,94]. For phase-matching in a plane-wave approximation (when most of the coherent emission is observed), the intrinsic intensity-dependent phase of HHG is not changing, and is thus not a factor. This critical ionization fraction is independent of gas pressure. Furthermore, if a pulse is not short enough, the gas may be ionized past the critical ionization fraction before any high harmonics in the desired wavelength range are emitted. Thus, the HHG process is extremely sensitive to the risetime of the driving laser pulse, but in a way that does not scale simply—conversion efficiency to a particular wavelength increases rapidly with shorter pulse duration, up to the point where this target wavelength is generated before ionization reaches the critical value. Even shorter pulses serve to increase the cut-off, without dramatically increasing the efficiency to any particular wavelength (and with a spectral broadening of the individual harmonic peaks). Phase-matching of HHG was first revealed using a few-cycle pulse, with the laser and the gas confined in a waveguide [92]. The waveguide adds a pressure-independent factor to the phase-matching expression, allowing for gas pressure-dependent control of the phase-matching conditions, resulting in a sharp and clear pressure optimization that was first to reveal the existence and mechanism behind phase-matching.

One important realization is that in phase-matching of HHG, the phase velocity of the driving laser generally sets the phase-matching conditions—the relative dispersion of the harmonics is small as they all see an index very near 1. As a result, phase-matching does not particularly restrict the bandwidth of the generated light as is the case with visible-wavelength NLO. This is a profound realization: in visible NLO, phase-matching parameters depend on frequency—not time—and the result is a narrowing of spectral bandwidth. In HHG, the phase-matching parameters are time-dependent—and this restricts emission in the time domain, making very short pulses possible, into the attosecond domain.

In its simplest manifestation, implemented as an efficient light source, the conditions for optimum conversion efficiency are clear:

1. The peak intensity of the focused beam should drive the ionization to the critical phase-matched ionization level at the peak of the pulse [73,92]. The actual optimum peak intensity depends primarily on the HHG gas medium, and on the pulse shape and duration. Overdriving the medium generally causes a degradation in the efficiency and beam quality (though the flux may still increase due to larger mode area), without increasing the observed cut-off substantially.
2. The focus region must sustain this intensity over a length corresponding to several XUV/soft X-ray absorption lengths of the generated light [95,96].
3. The region of dense gas must end as abruptly as possible before the laser intensity decreases significantly (which would shut-off the HHG but still allow the generated light to be re-absorbed). This also sets a lower limit on the confocal parameter of the driving laser, which ideally is several times longer than the gas interaction length to minimize intensity-phase coupling.

These conditions set a narrow optimum operating range that (depending on the gas and laser wavelength) results in conversion efficiencies approaching \(10^{-3}\) for UV laser up-conversion into the VUV, decreasing slowly to approximately less than \(10^{-7}\) for soft X-ray generation using mid-IR lasers. Phase-matched HHG will occur in general with a sufficiently short driving laser pulse, regardless of geometry. However, optimum conversion efficiency becomes progressively more difficult with lower driving pulse energy: a tighter focus is required, with a higher gas density. Confining the gas within a waveguide can serve to extend the interaction length, reducing
the pressure requirement while also confining the gas and making it easier to obtain higher target pressures. The relative benefit of using a waveguide varies. Generally, for VUV/XUV high harmonics using gasses such as xenon and krypton, the absorption cross-section at the target wavelength is quite high, and the ionization intensity is low, making HHG in a gas cell quite usable for HHG. However, for HHG in argon gas, a Cooper minimum in the absorption around 50 eV/30 nm makes it advantageous to use a waveguide for conversion, with a relatively sharp pressure-tuned phase-matching peak that reveals the existence of phase-matching [92].

HHG in a properly mode-matched waveguide can serve not only to optimize conversion efficiency, but also to minimize gas usage because the aperture and conductance of the gas flow is coincident with the propagation aperture of the laser. Phase-matching over an extended propagation length also optimizes the spatial coherence of the output [97]. The reason for this is that on-axis phase-matching is heavily favoured, and the waveguide maintains a relatively constant peak intensity over an extended propagation length. These constant phase-matching conditions also are favourable to the generation of isolated attosecond pulses, relaxing the requirements for very short driving pulse duration by taking advantage of the phase-matching [98]. Finally, regardless of conditions, the dramatically lower gas flow using a waveguide geometry obviates any need for gas recycling, which otherwise has proven an economic necessity in most cases other than HHG using (inexpensive) argon.

Late in the 2000s, the ability to generate tuneable ultrashort mid-IR pulses at high enough power for HHG led to the further realization that it is possible to phase-match the process into the soft X-ray region of the spectrum to generate useful flux. The basic high-harmonic cut-off relationship makes this clear: $h\nu_{\text{cut-off}} = I_p + 3.17 U_p \sim \propto I^2$, which reflects the fact that a longer trajectory returns the electron to the parent ion with higher energy. However, although the cut-off scaling is very favourable, the first efforts at HHG with mid-IR lasers also made clear the trade-off—that the re-collision interaction quickly diminishes [99]. Thus, under constant pressure conditions, the conversion efficiency was observed to decrease dramatically with increasing laser wavelength [100]. However, this scaling improves dramatically when phase-matching of the process is considered (figure 6). In general, the index of refraction of atomic gasses varies only slowly with laser wavelength in the infrared, and in the few-cycle tunnelling regime of ionization, the ionization rate is nearly independent of the peak intensity even while the high-harmonic cut-off energy is scaling quickly towards shorter wavelengths. Finally, for proper choice of gas medium, the soft X-ray absorption of the HHG light is dropping very quickly. Thus, the optimum density-length product increases quickly, with the need for multi-atmosphere pressures. Pressure optimization rapidly becomes problematic in all except the hollow waveguide geometry, which is why initial studies were discouraging [102–104].

When all these factors are considered, we can add two more conclusions for optimization of HHG:

1. Except in limited cases (such as for argon gas generating light in the Cooper minimum region), helium gas appears to be the most efficient HHG nonlinear medium. Inner-shell absorption in other species represents an incoherent loss mechanism, while (with no inner-shell absorption) helium becomes transparent into the X-ray spectral region. Atomic hydrogen would likely work even better were it practical to use at extremely high pressures. Molecular hydrogen and non-spherical molecules in general show lower efficiency due to the phase variations in emission resulting from orientation effects. However, it is possible that techniques such as molecular alignment can be exploited to improve efficiency in molecular systems [105].

2. The optimum driving laser wavelength corresponds to that required generate the desired HHG wavelength at the time during the pulse when helium gas is ionized to the critical ionization.

This conclusion is contrary to the ‘conventional wisdom’ prior to approximately 2008. The optimization of HHG in helium gas generally requires much higher pressures than were possible
in simple gas-jet or cell geometries, explaining this misperception. Figure 7 shows the estimated relative conversion efficiency assuming that the conversion efficiency is limited by the signal absorption [103]. However, it is important to understand that these simple scaling rules do not consider the effects of (i) diffraction of the laser beam; (ii) $v \times B$ forces that can cause the re-colliding electron to ‘miss’ the parent ion on return; (iii) group-velocity walk-off between the driving laser and the generated high harmonics; (iv) ionization loss of the driving laser; and (v) de-coherence of the re-colliding electron on ever-longer trajectories that can pass many neutral atoms in a high pressure medium. All these factors become more severe as one scales HHG towards shorter wavelengths; however, none of them appear fundamental. For example, (i) is ameliorated with waveguide propagation (the use of advanced low-loss waveguides [106] for HHG into the X-ray region of the spectrum may well be transformational); (ii) can be corrected with a longitudinal electric field, as is introduced in propagation [107], (iii) and (iv) could be ameliorated by tapering a waveguide, and (v) by operating with a large guided mode and a long propagation length. Thus, the ultimate wavelength limit for generation of coherent X-rays through HHG is not known, and may extend into the hard X-ray spectral region—for example, for advanced low-dose medical imaging and X-ray backscatter imaging.
Nevertheless, implementation is not trivial, as the optimum driving wavelength shifts into the THz spectral region, while the required peak intensity for ionization remains comparable, and focusing and diffraction will require higher pulse energies. There is a significant need for dramatically improved modelling of HHG in order to predictively optimize the process through engineering. The continued advancement is HHG sources towards useful hard X-rays is likely to be a decades-long endeavour involving modelling coupled with advances in laser technology.

(c) Other regimes of bright high-harmonic generation

By virtue of its simplicity, nearly all contemporary work is done using phase-matched HHG in the weakly ionized regime described above, unless observing the spectrum is the end goal in which case optimum efficiency is not critical. Nevertheless, there are a number of other regimes that have been explored as potentially useful. One is quasi-phase-matching. One concept identified very early in the field [108] is that the re-collision-induced intrinsic phase is susceptible to manipulation with a relatively weak laser field that need not be collinear with the driving laser; for example, a counter-propagating beam. This can allow for quasi-phase-matching. For example, periodic modulation of the diameter of the waveguide can scatter light in such a way as to enhance emission in an otherwise non-phase-matched regime [109,110]. While this technique has proven to be too alignment-sensitive to be practical, a very convincing demonstration of QPM of HHG used a counter-propagating pulse sequence in a waveguide for selective enhancement (by much greater than 100×) of harmonic orders with photon energy above the conventional phase-matching cut-off [111]. Importantly, although manipulation of the HHG phase requires an electromagnetic ‘structure’ created by a relatively long duration pulse, the peak intensity for creating this structure is orders of magnitude lower than that required for the driving pulse, as it need only induce a phase shift significant in relation to the desired HHG wavelength. These counter-propagating pulse experiments also served to probe the coherent build-up of the high-harmonic light. Other work has, for example, used multiple sources for QPM [112], and used weak fields to induce diffraction of the HHG beam [113].
Another regime of HHG recently revisited is the use of short-wavelength ultraviolet driving lasers [101]. In the conventional phase-matching regime, a UV driving laser would be expected to have a cut-off photon energy in the VUV at best, by virtue of the small ponderomotive potential from a high-frequency driving laser. However, the simplest picture of phase-matching is an increasingly poor approximation for shorter wavelength driving. From the perspective of the driving laser, the index of refraction of the neutral gas is higher, the negative contribution due to the field-ionized plasma is smaller and the contribution of the ions to the total phase-matching index of refraction is greater. All these factors diminish, and likely in some regimes cancel the phase mismatch that would otherwise be predicted. Along with the relatively high per-atom emission driven with UV light, high-harmonic light up to the water window region of the spectrum has been observed using UV drivers [101].

Finally, we note that at these high intensities, nonlinear reshaping of the driving laser pulse, in both the temporal and the spatial domain, can be important (e.g. [93])—but this is an area where theory as well as direct corroborating measurements becomes difficult and where much work remains to be done.

(d) Attosecond pulse generation

The spectrum of HHG is temporally coherent as a whole, and in the time domain consists of a series of attosecond bursts resulting from periodic re-collision events that follow the subcycle ionization of the atoms [114]. The first direct measurement of this was accomplished using the RABBITT technique [115], which is the interferometric case of the laser-assisted gas-phase photoionization process [116,117] instrumental to a large number of attosecond-science experiments. ‘Isolated’ attosecond pulses (IAPs) can be generated using HHG in the case where the emission is confined to a single electric field peak of the laser field. IAP using HHG was first suggested through polarization modulation of the driving laser field [118], although the scheme first demonstrated to generate an isolated attosecond pulse [119] followed the simpler approach of using a pulse short enough that the highest energy photons are emitted only during a single re-collision even [120]. These first experiments used pulses compressed to less than 10 fs using high-energy compression in a hollow waveguide [121]. Since that time, more-sophisticated polarization modulation techniques have made it possible to use pulses at 800 nm with the duration of approximately 25 fs [122]. It was also shown that the use of a properly phase-matched geometry also relaxes the pulse duration requirement for IAP generation [98].

However, the generation of isolated attosecond pulses remained quite difficult for quite some time, due to the need for both carrier envelope stabilization of high-energy pulses from the ultrafast Ti:sapphire amplifier (which requires interferometric stabilization in a high-power laser environment), and the need for pulse compression of the high-energy output of the amplifier system, which limits the pulse energy available and thus the HHG efficiency [121,123]. Overall, between the very significantly lower flux, and the additional technical complexity, many successful attosecond dynamics experiments have more simply made use of the attosecond pulse trains (APTs) that are inherent in the HHG process. As a generalization, the ability to generate a single attosecond pulse, as opposed to a train, is most useful in characterizing few-femtosecond phenomena. In the case of truly attosecond dynamics, the APT increases the data acquisition rate without introducing ambiguities, and in fact allows in some cases for a combination of spectral and temporal resolution that is not possible with IAPs [124].

More recently, as interest has shifted to HHG in the soft X-ray spectral region driven by mid-infrared lasers, the generation of isolated attosecond pulses may well be more the default for HHG rather than an extra effort. In driving HHG with longer wavelengths, the interval between optical cycles becomes progressively longer, while the effect of the plasma generation on the phase-matching conditions increases. Furthermore, the absorption depth of the generated light increases, making phase-matching more selective. The result is that the time window of phase-matched HHG shortens, resulting in stronger confinement of the emission to a single burst.
Figure 8. High-harmonic generation pulse duration as a function of driving laser wavelength. As the driving laser is tuned to longer wavelengths, spectral peaks merge and the spectrum becomes a continuum. The fact that this is a true continuum can be verified using Fourier transform spectroscopy. In the case where the HHG emission is presumed to be coherent, this verifies that the emission is in the form of a single attosecond pulse, from [126].

In experiment, the HHG spectral peaks are observed to rapidly become broader, while their spacing shrinks, rapidly merging into an apparent spectral supercontinuum [93,125]. Virtually, all HHG with driving laser wavelengths greater than approximately 1.5 µm is observed to be the supercontinuum characteristic of emission during a single attosecond pulse. However, verifying that the spectrum is indeed a true continuum requires spectral resolution beyond that generally obtained with a conventional diffraction grating spectrometer. High-resolution Fourier transform spectroscopy can more readily and directly verify a spectrum with the characteristics of a single pulse (figure 8) [126].

Regardless of the number of bursts in a high-harmonic pulse, each burst itself also exhibits a spectral chirp—the highest photon energies come first within a burst [91]. For HHG with infrared drivers, this chirp increases, so that the total duration of the single emitted burst is longer, but, in principle, can be recompressed to very short duration using dispersive optics.

(e) Control of spectrum and polarization

Another recent topic of considerable interest is in control over the polarization characteristics of high-harmonic light. For example, generation of HHG light with circular polarization was considered impossible for quite some time—when a circularly polarized driving laser is used to drive the HHG process, the ionized electron spirals away from its parent ion and no re-collision occur. However, by applying two fields with two colours of opposite polarization, re-collision trajectories in the two-dimensional plane can result in spectra with complex polarization states [127]. (The same configuration was explored in fact 20 years earlier [128], but the experimental work emphasized the generation, rather than the polarization state, of the emission.) More recently, a number of different methods for generating APTs, as well as isolated attosecond pulses, of circular polarization have been demonstrated [129–132]. Using tomographic reconstruction in polarization-dependent surface photoemission, the polarization state of these APTs can be reconstructed and compared with theory (figure 9).
Figure 9. (a) Circularly polarized HHG spectrum, from [129]. (b) Three-dimensional field plot of the experimentally reconstructed pulse train of circular HHG, compared with a numerical simulation, from [133].

(f) High-order harmonic generation applications: new science with a new light source

Current-generation table-top EUV light sources generate a coherent flux in the near-EUV comparable to a synchrotron source—but with much higher peak brightness and intrinsic synchronization to a laser. Further development of high-power laser source in the less than 100 fs regime, when suitable refined, will can continue to advance the HHG source flux [134]. However, current capabilities of HHG sources are more than adequate for experimental applications that are driving new scientific understanding in a variety of areas. The breadth of applicability ranges from basic discoveries—such as revealing hidden phases in quantum materials and the nature of excited-state electron dynamics—to the practical and technologically relevant, such as exploring fundamentally new regimes of heat transport, and for mechanical and thermal properties relevant to nanotechnology [135,136]. Importantly, applications frequently make use of the spatial coherence of HHG light (as well as its temporal coherence in the form of short pulses), for example, for unprecedented subwavelength resolution imaging in the EUV [137].

One important science application area for HHG sources has been in the study of fundamental charge dynamics in solids using time-resolved photoelectron spectroscopy (building on early studies by Haight et al. [138]), and combined with the laser-assisted photoelectric effect [139,140] to achieve extreme time resolution. Early attosecond studies applied isolated pulses to investigate the dynamics of photoelectron emission from solids [141], and the propagation time of electrons through surface layers [142]. Important recent studies investigated ‘final state’ effects in the photoemission process, and reported the first observation of approximately 200 as dynamics of highly excited states in solids [124]. This approach has also been used to directly examine
scattering versus screening interactions in the photoemission process, illustrating how this parameter is related to the band structure of the surface [143]. These measurements serve to validate solid-state models in new, non-equilibrium regimes.

Another important science application area for HHG sources has been in the study of spin and magnetization dynamics in magnetic materials using time-resolved absorption spectroscopy. By using the magneto-optic Kerr effect (MOKE) at the M-edge of many magnetic materials in the approximately 40–70 eV energy range, one can monitor element-specific magnetization dynamics. One surprising result from studies in common permalloy magnets found that the magnetization of the iron and nickel components of the material are coupled, but with a delay: following excitation, the magnetic order in iron is diminished first, and the nickel magnetization decays with the same timescale but with a delay. Studies of alloys of varying composition indicate that this delay, as short as 10 fs, relates directly to the exchange energy of the material [144]. More recent work shows evidence that the specific heat clamps the electron temperature immediately on heating of magnetic materials past their Curie temperature, on timescales much faster than the approximately 200 fs dynamics apparent from visible and EUV MOKE experiments [145], and illustrate the role of spin currents in the ultrafast demagnetization process [146].

HHG sources have also been used to probe the acoustic, mechanical and thermal properties of nanostructured materials, frequently with unexpected results. By diffracting coherent ultrafast EUV pulses from a periodic nanostructure, thermal and acoustic dynamics can be monitored via changes in diffraction resulting from thermal expansion. EUV pulses make it possible to probe dynamics at small length scales where (i) thermal conductivity becomes non-diffusive because the phonon mean-free path in many materials is in the 10s to 100s of nanometre range; and (ii) where over-layers just a few nanometre thick influence acoustic wave propagation. In the former case, the most remarkable counterintuitive result to date is that heat conduction from a periodic array of nanostructures can be faster for a closely spaced array than from a structure where the heat sources are more widely separated [147]. In the latter case, by monitoring both longitudinal and surface acoustic wave propagation, both Young’s modulus and the Poisson ratio of very thin films can be determined for the first time, verifying that indeed the connectedness, at the molecular level, of very thin films profoundly affects acoustic properties [148].

(g) High-order harmonic generation outlook

Progress in the development and scientific application of HHG laser sources over the past 2 decades has been extraordinary. Further advances continue in several areas. First and likely foremost, the use of mid-IR (1.5–4 μm) ultrafast lasers makes possible the extension of HHG sources to the soft X-ray region, even to greater than 1 keV. It is possible that intense light pulses at greater than 10 μm wavelength, and in the THz region of the spectrum, will generate coherent hard X-rays [107]. Second, new regimes of HHG that take advantage of new propagation physics (e.g. using deep-UV lasers for HHG in more highly ionized gasses [101]) may reveal novel more-efficient generation methods. Third, the generation of new polarization states represents new physics as well as an enabler for many applications [127,133]. Fourth, as ultrafast laser technology becomes more flexible and robust, precise control over the spectrum of the emission will become routine. The unique combination of quantum coherence and macroscopic physics has made it possible both to optimize emission to a particular harmonic order using pulse shaping [82] and to generate the smooth supercontinuum spectrum of a single attosecond burst [119,149].

4. Laser-based plasma X-rays

Laser-driven plasma sources represent an important laboratory-based alternative to large-scale facilities for ultrafast pulses in the hard X-ray range. Recent advances in femtosecond optical laser drivers and optimized target geometries (figure 10) now enable the generation of 100 fs hard X-ray pulses (temporally and spatially incoherent) at kilohertz repetition rates with negligible timing jitter relative to optical pulses from the drive laser [150–156]. The total generated X-ray
Debye–Scherrer rings

$E$ (V/nm)

$K_{\alpha}$ photons per shot in solid angle $4\pi$

$\lambda = 3.9 \mu$m

$\lambda = 0.8 \mu$m

$\Delta I(hkl)/I_0(hkl)$ (%)

$q = \sin(\theta)/\lambda(\text{Å}^{-1})$

$\Delta I(111)/(111)$ (%)

Figure 10. (a) Laser–target interaction geometry and X-ray generation process in laser-driven femtosecond hard X-ray sources. (b) Generated X-ray flux per surface area (symbols), i.e. $K_{\alpha}$ photons per shot into solid angle $4\pi$ for 50 fs pulses (beam diameter $d_{\text{FWHM}} = 2.6 \mu$m) centred at $\lambda = 0.8 \mu$m (blue) or 80 fs pulses ($d_{\text{FWHM}} = 21 \mu$m) at 3.9 \mu m (red) as a function of the electric field amplitude perpendicular to the metal surface (experimental data from [150]). Solid lines: model calculations including space-charge and radiative-coupling effects, dashed lines: calculations without the latter effects. (c) The Debye–Scherrer pattern (partial ellipses) of an LiH powder measured with copper $K_{\alpha}$ radiation from a laser-driven femtosecond hard X-ray source. (d) Intensity integrated along the rings as a function of the scattering vector $q$ with Bragg reflections as indicated. (e) Relative intensity changes of the (111) reflection as a function of the pump–probe delay when polarizing off-resonantly the LiH crystallites with 50 fs pump pulses centred at $\lambda = 0.8 \mu$m.

Flux is on the order of $5 \times 10^{10}$ Cu $K_{\alpha}$ photons per second (8.04 keV photon energy), of which several $10^{-4}$ are typically collected with X-ray optics and focused onto a sample.

Table-top ultrafast plasma X-ray sources have found widespread application in femtosecond time-resolved X-ray diffraction studies of photo-induced structural dynamics in single crystals [157–160] and polycrystalline powders. Femtosecond powder diffraction is a particularly effective method for probing intensity changes of a multitude of Bragg reflections simultaneously. Differential electron density maps $\Delta \rho(r,t)$ are derived from the transient diffraction patterns, thereby providing direct information on coherent atomic motions and changes in electronic charge distributions between the atomic or ionic entities in the unit cell of the material [161–171]. The time resolution of such pump–probe diffraction studies is currently approximately 100 fs [155], and the spatial resolution of the reconstructed electron density maps is approximately 50 pm, essentially determined by the X-ray wavelength and by the diffraction angle spanned by the X-ray area detector as shown in figure 10c,d. As current X-ray detector technology provides almost 100% quantum efficiency and single-photon-counting sensitivity, femtosecond X-ray powder diffraction experiments [161–171] or experiments using the rotation method [163] can be performed close to the shot-noise limit [156]. Typical experiments use drive laser pulses at 800 nm from Ti:sapphire laser systems operating at 1 kHz repetition rates, which allows for detection of relative intensity changes $\Delta I(hkl)/I_0(hkl)$ as small as several $10^{-3}$ [164,170].

There are three key steps in the laser-plasma hard X-ray generation process: (i) creation of free electrons by field-induced tunnelling from the target into vacuum (figure 10a, step 1), (ii) electron acceleration in the vacuum by the strong laser field (figure 10a, step 2), and (iii) electron re-entrance into the target, collisional inner-shell ionization and X-ray emission by a radiative transition of an outer-shell electron into the unoccupied inner shell (figure 10a, step 3). This scheme requires a strong electric field component perpendicular to the target surface as exists for the $p$-polarized driving field under a large angle of incidence $\theta$ as shown in figure 10a.
The scientific need for higher flux in the hard X-ray regime has motivated the development of advanced schemes based on optical drivers at longer mid-infrared wavelengths to provide stronger optical accelerating fields [150,155]. In recent first experiments, sub-100 fs optical pulses with up to 15 mJ pulse$^{-1}$ at a centre wavelength of 3.9 µm were generated in an optical parametric chirped-pulse amplification (OPCPA) laser system at a 20 Hz repetition rate. The typical laser–target interaction geometry and X-ray generation process is sketched in figure 10a. The p-polarized mid-infrared pulses are focused onto a Cu tape target ($D_{\text{tape}} = 20$ µm thick), in vacuum, with an adjustable angle of incidence $\theta$ as shown in figure 10a. The target is translated at 5 cm s$^{-1}$ to provide a fresh target volume for each drive laser pulse. In this interaction geometry, a standing optical wave is formed at the front target surface by constructive interference between the incoming and the partly reflected laser beam, which results in a peak intensity of several $10^{16}$ W cm$^{-2}$ and a peak electric field of 450 V nm$^{-1}$ for an incident pulse energy of $W_P = 15$ mJ ($\theta = 59^\circ$). The optical spot size on the target (derived from knife-edge measurements) is $d_{\text{FWHM}} = 21$ µm. The generated X-rays are collected in a transmission geometry in the forward direction and detected with a calibrated energy-resolving CdTe detector.

In figure 10b, the generated X-ray flux per surface area (i.e. $K_\alpha$ photons per shot into solid angle $4\pi$) is compared for the mid-infrared driver (red symbols) with the X-ray flux from a source driven by 50 fs pulses centered at $\lambda = 0.8$ µm from a Ti : sapphire drive laser (blue symbols). The experimental data are plotted as a function of the electric field amplitude perpendicular to the metal surface (experimental data from [150]). The comparably long optical period of the mid-infrared laser field allows for accelerating electrons from the Cu target to very high kinetic energies and for generating a characteristic $K_\alpha$ flux of $10^9$ photons per pulse, 25-fold more than with the 800 nm drive laser. Moreover, there is a rescaling of the X-ray output towards smaller driving fields.

Theoretical simulations accurately account for the experimental results over a wide range of driving fields, and predict a further enhancement of X-ray flux with larger driving field [155]. The solid lines in figure 10b show model calculations including space-charge and radiative-coupling effects, in very good agreement with the experiments. By contrast, calculations without the latter mechanisms predict a simple $\lambda^2$ scaling law of X-ray intensity with the driving wavelength $\lambda$, which breaks down for very long wavelengths and large angles of incidence, $\theta > 60^\circ$. The theoretical model [155] includes tunnelling of electrons from the metal to the vacuum as a first step, an extension of the vacuum-heating model by Brunel [172] which considers the metal as an inexhaustible electron reservoir, an approximation that breaks down for high electric fields perpendicular to the surface.

Figure 10c–e displays results from a femtosecond X-ray powder diffraction experiment which demonstrate the high time resolution of such pump–probe studies [168]. Field-induced intensity changes were studied on different Bragg reflections $(hkl)$ from LiH as a function of the pump–probe delay. Non-resonant 50 fs pump pulses centered at $\lambda = 0.8$ µm electronically polarize the LiH crystallites. The stationary Debye–Scherrer pattern (partial ellipses) of an LiH powder is shown in figure 10c. Intensity integrated along the rings is plotted as a function of the scattering vector $q$ in figure 10d (Bragg reflections as indicated). Relative intensity changes of the (111) reflection as a function of the pump–probe delay are presented in figure 10e. The time trace follows the cross-correlation between the 800 nm pump and the X-ray probe pulses and demonstrates a time resolution on the order of 100 fs.

Currently, a novel mid-infrared OPCPA driver providing millijoule pulses at a centre wavelength of 5 µm and kilohertz repetition rates [173] is being implemented at the Max-Born-Institute in Berlin to generate an even higher ultrafast hard X-ray flux of high stability for future diffraction experiments.

5. Synchrotron-based methods

X-ray pulses from electron storage rings offer important science opportunities for time-resolved X-ray science with high spectral resolution and high stability owing to the inherent features
Figure 11. Illustration of femtosecond X-ray generation in a dispersive arc of a storage ring. An fs-laser pulse, co-propagating with the e-beam in a resonantly tuned wiggler, modulates the electron energy. A subsequent dispersive dipole magnet transversely separates the energy-modulated electrons (red) from the main e-bunch (blue). X-rays from unperturbed electrons are blocked and only femtosecond X-rays are directed to the beamline optics. The length of the X-ray pulses is determined by the laser pulse length, elongated by the dispersive element which leads to X-ray pulses of approximately 100 fs FWHM. (Online version in colour.)

...of synchrotron radiation from stored e-beams in stable closed orbits. Moreover, the maximum repetition rate of such X-ray pulses is intrinsically high as determined by the synchrotron RF drive which is typically approximately 500 MHz. However, in an electron storage ring, accelerator physics limit the achievable electron bunch duration and thereby limit the achievable time resolution of experiments with synchrotron-based short X-ray pulses.

The e-bunch length in a storage ring is a consequence of the equilibrium particle dynamics given by key accelerator design parameters like ring circumference, energy, RF frequency and RF voltage that define the longitudinal phase-space [174]. Modern third-generation synchrotron facilities are designed to maximize both the average beam brightness and the e-bunch storage lifetime (see [174]), which typically results in rms bunch lengths of the order of a few 10 ps to a few 100 ps. This readily enables sensitive time-resolved X-ray studies of chemical [175,176], structural [177] and magnetic [178,179] dynamics on the subnanosecond timescale. In the following, we highlight a few key methods that have been developed and implemented to extend the time resolution at synchrotron sources to the few-picosecond and sub-picosecond regimes.

(a) Laser manipulation of e-beams

As originally proposed by Zholents & Zolotorev [180] in 1996, femtosecond synchrotron pulses can be generated by using a femtosecond laser pulse to modulate the energy of a femtosecond slice of a long relativistic electron bunch in a storage ring. These modulated electrons can then be used to generate femtosecond X-rays that can be separated from the long-pulse background [5,6,181]. Figure 11 shows a schematic of a typical implementation of this approach where a femtosecond laser pulse of approximately 1 mJ energy co-propagates in a wiggler which is tuned to resonance (i.e. emitting undulator radiation at the laser wavelength, 800 nm, on-axis). The electrons in the temporal region of the e-bunch overlapping with the laser field are modulated in energy (i.e. accelerated and decelerated, depending on the optical phase) by an amount $\Delta \gamma / \gamma \approx \pm 1\%$. Following dispersive elements in the storage-ring lattice then separate the modulated electrons in transverse phase-space. In a subsequent undulator, these modulated electrons then emit a cone of radiation that is separable from the long-pulse background in the transverse position and/or emission angle. The radiated ultrashort X-ray pulse of few 100 fs FWHM is naturally synchronized on a sub-100 fs scale to the modulating laser pulse.

The generated femtosecond X-ray pulses share all the attributes of conventional synchrotron pulses with the full flexibility and peak brightness of undulator radiation, including continuous tunability and complete polarization control. The average spectral flux and brightness of the femtosecond X-rays scales from the full synchrotron flux and brightness by three factors: $\eta_1 = \tau_{\text{laser}} / \tau_{\text{e-bunch}} \approx 10^{-3}$, $\eta_2 = f_{\text{laser}} / f_{\text{e-bunch}} \approx 10^{-4}$ and $\eta_3 \approx 0.1$, where $\tau_{\text{laser}}$ and $\tau_{\text{e-bunch}}$ are the laser pulse and electron bunch durations, respectively; $f_{\text{laser}}$ and $f_{\text{e-bunch}}$ are the laser and...
electron bunch repetition rates, respectively; and $\eta_3$ accounts for the fraction of electrons that are in the proper phase of the laser pulse to get the maximum energy exchange for creating the large transverse separation. Synchrotron radiation damping (typically $T_{\text{damp}} \sim 10\,\text{ms}$) provides for recovery of the electron beam between interactions. In principle, the maximum repetition rate is given by $3N_{\text{e-bunch}}/T_{\text{damp}} \sim 100\,\text{kHz}$, where $N_{\text{e-bunch}}$ is the number of bunches in the storage ring (and assuming that the laser interacts sequentially with each of the $N_{\text{e-bunch}}$ bunches) [5,180,182,183]. Femtosecond hard and soft X-ray beamlines operating in the few-kilohertz regime have been implemented at synchrotron storage rings, including the Advanced Light Source [5], BESSY II [6,184], Swiss Light Source [7] and most recently at SOLEIL [185].

(b) Low-\(\alpha\) operation

Another approach to extend the time resolution of storage-ring X-ray sources to the few-picosecond regime, while exploiting all of the stored electrons at the full repetition rate, relies on a different stable closed e-bunch orbit with a minimized energy/momentum correlation [186,187]. The momentum-compaction factor, $\alpha_c$, relates differential changes in electron orbital path length ($\Delta L$) with differential changes in electron energy ($\Delta E$) as given by: $\Delta L/L = \alpha_c \Delta E/E$. In the ‘zero-current’ limit (i.e. ignoring collective electron effects), the nominal e-bunch length scales as: $\sigma_{\text{e-bunch}} \sim \sqrt{\alpha_c/V'}$, where $V'$ is the temporal gradient of the accelerating RF voltage. The scaling of e-bunch length with $\alpha_c$ and e-bunch current is illustrated by the red curves in figure 12, taking the example of BESSY II [187]. The interaction of electrons with their own coherent synchrotron radiation (CSR), and with the impedance of the storage-ring vacuum chamber, leads to a nonlinear increase in $\sigma_{\text{e-bunch}}$ with e-bunch current. This e-bunch length asymptotically approaches the CSR-instability threshold [188] at high current [189,190]. Thus, short e-bunch lengths (and few-picosecond X-ray pulses) are generated only at small $\alpha_c$ and at low currents (few $\mu\text{A}$/bunch compared to typical approx. 1 mA bunch$^{-1}$). Compared to the typical operating mode (‘user optics’ with $\alpha_c = 7.3 \times 10^{-4}$) X-ray pulses of approximately 2 ps duration can be generated in an optimized low-$\alpha_c$ mode (‘THz optics’ with $\alpha_c = 3.5 \times 10^{-5}$) [187].

(c) Variable pulse length storage ring

As indicated above, a higher RF voltage temporal gradient, $V'$, from superconducting RF cavities can also be used to shorten the electron bunches [187], and thereby avoid the drawbacks of a
modified optics (\(\alpha_c\)) while preserving higher storage-ring current as indicated by the blue curves in figure 12. Thus, short X-ray pulses (e-bunches) can be generated at approximately 100-fold higher average brightness (e-bunch current) as determined by the increase in \(V'\).

The Helmholtz-Zentrum Berlin is now implementing a new scheme, BESSY VSR (as summarized in table 2) [191] to provide short pulses (1.7 ps) and long pulses (15 ps) simultaneously in a mixed filling pattern, while preserving the average brightness so that users can select pulse durations on demand to suit their experiments. Briefly, BESSY VSR is based on an idea by Jankowiak & Wüstefeld [187] to use two additional superconducting RF cavity units, one operating at the third harmonic of 500 MHz (1.5 GHz) and the other operating at the 7/2 harmonic (1.75 GHz). These two beating frequencies (with appropriately adjusted amplitude and phase) can be made to cancel the RF voltage for every second e-bunch (250 MHz), thus generating long bunches at 250 MHz (as dictated by the 500 MHz RF system) interleaved with short bunches (where the RF voltages add constructively). The fill pattern for the short bunches (1.7 ps at 0.8 mA as shown in figure 12) will be limited to few-megahertz repetition rates in order to limit the wakefield heating effects on storage-ring components, while long bunches will be limited to few-megahertz repetition rates in order to preserve the average brightness [191]. Even shorter bunches (\(\sigma_{e\text{-bunch}} \sim 0.4\) ps rms at 0.04 mA bunch\(^{-1}\)) can be achieved combining high \(V'\) with low-\(\alpha_c\) at expense of average brightness [192].

Related to the VSR approach is the proposed RF orbit-deflection scheme which relies on RF ‘crab cavities’ to tilt a long e-bunch in the vertical–longitudinal (y–z) plane. By using a slit in the X-ray beam further downstream, X-rays only from a short (approx. 1 ps) slice of the e-bunch can be selected at dedicated beamlines [193–195] before restoring the e-bunch tilt to the original orbit. This approach is under consideration at a number of facilities, but at present there are no definite plans to implement it.

(d) Pulse separation schemes

Two important developments for time-resolved X-ray methods at synchrotron storage rings are new methods for selecting specific pulses from the nominal 500 MHz train. The pseudo-single-bunch operational mode, developed at the Advanced Light Source [196], employs a fast-deflecting pulsed stripline kicker to perturb a single e-bunch in transverse direction in order to separate the X-rays generated from a single e-bunch at kilohertz repetition rates. More recently, it has been

<table>
<thead>
<tr>
<th>facility</th>
<th>peak brilliance</th>
<th>average brilliance</th>
<th>number of bunches</th>
<th>pulse duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>BESSY II</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>standard</td>
<td>(6.1 \times 10^{21})</td>
<td>(4 \times 10^{18}) (300 mA)</td>
<td>350</td>
<td>15</td>
</tr>
<tr>
<td>low-(\alpha)</td>
<td>(1.9 \times 10^{20})</td>
<td>(2.5 \times 10^{17}) (15 mA)</td>
<td>350</td>
<td>3</td>
</tr>
<tr>
<td>BESSY VSR</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>total</td>
<td>varying</td>
<td>(4 \times 10^{18}) (300 mA)</td>
<td>varying</td>
<td>varying</td>
</tr>
<tr>
<td>std. long pulses</td>
<td>(1.2 \times 10^{22})</td>
<td>(3.3 \times 10^{19}) (248 mA)</td>
<td>150</td>
<td>15</td>
</tr>
<tr>
<td>std. short pulses</td>
<td>(1.7 \times 10^{22})</td>
<td>(3.6 \times 10^{18}) (27 mA)</td>
<td>150</td>
<td>1.1</td>
</tr>
<tr>
<td>long camshaft</td>
<td>(3.95 \times 10^{22})</td>
<td>(1.3 \times 10^{18}) (10 mA)</td>
<td>1</td>
<td>27</td>
</tr>
<tr>
<td>short camshaft</td>
<td>(5.1 \times 10^{22})</td>
<td>(1 \times 10^{19}) (0.8 mA)</td>
<td>1</td>
<td>1.7</td>
</tr>
<tr>
<td>low-(\alpha)</td>
<td>(2.2 \times 10^{22})</td>
<td>(1.2 \times 10^{18}) (7.5 mA)</td>
<td>150</td>
<td>0.4</td>
</tr>
</tbody>
</table>
demonstrated at BESSY II [197] that a second stable e-bunch orbit can be supported by operating on a horizontal third-order resonance. The transverse resonant island bucket (TRIB) mode is a stable closed orbit winding around the standard core e-bunch orbit and closing after \( n \) turns, whereby \( n \) is the order of the resonance. TRIB is presently being developed as a bunch separation scheme [198,199], with the first promising results exploiting different X-ray emission angles for the two orbits in an undulator. The TRIB scheme seems compatible with ‘top-up’ operation and looks promising for a future BESSY VSR user operation with flexible pulse selection.

6. Summary

Recent advances in ultrafast X-ray sources promise to transform the field of X-ray science in the twenty-first century. The generation of high-brightness (in some cases, fully coherent) ultrafast X-ray pulses spanning the EUV, soft X-ray and hard X-ray ranges now enable sensitive probing of structural dynamics in matter on the fundamental timescales of atomic motion. The development of tuneable subfemtosecond coherent X-ray sources further enables element-specific probing of electronic structure, charge dynamics and coupling to atomic structure on fundamental timescales of electronic motion. Synchrotron-based X-ray sources have been the workhorse in the field of X-ray science for nearly half a century, and effective methods have been developed to extend the time resolution of these sources to the sub-picosecond regime. Table-top plasma X-ray sources have been refined to become compact, reliable sources of less than 100 fs hard X-ray pulses at the laboratory scale, and table-top HHG XUV sources have opened the attosecond time domain. Most significantly, XFELs have been demonstrated, and are proliferating worldwide. They provide tuneable coherent X-ray beams, many orders of magnitude brighter than any previously available X-ray source, with pulse durations that are now extending to the attosecond domain.

Data accessibility. This article has no additional data.

Competing interests. The authors declare that they have no competing interests.

Funding. This work was supported by the US Department of Energy, Office of Science, Basic Energy Sciences under Contract no. DEAC02-76SF00515 (R.S., Z.H.). The research concerning laser-plasma sources was supported by the Deutsche Forschungsgemeinschaft (grant no. WO 558/13-2). H.K. and M.M. acknowledge funding from the US National Science Foundation, the US Department of Energy and the Defence Advanced Research Projects Agency.

Acknowledgements. K.H. is indebted to A. Jankowiak and P. Goslawski for valuable contributions. The research concerning laser-plasma sources was done in collaboration with the group of A. Baltuska, University of Vienna, Austria.

References

45. Tschentscher T. 2011 Layout of the X-ray systems at the European XFEL.


174. Wiedemann H. 1999 *Particle accelerator physics II: nonlinear and higher-order beam dynamics*. Berlin, Germany: Springer.


