The fundamental building blocks that constitute all matter – animate and inanimate – interact on extremely small (sub-)atomic distances (1 Å = 10^{-10} m) and timescales (1 fs = 10^{-15} s). Unravelling these ultrasmall as well as ultrafast dynamics would not only expand our understanding of the fascinating nanoworld but may also lay the groundwork for the development of innovative nanotechnology.

Lightwave-driven scanning tunnelling microscopy (LW-STM) can be used to observe and even control these elementary dynamics, which lie at the core of all matter. This approach complements the remarkable atomic spatial resolution achieved by conventional STM with a simultaneous ultrashort temporal resolution by employing the fastest transient bias voltage nature has to offer: the oscillating carrier wave of light. By coupling a light pulse, whose duration is shorter than a single cycle, into the tunnelling junction, the tunnelling window into a specific energy level can be selectively opened for durations as short as 100 fs.

Employing this sophisticated technique, we capture a nanomovie of the sub-picosecond (1 ps = 10^{-12} s) coherent dynamics of a single-molecule rotator upon the application of local ultrafast forces, demonstrating the coherent control over a molecular switching reaction. Moreover, the exact shape of the electromagnetic near-field waveform within the atomic-scale tunnelling junction is traced quantitatively for the first time, paving the way for the predictive design of nanophotonic structures.

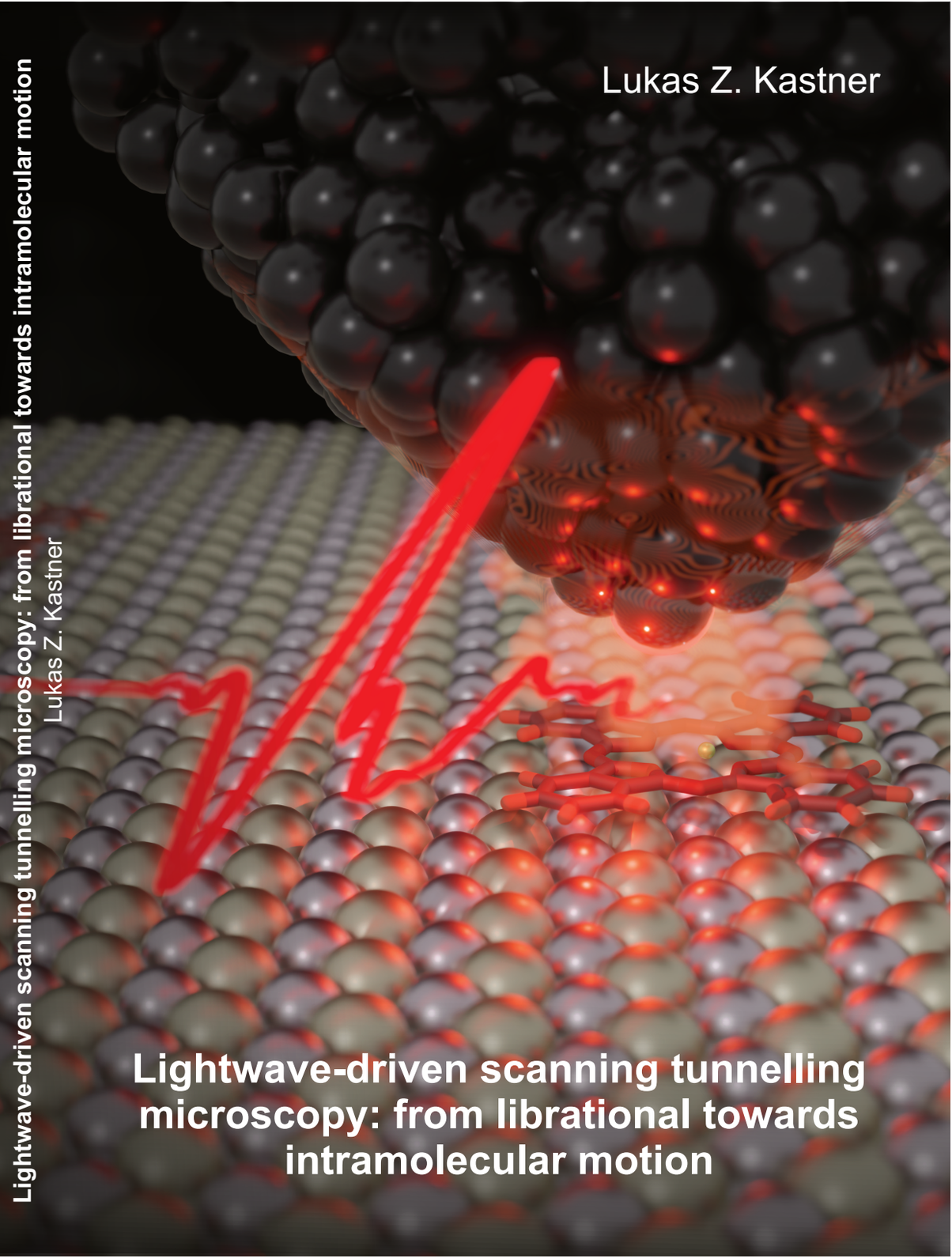
As all motion is intricately linked to changes in the energetic structure of a system, we aimed to further elucidate nanoscopic phenomena by visualising the temporal evolution of energy levels after ultrafast excitation. To this end, we investigate an atomic vacancy within a tungsten-diselenide monolayer using LW-STM. The first observation of shifts of localised energy levels at an atomic vacancy with combined Ångström spatial, sub-picosecond temporal and 10 meV energy resolution is demonstrated.

Finally, to enhance the capabilities of LW-STM, a cutting-edge setup consisting of an improved STM scan head and a laser source operating at higher repetition rates, field-strengths, and frequencies is introduced. The use of mid-infrared frequencies allows for the temporal confinement of tunnelling events to a window less than 10 fs wide. In combination with a tunable source of ultrashort visible pulses, this novel setup will allow for the real-space investigation of intramolecular vibrations, electronic motion, and the complex interaction of various degrees of freedom.

Universität Regensburg
Institut für Experimentelle und Angewandte Physik
Lehrstuhl Prof. Dr. Rupert Huber

Lightwave-driven scanning tunnelling microscopy: from librational towards intramolecular motion
Lukas Z. Kastner

Lukas Z. Kastner



Lightwave-driven scanning tunnelling microscopy: from librational towards intramolecular motion

Cover: Artistic visualisation of a sub-cycle light pulse (left) impinging on a tunnelling junction between a tungsten tip (black spheres) and a single-molecule rotator (magnesium-phthalocyanine) adsorbed on a sodium-chloride lattice (bright spheres).

Lightwave-driven scanning tunnelling microscopy: from librational towards intramolecular motion

Dissertation
zur Erlangung des Doktorgrades der
Naturwissenschaften
(Dr. rer. nat.)
der Fakultät für Physik der Universität Regensburg
vorgelegt von
Lukas Zheyi Kastner
aus Regensburg
im Jahr 2023

Das Promotionsgesuch wurde eingereicht am: 05.07.2023
Die Arbeit wurde angeleitet von: Prof. Dr. Rupert Huber.

Prüfungsausschuss:	Vorsitzender:	Prof. Dr. Klaus Richter
	1. Gutachter:	Prof. Dr. Rupert Huber
	2. Gutachter:	Prof. Dr. Jascha Repp
	weiterer Prüfer:	Prof. Dr. Christian Schüller

Contents

1 Tracking the building blocks of condensed matter with ultrafast nanoscopy **1**

2 Lightwave-driven scanning tunnelling microscopy **7**

2.1 Introduction to conventional scanning tunnelling microscopy techniques 8

2.2 Laser-assisted scanning tunnelling microscopy 14

2.3 Ultrafast state-selective lightwave-driven scanning tunnelling microscopy 17

2.4 Shot-to-shot detection of single-molecule switching statistics 27

2.5 Sampling ultrastrong atomically confined light fields 36

2.6 Controlled excitation of librational motion 44

2.7 Towards lightwave scanning tunnelling microscopy of non-classical quantum motion 49

3 Controlling phononic motion and ultrafast energy level shifts of atomic defect levels in monolayer tungsten diselenide **57**

3.1 Monolayer transition-metal dichalcogenides 59

3.2 Exciting phononic motion in a monolayer WSe_2 65

3.3 Numerical calculations to elucidate the origin of current oscillations . 70

3.4 Ultrafast snapshots of the local density of states 73

3.5 Exciting and monitoring ultrafast shifts of single-defect energy levels . 81

3.6 Numerical calculations on the energy level shifts 84

3.7 Preliminary data on optical excitation of WSe_2 monolayers and homobilayers 85

4 Femtosecond LW-STM: towards the visualisation of intramolecular motion 93

4.1 Optimisation of ultrashort near-infrared laser pulses 98
4.2 Pulse compression in Herriott-type multi-pass cells 111
4.3 Generation of single-cycle mid-infrared pulses 130
4.4 Michelson interferometer and mid-infrared beam steering 147
4.5 Generation of ultrashort optical pump pulses 152
4.6 Optimised STM scan head for lightwave-driven scanning tunnelling microscopy 164

5 Summary and outlook 173

Appendices 179

A Single-molecule oscillator 179

A.1 Spatial map of the switching probability 179
A.2 Extended data on quantitative sampling of the near field 179
A.2.1 Far-to-near-field transfer function 179
A.2.2 Simulation: Retardation effects in the near field 181
A.3 Field-induced in-plane rotation 182
A.4 Extended data on non-classical oscillator states 184
A.4.1 Bootstrapping 184

B Single-atom vacancy in monolayer WSe_2 187

B.1 Ultrafast tunnelling model 187
B.2 Analysis of femtosecond snapshots 189
B.3 Tunnelling decay constant 190
B.4 Long-lived defect oscillations 191
B.5 DFT calculations of ultrafast energy shifts of localised defect levels . 192
B.6 Characterisation of the ultrafast voltage transient 194
B.7 Bias-modulation for quicker spectroscopy measurements 196

C Ultrashort laser pulses for MIR-STM **201**
C.1 Extended data on the amplifier characterisation 201
C.1.1 Spatial homogeneity 201
C.1.2 Pulse contrast 204
C.2 Extended data on the Herriott cells 205
C.2.1 Saturation effect in Herriott cells. 205
C.2.2 Optimised spectral broadening in Herriott cells 206
C.2.3 Compression of the HC2 output pulses 209
C.3 Extended data on mid-infrared generation 211
C.3.1 Simulation of optical rectification 211
C.3.2 Mid-infrared generation focus 214
C.3.3 Optimisation of mid-infrared generation 214

D Publications **221**

Bibliography **225**

Acknowledgements **251**

Chapter 1

Tracking the building blocks of condensed matter with ultrafast nanoscopy

"Dass ich erkenne, was die Welt
Im Innersten zusammenhält."

"That I may grasp forevermore
What binds all matter at its core."
(loose translation)
Faust, Part I (1808), Johann Wolfgang von Goethe

In the early 19th century, Goethe's Faust was not the first yet would not be the last persona to muse about the inner workings of the world around us. Already in ancient Greece, great philosophers like Leucippus pondered such fundamental questions, and humanity's persistent endeavour to understand the elementary building blocks of matter has only grown with the technological advancement of recent times. To satisfy this inherent curiosity of mankind as well as to continuously drive innovation in quantum-, nano- and biotechnologies, we need to find ways to delve ever deeper into the nanocosmos. This pursuit of innovation besides pure curiosity is motivated by the desire to revolutionise the fields of nano- and life sciences and unlock new possibilities in areas such as electronics, quantum information processing, chemical reaction control, photovoltaics, optogenetics, molecular nanotechnology, and beyond. In the 19th and 20th centuries, the concept of atoms, molecules, and electrons, which

had so far been only the subject of philosophical speculation, was finally backed with experimental evidence [Dal08, Avo11, Tho01] and became the prime description of all matter – both living and non-living – around us. Direct visualisation of these elementary building blocks became a reality only much later due to their minuscule size in the range of 100 pm (1 pm = 10^{-12} m) to nanometres (1 nm = 10^{-9} m). Facilitated by its resolution on the order of tens of nanometres, high-resolution microscopy (Nobel Prizes in Physics in 1986 [Nob86] as well as Chemistry in 2014 [Nob14] and 2017 [Nob17]) has allowed researchers a glance into this whole new world that is dictated by aspects of both classical and quantum science. Using a plethora of scanning probe techniques such as scanning tunnelling microscopy (STM) or atomic force microscopy (AFM), the exploration of structures on even smaller length scales of single atoms – typically defined as around one Ångström (1 Å = 10^{-10} m) – was made possible. This investigation of local physical properties at the nanoscale helps in disentangling the various (and often competing) effects that influence the electronic, optical, and geometric properties of materials.

However, the microscopic world is never static but inherently dynamic, and it is the dynamics that govern most phenomena. Energy is unceasingly exchanged between different subsystems or modes, atomic nuclei vibrate, electronic ordering arises and disappears, charge carriers relocate and bind to each other, and more. Thus, to answer the pressing questions of modern science, from unlocking the potential of quantum materials to controlling biochemical reactions, it is crucial to directly observe the intricate interplay between the nanoscale building blocks of matter with high spatial, temporal, and potentially even energy resolution. Yet, this task is incredibly demanding as these elementary processes unfold within femtoseconds (1 fs = 10^{-15} s) or even attoseconds (1 as = 10^{-18} s). Watching molecules move and react with one another has been a fantasy of natural scientists since the concept of these tiny fundamental objects of matter was postulated by Greek philosophers such as Leucippus and his disciple Democritus [Lae53]. Thus, the development of innovative approaches that enable the simultaneous resolution of ultrashort time as well as length scales stands as one of the pivotal scientific and technological challenges of our time.

Since more than 50 years, researchers in the field of ultrashort laser physics have continuously pursued the generation of ever shorter laser pulses. The invention of the titanium:sapphire laser, initially emitting nanosecond pulses [Mou82], significantly

advanced the field. Nowadays, high-power lasers with picosecond, femtosecond, and even attosecond pulse durations are available [Kra09]. Employing such ultrashort pulses for the study of condensed matter has allowed for numerous intriguing achievements, including the extraction of time-dependent dielectric functions with ultrafast temporal resolution [Eis14] and the driving of coherent interband polarisation in semiconductors [Sch14].

Yet, the spot size, to which one can focus light and thus is able to extract information on the local properties of the sample, is limited by the diffraction limit of half the wavelength of the utilised radiation. First formulated by Ernst Abbe [Abb73], this principle limits the spatial resolution in ultrafast far-field photonics to at best a few tens of nanometres even when using the short wavelengths of extreme ultraviolet light. Thus, the fascinating nanoscale variations in physical properties and their underlying origins cannot be resolved in purely laser-based microscopy.

In parallel to the advent of ultrafast laser physics, the field of scanning probe microscopy has become the standard approach for the investigation of nanoscopic and atom-sized variations in the topography and electronic properties of various sample surfaces. However, these techniques were limited by the bandwidth of the electronics used – typically in the gigahertz range –, resulting in a temporal resolution in the nanosecond regime at best [Lot10].

Thus, the reconciliation of the seemingly conflicting goals of ultrafast temporal precision and sub-nanometre spatial resolution to ultimately capture actual "movies from the nanoworld" [Pep17] requires us to simultaneously harness the advantages of multiple cutting-edge fields.

After more than 50 years of development in the two fields, femtosecond optics and scanning probe microscopy, the combination of their highest temporal and spatial resolution has only recently become attainable [Coc16]. State-of-the-art techniques such as lightwave scanning tunnelling microscopy (LW-STM), ultrafast scanning-type near-field optical microscopy (U-SNOM) [Eis14], ultrafast atomic force microscopy (U-AFM) [Sch17b], and ultrafast transmission electron microscopy (U-TEM) [Fei17] have paved the way for achieving unprecedented simultaneous spatial and temporal resolutions (see Fig. 1.1). However, simultaneous atomic (sub-Ångström) spatial resolution and 10 fs temporal resolution has not been achieved with these methods so far. Among the existing techniques, U-SNOM and LW-STM stand out particularly, hav-

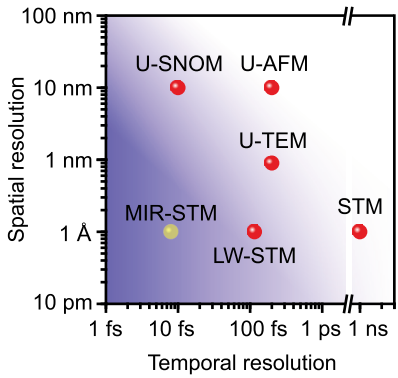


Figure 1.1 | Ultrafast nanoscopy techniques. Existing state-of-the-art ultrafast nanoscopy systems (red) such as ultrafast scattering-type near-field microscopy (U-SNOM), ultrafast transmission electron microscopy (U-TEM), and ultrafast atomic force microscopy (U-AFM) reach temporal resolutions of tens to hundreds of femtoseconds. By extending a conventional STM into the ultrafast regime using light-wave biasing, a time resolution of 100 fs has been shown in LW-STM. A mid-infrared driven STM (MIR-STM, yellow) can push the temporal resolution to below 10 fs.

ing enabled spatio-temporal resolutions of 10 nm/10 fs [Eis14] and 1 Å/100 fs [Coc16], respectively. Thus, they hold immense potential for pushing the boundaries towards a simultaneous spatio-temporal resolution better than 1 Å and 10 fs, possibly enabling the observation of intramolecular vibrations as well as single-electron motion within solids and individual molecules. By employing such innovative ultrafast nanoscopy techniques, glimpses into the ultrafast microscopic dynamics could potentially be made a reality. Thanks to these advanced approaches, we now are on the brink of unravelling the fundamental interactions that bind our world together – a feat that would surely impress luminaries like Faust, Leucippus, and Democritus.

Presenting the intriguing insights achieved using sophisticated nanoscopy experimental designs, the thesis at hand is structured in the following way: In **Chapter 2**, I will introduce the underlying concepts of STM and single-orbital imaging (sec. 2.1), before elaborating on the recent breakthroughs in state-of-the-art LW-STM systems (sec. 2.2). An innovative approach confining state-selective tunnelling to femtosecond timescales allowed for the capturing of the first ultrafast movie of the dynamics occurring in the quantum world after transfer of a single electron from the highest occupied state of a molecule into the STM tip (sec. 2.3).

Moreover, tunnelling in the opposite direction is possible as well. I will demonstrate how a molecular switch could be toggled from one geometry into another by injection of a single electron into a specific empty state (sec. 2.4). As ultrafast light-matter

interaction not only depends on the pulse duration and field strength but the exact waveform used, precisely mapping out the temporal evolution of electric field transients is of utmost importance. So far, reading out the oscillating electric near field within nanoscale structures has been highly challenging. I will present the first quantitative extraction of near-field transients within the Ångström-sized tip-sample junction utilising the molecular switch introduced in the previous section as a nanoscopic voltage gauge (sec. 2.5). Furthermore, apart from single-electron injection or extraction, the ultrafast bias pulses were also used to directly and selectively control single-molecule dynamics via femtosecond atomic forces [Pel20b] (sec. 2.6). The access to the nanocosmos offered by LW-STM due to its femtosecond, single-electron, state-selective control on atomic length scales raises the question whether this technique can be used to probe remnants of quantum behaviour: As I will discuss in section 2.7, it could indeed be possible to engineer non-classical librational quantum motion of a single molecule oscillator using femtosecond forces and probe its dynamics with single-electron tunnelling within LW-STM.

Motion of the atomic nuclei does not only occur in molecular systems but also in more extended crystalline structures, as will be closely elaborated on in **Chapter 3**. The ubiquitous collective vibrations in solids known as phonons strongly influence the physical properties of a material. Here, I focus on an atomically-thin tungsten diselenide monolayer. First, the material class will be introduced in section 3.1. Subsequently, I will present our capability to excite a vertical phonon mode whose effect on the monolayer was investigated with sub-picosecond resolution. As the evidence will indicate, strong local differences in oscillation behaviour arise when the periodicity of the crystal lattice is disrupted by a single missing atom in the upper selenium sheet of the monolayer (sec. 3.2). Numerical calculations reveal a vertical phonon mode as a possible explanation for these variations (sec. 3.3). Extending the capabilities of LW-STM to ultrafast time-resolved lightwave-driven spectroscopy allows us to investigate changes of the local density of states of nanoscopic entities (sec. 3.4). Specifically, we found that the defect levels arising at the location of a selenium vacancy strongly shift in energy on sub-picosecond timescales following an excitation pulse (sec. 3.5). Possible origins of these shifts lie in changes of intralayer distances and image charge effects (sec. 3.6). Besides investigating the dynamics excited by a terahertz pulse, moving towards higher energy photons in the visible

regime is highly appealing. In section 3.7, I will present our preliminary results of optically pumped mono- and bilayer tungsten diselenide samples and the effect on the tunnelling current.

For the visualisation of processes such as intramolecular vibrations or charge carrier dynamics, reaching a time resolution on the few to ten femtoseconds range is imperative. In **Chapter 4**, I will showcase the concept and realisation of a mid-infrared-driven STM, which is predicted to achieve a remarkable 6 fs temporal resolution by increasing the centre frequency of the employed laser pulses from 1 THz to 25 THz. A thorough description of the intricate optical setup including high-power laser amplifier (sec. 4.1), spectral broadening scheme (sec. 4.2), and nonlinear optical conversion processes (sec. 4.3) to create suitable ultrashort pulses is provided. Within a Michelson interferometer setup, pairs of mid-infrared pulses, which can be individually tuned and delayed with respect to one another, are created, allowing for pump-probe experiments (sec. 4.4). The response of a material to an external optical stimulus depends on the pulse energy as well as the pulse duration and frequency used. To facilitate a broad range of excitation mechanisms, several different sources of laser pulses are set up. I will introduce how we create ultrashort pulses from the mid-infrared to the visible spectral region (sec. 4.5), building on a near-infrared laser amplifier. From the experience and knowledge collected in the first-generation LW-STM, further improvements to the STM scan head were implemented, which are briefly outlined in section 4.6.

In **Chapter 5**, our recent successes in visualising and controlling vibrational dynamics of nanoscopic entities using a state-of-the-art LW-STM using terahertz pulses as ultrashort transient bias are summarised. Moreover, I will provide an outlook on the wide range of possible interesting sample systems and physics that can be explored utilising the remarkable spatio-temporal resolution the new LW-STM has to offer.

Chapter 2

Lightwave-driven scanning tunnelling microscopy

The development of scanning tunnelling microscopy (STM) by Gerd Binnig and co-workers in 1981 has ushered in a new era in nanoscience [Bin82], offering a powerful tool to investigate the electronic properties of nanostructures at the atomic level – meaning length scales in the nanometre and Ångström range. This unprecedented spatial resolution revolutionised the fields of surface science and engineering, as it enables the imaging of individual atoms, which was once deemed to be a mere dream. Beyond observation, STM also permits the direct manipulation of nanoscale objects with sub-Ångström precision, leading to the creation of artificial structures and atoms with novel and intriguing electronic properties [Cro93, Hei02, Str04, Sti21]. To that end, an enormous amount of research and technological developments have been reported over the years, testing various schemes for ultrafast control of the tunnelling process in an effort to enhance the time resolution of STM. One such approach involves the use of ultrashort laser pulses in the terahertz (THz) domain focused onto the tip-sample junction. These pulses generate a rapidly oscillating electric field within the tunnelling gap, acting as an ultrafast bias voltage that drives electron tunnelling currents across the junction. In a seminal experiment conducted by Cocker *et al.* [Coc13] in 2013, the effectiveness of this approach was demonstrated. The authors used this technique to investigate semiconductor nanodots with a spatial resolution of 2 nm by driving currents of a few thousand electrons across the

tunnelling junction with a time resolution as short as 500 fs. Replacing an arbitrary nonlinearity in the current-bias characteristic (I–V curve) of the junction with state-selective tunnelling, only three years later, researchers recorded the first femtosecond movie of a single molecule excited by extracting a single electron from its highest occupied molecular orbital (HOMO) [Coc16]. This groundbreaking work demonstrated single-electron tunnelling into a select molecular orbital with combined atomic spatial and 100 fs temporal resolution. It led to a paradigm shift, offering even more precise handles on the tunnelling current.
Over the last decade, numerous high-ranking studies followed up with ultrafast atomic-scale imaging and even control in STM, revolutionising our perspective of the nanocosmos [Yos16, Jel17, Yos19, Pel20b, Yos21, Amm21, She22]. The development of suitable laser sources and their characterisation in the near field has made considerable progress, allowing for an even more precise control of LW-driven currents [Mül20, Pel21, Abd21, Ara22b, Ara22a]. Furthermore, significant computational advancements have made it possible to model ultrafast tunnelling currents more accurately [Luo20, Amm22].
Only through the combination of two state-of-the-art systems – a high-resolution scanning probe microscope and an elaborate ultrafast photonics setup –, these experiments and the invaluable insights they provide were made a reality. The following sections offer an introduction into the prerequisites and mode of operation of our cutting-edge LW-STM setup by initially introducing conventional all-electronic scanning tunnelling microscopy.

2.1 Introduction to conventional scanning tunnelling microscopy techniques

The core part of an STM is its tunnelling circuit, consisting of a probe tip in tunnelling contact with a conductive sample, a current amplifier, and a voltage source. Figure 2.1 provides a schematic illustration of the basic idea. A tapered metallic wire of a few hundred microns in diameter is placed in close proximity to a conductive sample. The vacuum junction can be conceived as a potential barrier, which prevents direct electron flow between the two electrodes (tip and sample). Due to the quantum

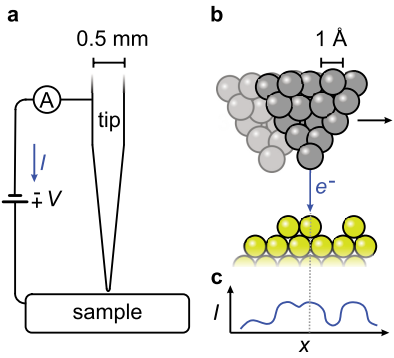


Figure 2.1 | Atomic resolution in STM. **a**, A schematic diagram of an STM measuring circuit is depicted. Applying a bias voltage V between tip and sample, which are in close proximity to each other, leads to a tunnelling current I. **b**, In a close-up sketch of the tunnelling junction, atoms of the tip and the sample are represented as grey and yellow spheres, respectively. The tunnelling probability for electrons critically depends on the topography and electronic structure of the sample. **c**, Recording the tunnelling current I as the tip position x varies, atom-scale features of the sample surface are detected.

mechanical tunnel effect, there is, nonetheless, a non-zero probability for an electron to pass the barrier. A voltage applied across the junction induces a measurable tunnelling current between apex and sample atoms (Fig. 2.1b). In the most simplistic view, the measured current encodes the tip-sample separation due to the strong dependence of the tunnelling probability on the thickness of the potential barrier. Scanning the tip across the sample by means of a piezo-driven stage provides insights on the topography of the sample.

Importantly, for the most stable and noiseless operation of an STM, working in an ultrahigh vacuum (UHV) below 10^{-10} mbar and at cryogenic temperatures is of paramount importance. Cooling the STM scan head by means of liquefied nitrogen and helium to below 8 K reduces all thermal motion to allow for a much more precise measurement. Moreover, working at cryogenic temperatures, the distribution of electrons within the electrodes (tip and sample) becomes particularly simple. While at a finite temperature, the electron distribution is given by the Fermi-Dirac distribution – a function of energy ϵ and temperature T:

$$f(\epsilon) = \left(1 + \exp\left(\frac{\epsilon - E_\mathrm{F}}{k_\mathrm{B}T}\right)\right)^{-1} \tag{2.1}$$

with the Fermi energy E_F and the Boltzmann constant k_B, at 0 K this function simplifies to a step function. This means that all states up to E_F are filled and all above

remain unoccupied. In Fig. 2.2a, the dark grey filled areas mark the filled states in the tip and the sample, respectively. The work function, denoted as Φ_t (Φ_s), is necessary to promote an electron from the tip (sample) into the vacuum (see Fig. 2.2, light grey areas).
Applying a bias voltage V across the junction (see Fig. 2.2b), the filled states of the sample are lowered in energy by elementary charge e times V, leading to the distribution $f(E-eV)$ displayed in Fig. 2.2b. The energy landscape is tilted, allowing electrons to tunnel from filled tip states into unoccupied sample states in spite of possessing less energy than the work function of the tip Φ_t. For higher-lying filled tip states the effective barrier height is lower, leading to a larger tunnelling probability than for lower-lying states as indicated by the thickness of the horizontal arrows.
The probability for an electron to tunnel through the vacuum between the two electrodes depends on the density of occupied and free energy states as well as on a complex interaction of the quantum mechanical wave functions of tip and sample. A theoretical description employing reasonable simplifications for the tunnelling probability was developed by Bardeen [Bar61] using a time-dependent perturbation approach. This model has become the most widely applied theoretical method for studying the metal-insulator-metal, i.e., tip-vacuum-sample, tunnelling junction. Within this theoretical framework, the interaction between tip and sample is considered weak and therefore negligible. Thus, using a separation approach, the stationary Schrödinger equation can be solved for the two subsystems separately. The probability for electron transfer is then given by the wave-function overlap, encoded in the tunnelling matrix element M [Dir27]. The total tunnelling current under biasing is obtained by integration over all energies:

$$I(V) = \frac{4\pi e}{\hbar}\int_{-\infty}^{\infty} d\epsilon\Big(f(E_\mathrm{F}-eV+\epsilon)-f(E_\mathrm{F}+\epsilon)\Big)\cdot\rho_\mathrm{s}(E_\mathrm{F}-eV+\epsilon)\cdot\rho_\mathrm{t}(E_\mathrm{F}+\epsilon)\cdot|M|^2 \quad (2.2)$$

The parameters $\hbar$, ρ_s, and ρ_t denote the reduced Planck's constant and the local density of states (LDOS) of sample and tip, respectively. For cryogenic temperatures, where the Fermi distribution $f(\epsilon)$ can be approximated by a step function, the calculation is simplified to:

$$I(V) = \frac{4\pi e}{\hbar}\int_{0}^{eV} d\epsilon\rho_\mathrm{s}(E_\mathrm{F}-eV+\epsilon)\cdot\rho_\mathrm{t}(E_\mathrm{F}+\epsilon)\cdot|M|^2 \quad (2.3)$$

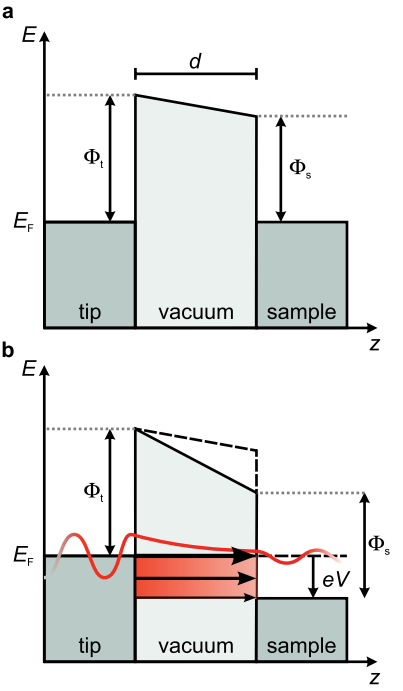


Figure 2.2 | Potential energy diagram. The schematic energy diagram of the tunnelling junction illustrates that the tip and the sample are separated by vacuum forming a potential barrier of thickness d. **a**, Without an applied bias voltage and at a temperature of 0 K, the states in tip and sample (dark grey areas) are filled up to the Fermi energy E_F. The work functions of tip and sample are abbreviated as Φ_t and Φ_s, respectively. **b**, Applying a bias voltage to the sample shifts the sample energy, tilting the potential landscape by $|eV|$. This opens up a window for electrons to tunnel from filled tip states into unoccupied sample states (wave function sketched in red). The horizontal arrows of varying thickness indicate the variations in tunnelling probability based on the different barrier heights experienced by electrons at different energy levels.

Spatial resolution in scanning tunnelling microscopy. If the bias voltage is kept constant while scanning over the sample, both changes in the LDOS of the sample and the wave-function overlap influence the measured tunnelling current. Using simplifications such as assuming a spherical s-type tip-wave function, low bias voltages, low temperatures, and dominating tunnelling contributions from electrons close to the Fermi energy of the tip, Tersoff and Hamann [Ter83, Ter85] and subsequently Lang *et al.* [Lan86] were able to show that the tunnelling current can be rewritten as:

$$I(V) \propto \int_0^{eV} d\epsilon \rho_s(E_F - eV + \epsilon) \cdot \rho_t(E_F + \epsilon) \cdot \exp(-2\kappa d) \tag{2.4}$$

with the exponential decay constant:

$$\kappa = \sqrt{\frac{2m_\mathrm{e}}{\hbar^2}\left(\frac{\Phi_\mathrm{t}+\Phi_\mathrm{s}}{2}\right)} \tag{2.5}$$

As the work functions of tip and sample enter the decay constant, STM operation possesses a strong material sensitivity. Moreover, due to the exponential dependence of the current on the separation d between tip and sample, STM is able to provide supreme sub-Ångström spatial resolution.

Energy resolution in scanning tunnelling spectroscopy. Keeping the tip stationary ($d = \text{const.}$), equation 2.4 can be simplified to:

$$I(V) \propto \int_0^{eV} \rho_\mathrm{s}(E_\mathrm{F} - eV + \epsilon)\rho_\mathrm{t}(E_\mathrm{F} + \epsilon)d\epsilon \tag{2.6}$$

Hence, measuring the tunnelling current as a function of the bias voltage provides insights into the electronic properties of the junction. This measurement scheme, called scanning tunnelling spectroscopy (STS), is able to resolve the local differential conductance of the tip-sample junction. Utilising a metallic tip with an approximately constant density of states, the LDOS of the sample can be extracted in STS as:

$$\frac{\mathrm{d}I}{\mathrm{d}V} \propto \rho_\mathrm{s}(E_\mathrm{F} - eV) \tag{2.7}$$

This spectroscopic capability of STM allows for the visualisation of how the energy levels of atoms, molecules, or crystals vary in real-space with atom-scale precision.

State-selective tunnelling in scanning tunnelling microscopy. Atoms and molecules constitute the building blocks of all matter, not only of any (inorganic) material in the inanimate world but also of all compounds and tissues relevant for chemistry, biology, and medicine. It is therefore of utmost interest to investigate the dynamics of small single molecules, oligomers, and more complex aggregates to shed light on the elementary processes happening all around us in everyday life. Specifically, one large and highly relevant field of interest centres around (bio-)chemical reactions. The examination of reactions by real-time videographic imaging and potentially even a targeted manipulation would provide tremendous potential for various

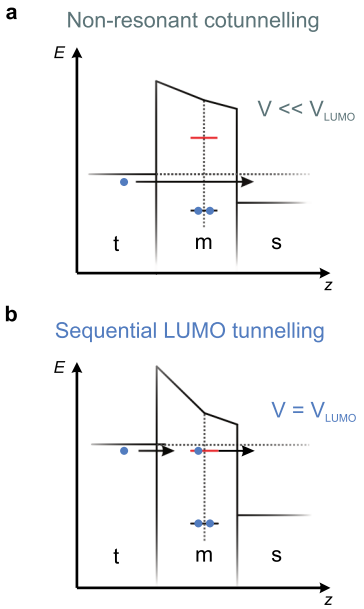


Figure 2.3 | Non-resonant vs. sequential tunnelling through a single molecular orbital. A schematic of the energy landscape of an STM tunnelling junction is depicted. A molecule (m) with filled (black line with two blue spheres) and unoccupied (bare red line) orbitals is positioned between the tip (t) and the substrate (s). **a**, By applying a small bias voltage V between the tip and the substrate, the tip potential is increased relative to the sample's Fermi energy. As a result, a net electron flow directly from the tip to the substrate is initiated (black arrow). **b**, At specific voltages, when the Fermi energy of the tip aligns with the energy of a molecular orbital (here, LUMO, red line), temporary charging of the molecule is possible, opening an additional two-step tunnelling channel (two black arrows) with a higher tunnelling probability.

applications. As a first step towards this goal, it is necessary to provide snapshots of the momentary structure of electronic orbitals. When operated under the suitable conditions, STM provides the possibility to map out the electron density of single molecules with Ångström precision.

Since STM relies on the non-zero conductance of its tunnelling junction to generate an image, most STM studies of individual molecules have thus far focused on molecules deposited on metals or semiconductors. This will inevitably lead to a perturbation of the electronic structure of the molecule by the presence of substrate electrons. To study the unperturbed electronic properties of molecules using STM, the sample must be electronically decoupled from the prerequisite conductive substrate. This is fulfilled by growing few monolayers of insulating materials – mostly salts such as sodium chloride (NaCl), potassium bromide (KBr), or silver chloride (AgCl) – in-situ on top of the metallic substrate prior to depositing the molecules [Rep05]. As a result, the molecular orbitals cannot hybridise with the metal and are

only minimally altered with respect to the free-space case, allowing for the study of the inherent electronic properties of the molecules. As these molecules exhibit sharply defined energy levels, different tunnelling situations must be considered. Figure 2.3 illustrates how non-resonant and sequential tunnelling through the lowest unoccupied molecular orbital (LUMO) differ from one another. Metallic tip (t) and substrate (s) are separated by tunnelling barriers in the form of the vacuum and the salt layers grown on the substrate. The molecule (m) adsorbed on top of the salt is marked by the horizontal lines indicating the HOMO in black and the LUMO in red. By applying a bias voltage to the tip, we elevate the Fermi energy of electrons therein, resulting in a predominantly unidirectional electron flow toward the sample. While $V \ll V_{\mathrm{LUMO}}$, electrons must pass the combined thick potential barrier and tunnel directly from tip to substrate (see Fig. 2.3a), bypassing the molecule in a so-called cotunnelling process. If the bias is large enough ($V = V_{\mathrm{LUMO}}$), a tunnelling channel opens such that electrons from the tip can charge the molecule. Subsequently, after a statistical lifetime of the charged state, the electron can tunnel through the thin insulating layer into the substrate. The sequential tunnelling current is considerably larger than the non-resonant cotunnelling due to the higher density of states and the effectively thinner as well as lower potential barriers. This type of level alignment between tip and molecular resonances can be attained by static or modulated voltage sources as well as ultrashort light pulses as is done in LW-STM.

2.2 Laser-assisted scanning tunnelling microscopy

In conventional STM, which will be referred to as DC-STM ("direct-current STM") in the remainder of this thesis, the temporal resolution is typically limited by the frequency of the voltage source. In all-electronic biasing applications, the length of the voltage pulses is determined by the bandwidth of the electronic circuity inside the STM, which currently allows for pulse durations in the nanosecond regime at best [Lot10]. Improvements beyond this limit can be accomplished by adding ultrafast photonics into the electronic circuitry in several different manners. For instance, photoconductive switches can be used to gate the STM circuitry with femtosecond near-infrared pulses, allowing for the generation of voltage pulses on the order of 650 fs [Wei93]. Alternatively, the voltage transients can be created directly on the

sample and transmitted to the STM junction by means of striplines [Nun93]. This so-called junction-mixing STM allows for the confinement of tunnelling currents within picosecond time frames. Apart from modulating the bias voltage with laser pulses, another possible approach makes use of direct illumination of the STM junction. As described by Hamers and Cahill [Ham90] or McEllistrem *et al.* [McE93], pulsed optical excitation of the junction can induce surface photovoltages that affect the local tunnelling barrier and thus the current on picosecond timescales. Specht *et al.* [Spe92] employed an STM tip as a localised probe to investigate resonantly excited surface plasmons, resulting in the creation of a near-field optical image with a resolution of 3 nm. Surface plasmon excitations have also been harnessed by Rücker *et al.* [Rüc92] to enhance the contrast in an STM image. Even the nanoscale dynamics of excitons in atomically-thin materials have been investigated on timescales of around 100 ps [Mog22].

However, the analysis of tunnelling currents obtained in STM junctions excited by light pulses in the visible regime is challenging. The interaction between light and the electronic systems of the tip and sample can give rise to effects such as local electron heating and multiphoton excitation, causing the systems to deviate from equilibrium on ultrashort timescales. Consequently, extracting pure dynamics of the sample in a photoexcited scanning tunnelling microscope is a complex task. Unlike in DC-STM, where the observed properties directly reflect the sample's characteristics, in a photoexcited STM, it becomes necessary to distinguish and separate the non-equilibrium dynamics (heating, multi-photon excitation, etc.) of the microscope and sample. Additionally, the presence of plasmonic resonances in metallic tips and substrates at optical frequencies poses challenges for efficient coupling of ultrashort pulses into the junction.

A different approach to improve the temporal resolution in STM relies on directly utilising the fastest voltage pulse nature has to offer, namely the carrier wave of light. The electric near field of a light pulse focused onto the STM junction may act as an ultrafast alternating bias voltage dropping across the gap. When properly tuned, this may drive electron tunnelling solely in the presence of the pulse or even on sub-cycle timescales – precisely the control mechanism desired – in lightwave STM.

In general, when light impinges on matter, due to its quantum mechanical nature, it can be described in the particle or wave picture. Electromagnetic radiation can

either be absorbed, leading to quantum transitions, or excite classical field-driven dynamics. The Keldysh parameter κ provides a measure of the balance between these two regimes [Kel65]:

$$\kappa = \sqrt{\frac{\varepsilon}{2U_\mathrm{P}}} = \frac{\omega\sqrt{2m_\mathrm{e}\varepsilon}}{eE} \tag{2.8}$$

wherein ε denotes the excitation or ionisation energy and:

$$U_\mathrm{P} = \frac{e^2E^2}{4m_\mathrm{e}\omega^2} \tag{2.9}$$

the ponderomotive energy, which describes the kinetic energy of a free electron with mass m_e after acceleration by one optical cycle of a driving field E with angular frequency ω. At high optical frequencies ω, the photon energy $\hbar\omega$ is large and light-matter interaction can be described as (multi-)photon processes. In this regime, where according to equation 2.8 the Keldysh parameter κ is larger than one, the particle picture is most suitable. For lower frequencies and/or very large electric fields ($\kappa < 1$), the wave character of light dominates, which can accelerate charges on sub-cycle timescales. The carrier wave of light can thus be utilised as an ultrafast bias pulse driving currents across the tunnelling junction. High-intensity light in the terahertz spectral range is especially suitable for this application, as the achievable temporal resolution, determined by the oscillation period and the pulse duration, is still on the scale of elementary excitations of matter such as phonons, plasmons, or magnons but at the same time κ remains well below one (high fields and simultaneous low photon energies).

In 2013, Cocker *et al.* [Coc13] demonstrated for the first time that when terahertz pulses are precisely focused onto the tunnelling junction, their oscillating carrier wave is indeed able to drive tunnelling within its pulse duration. They successfully used this technique to visualise carrier dynamics in semiconductor nanodots with a combined spatial and temporal resolution of 2 nm and 500 fs, respectively.

2.3 Ultrafast state-selective lightwave-driven scanning tunnelling microscopy

Bringing LW-STM into the thrilling single-electron and strictly sub-cycle tunnelling regime, LW-STM with unprecedented speed and control was demonstrated in 2016 [Coc16]. Within a collaboration between the groups of Prof. Dr. Rupert Huber and Prof. Dr. Jascha Repp, a custom-tailored optical setup and a cryogenic low-noise scanning tunnelling microscope were developed. This LW-STM made tunnelling of single electrons into a select molecular orbital with combined atomic spatial and femtosecond temporal precision a reality. The exceptional temporal resolution is achieved by confining the tunnelling window into a sharp electronic resonance to timescales shorter than the pulse duration. The schematic in Fig. 2.4 illustrates the principle of state-selective LW-driven tunnelling. The terahertz near field acts as an ultrafast tip-confined voltage pulse (blue time trace). An exemplary molecule with the LUMO resonance (orange) at $V_{\mathrm{LUMO}} = 1.2\,V$ is placed in the tunnelling junction. Only bias voltages exceeding a certain threshold voltage (dashed line, $\frac{1}{e^2}$ of the LUMO peak height) and thus only a fraction of the transient (filled blue region) leads to a significant tunnelling probability and thus contributes to the sequential tunnelling current. This window is considerably shorter than the pulse duration, allowing for the confinement of tunnelling events to timescales on the order of 100 fs. Moreover, since this process is state-selective, it offers unprecedented opportunities for ultrafast atomic-scale imaging and spectroscopic analysis of atoms, molecules, crystals, and more.

The full significance of this improved lightwave-controlled STM technique lies in its capability to record sub-picosecond dynamics of single molecules in real space using a pump-probe scheme. A first terahertz pulse (pump) can be utilised to either extract an electron from the HOMO of a molecule or inject one into its LUMO depending on the polarity of the voltage pulse. Of course, the same applies to solids and the corresponding valence/conductance bands.

In the groundbreaking work by Cocker *et al.* [Coc16], the terahertz pulse was used to transiently charge a pentacene molecule adsorbed on a monolayer NaCl substrate positively by removing a single electron from the HOMO. The charging process induces a sudden change in the electronic state and the Coulomb and van der Waals

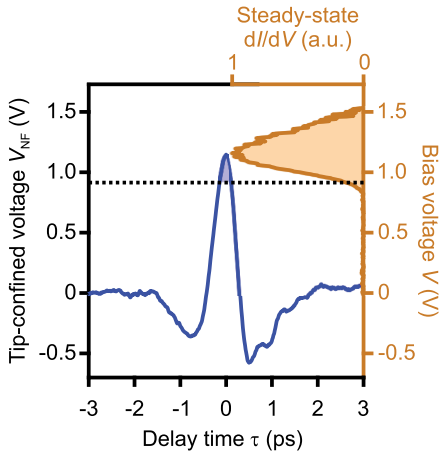


Figure 2.4 | Ultrashort tunnelling window in LW-STM. A terahertz pulse can act as an ultrafast transient voltage, which opens a sequential tunnelling channel through the LUMO (orange shaded region, peak at 1.2 V). Only instantaneous voltages exceeding a certain threshold (here the voltage where the conductance exceeds $\frac{1}{e^2}$ of the LUMO peak height) contribute considerable tunnelling currents, which effectively confines the tunnelling window to a fraction of the pulse duration (blue shaded region).

interactions between the molecule and the substrate are modified by the additional charge. The sudden change of the potential landscape leads to a vertical displacement of the molecule [Par00]. The molecule rapidly discharges through the NaCl monolayer on the timescale of 100 fs [Ste14], leaving it electronically neutral but still oscillating vertically. This displacement initiates a vertical vibration of the molecule, resulting in periodic changes in the tip-sample separation. The variation in tip-sample distance was probed by a second terahertz pulse, measuring the tunnelling probability as a function of the delay time between the two pulses. By repeating the pump-probe measurement for different tip positions and delay times, the first molecular movie capturing the oscillation of a single molecular orbital could be recorded.

As this state-of-the-art technique was used for a large fraction of the work presented in this thesis, the following subsections provide a brief summary of the utilised setup and its relevant components. For a more detailed description, the reader may be referred to the dissertation of Dr. Dominik Peller [Pel20a].

Generation and manipulation of sub-cycle terahertz pulses. To drive measurable tunnelling currents across the junction on ultrashort timescales, the laser pulses injected into the STM must fulfil several requirements. First, as the employed electronics are considerably slower than the laser repetition rate, charge accumulation

must occur before a measurable electronic signal arises. More precisely, this means that subsequent laser shots must drive electrons in the same direction. Therefore, the carrier envelope phase (CEP) ϕ_{CE} of the terahertz pulse train must be constant. The CEP gives the temporal evolution of the shift between the maximum of the intensity envelope and the carrier wave of the pulse (for details see Fig. 4.16). For example, $\phi_{\mathrm{CE}} = 0$ corresponds to a cosine-like oscillating electric field underneath the envelope, whereas $\phi_{\mathrm{CE}} = \frac{\pi}{2}$ indicates a sine shape and $\phi_{\mathrm{CE}} = -\pi$ corresponds to a negative cosine (see Fig. 4.16a). Second, as mentioned above, the Keldysh parameter κ should remain below one to ensure that the classical picture of electromagnetic radiation is applicable and the carrier wave of light can be directly used as an ultrafast bias. Third, the peak voltage in the near field must be comparable to energy levels within the sample to allow for sequential tunnelling. Most resonances, such as valence/conduction band onsets or molecular orbitals (HOMO/LUMO) lie on the order of 1 eV from the Fermi energy. With a typical tip-sample distance of a few Ångström, near fields of $E \approx 0.1\,\mathrm{V/Å}$ are necessary.

These requirements can be satisfied in our custom-tailored optical setup. A high-stability ytterbium-doped potassium gadolinium tungstate (Yb:KGW) femtosecond regenerative amplifier (Light Conversion, Pharos) acts as the source of all radiation necessary in our LW-STM (see Fig. 2.5). The commercially available laser amplifier outputs near-infrared radiation centred around 1030 nm with a pulse duration of 240 fs (full width a half maximum of the intensity envelope) at an adjustable repetition rate between 610 kHz and single shot operation (in steps of integer dividers of 610 kHz). The output-pulse energy of 16 µJ is split up into two parts by means of a 90/10 beamsplitter. The majority of the pulse energy is utilised for the generation of CEP-stable terahertz pulses by harnessing the nonlinear optical process of optical rectification within a lithium niobate ($LiNbO_3$) crystal (see Fig. 2.5b). Details on nonlinear optical effects are included in the explanations following equation 2.10. In an elaborate pulse-shaping setup (Fig. 2.5c), out of each terahertz pulse a pair of terahertz transients can be created. Using a germanium wafer, each incident pulse is split into two approximately equally strong parts where in each arm the field strength can be tuned by means of computer-controlled wire-grid polarisers. The pulses in each pair can be delayed relative to each other using a mechanical delay stage allowing for THz pump – THz probe measurements. To obtain ultrafast access

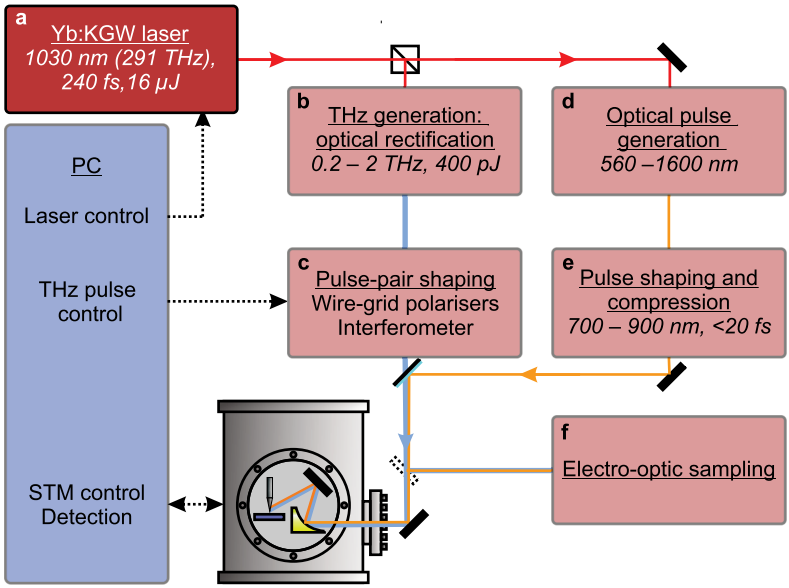


Figure 2.5 | Schematic of the LW-STM setup. CEP-stable terahertz and ultrashort optical pulses are generated and applied in LW-STM to drive ultrafast currents and excite the sample, respectively. **a**, A commercial Yb:KGW solid state laser emits near-infrared pulses centred around 1030 nm with a pulse duration of 240 fs. Multiple nonlinear optical processes are employed to convert the near-infrared radiation to the spectral ranges necessary for the operation of our LW-STM. **b**, Exploiting optical rectification in the nonlinear crystal lithium niobate, sub-cycle terahertz pulses spanning frequencies of 0−2 THz are generated. **c**, The terahertz pulses are split into pairs of pulses in an interferometer, which can be delayed with respect to each other and individually tailored in field strength as well as polarisation using wire-grid polarisers. The pairs of terahertz pulses are subsequently injected into the STM vacuum chamber and focused onto the STM tip, where they act as ultrafast bias pulses. **d**, To generate ultrashort laser pulses of durations less than 20 fs, a small fraction of the near-infrared pulse energy is split off before the optical rectification process to generate a supercontinuum spanning wavelengths of 560−1600 nm. **e**, The broadband spectrum can be shaped and compressed to provide ultrashort pulses with centre wavelengths between 560 nm and 980 nm. **f**, The visible pulses can be either used to electro-optically sample the terahertz pulses in the time-domain or alternatively be focused into the tip-sample junction to optically excite the sample. Terahertz pulse shaping, data acquisition, as well as laser and STM operation are computer-controlled remotely.

to both filled and unoccupied quantum states in the sample, it is necessary that both positive and negative transient voltages can be applied across the tip-sample junction. This is achieved by changing the polarity of the terahertz pulses. To flip the field direction of both pulses, we make use of the phase shift upon reflection off a metallic surface. The reflected light exhibits an additional phase shift of π with respect to the incident pulse. By increasing the number of reflections, experienced by the pulses propagating from generation to experiment, by an uneven number, the polarity can be flipped. Subsequently, the terahertz pulse pairs can be coupled into the STM through sapphire windows placed in the vacuum chamber as well as the helium and nitrogen shields. A 90°-off-axis parabolic mirror focuses the laser radiation into the tip-sample junction. To allow for steering of the focal spot, the orientation of an additional planar gold mirror placed between the parabolic mirror and the sample is addressable using piezo-electric motors. A simplified illustration of the focusing geometry id depicted in the bottom left corner of Fig. 2.5. To adjust the position of the STM tip in all three dimensions, a Besocke-type scanning geometry [Bes87] was chosen. In contrast to the original publication by Besocke *et al.*, in our case the sample is stationary and the tip is scanned. To accommodate this, the tip is attached to a rampring that comes to rest on top of three piezo stacks and can be moved in all directions. The sample is inserted below the tip and a bias voltage can be applied between tip and sample leading to a tunnelling current. To ensure a high signal-to-noise ratio, the tunnelling current is amplified directly on the scan head. By placing the amplifier at liquid helium temperature and in close proximity to the junction, electrical noise is minimised. The signal is then additionally amplified at room temperature and sent to a controller where the final signal is eventually recorded for further analyses on a computer.

With the remaining 1.6 µJ of near-infrared pulse energy, which was not used for terahertz generation, ultrashort pulses in the visible spectral region are created. Focusing near-infrared pulses under suitable conditions into a nonlinear crystal, a broadband supercontinuum spectrum spanning wavelengths of 560 nm to 1600 nm, is generated within a sapphire window with a thickness of 5 mm (see Fig. 2.5d). By correct spectral shaping of the spectrum and dispersion compensation using a prism compressor setup, pulses centred around a wavelength of 800 nm with a pulse duration of 15 fs can be established. These pulses are then used for either optical excitation of the sample

or field-resolved detection of the terahertz waveform in an electro-optic sampling setup (see Fig. 2.5f and sec. 2.3). Details on supercontinuum generation and prism compressor setups can be found in section 4.5.

Nonlinear polarisation. As nonlinear optics lies at the heart of numerous processes described within this thesis, it shall be briefly introduced in the following. By exploiting the nonlinear polarisation response of matter subjected to intense laser fields, a plethora of interesting conversion processes can occur. In the low-field regime, electrons and (to a much smaller extent) ions within the illuminated material follow the external driving field E linearly, meaning that the polarisation response P is directly proportional to the electric field:

$$P(E) = \varepsilon_0 \chi^{(1)} E$$

The natural constant ε_0 denotes the vacuum permittivity and $\chi^{(1)}$ stands for the material-specific first-order electric susceptibility. For larger field strengths, such as under intense laser radiation, this proportionality is superseded by a more complex relation involving non-negligible higher-order terms. The polarisation P can be expanded in a Taylor series as:

$$P(E) = \epsilon_0 \Big(\chi^{(1)} E + \chi^{(2)} E^2 + \chi^{(3)} E^3 + ... \Big) \tag{2.10}$$

with the material-dependent parameters of first-order, second-order and third-order susceptibilities, $\chi^{(1)}, \chi^{(2)}$, and $\chi^{(3)}$, respectively. The quadratic term entails a mixing of frequencies of the incident oscillating driving fields. According to Maxwell's equations, the oscillating polarisation/dipole leads to emission of radiation that can comprise frequencies that were not present in the incident spectra. For a non-centrosymmetric crystal, the first non-vanishing higher-order polarisation is the second-order contribution leading to effects such as second-harmonic generation (SHG), difference-frequency generation (DFG), and sum-frequency generation (SFG). The former two are of particular interest within the scope of this thesis and will be further elaborated on in this chapter and in section 4.5.

The terms proportional to the third-order nonlinear susceptibility $\chi^{(3)}$ induce further intriguing effects that are exploited within this work for self-focusing (see sec. 4.2)

and spectral broadening (see sec. 4.2).
The second-order nonlinear process of difference-frequency generation can be explained by the mixing of two optical frequencies ω_1 (the so-called "pump" frequency) and ω_2 (the "signal" frequency), leading to the emission of light at $\omega_3 = \omega_1 - \omega_2$ (the "idler" frequency). By using DFG, completely new frequencies can be generated and existing spectral components can be amplified (see sec. 4.5). When a broadband spectrum of near-infrared frequencies is used for DFG, pulses in the terahertz spectral region can be generated. The maximum terahertz frequency is ultimately limited by the bandwidth of the near-infrared pulses used. Importantly, intra-pulse difference-frequency generation, also known as optical rectification (OR), creates inherently CEP-stable terahertz pulses. This pulse-to-pulse stability of the polarity and waveform makes them applicable as ultrafast transient bias in the fields of LW-STM or lightwave electronics [Sch14, Lan18].
The near-infrared output pulses of the solid-state laser amplifier, used within my work, feature a bandwidth of 2.6 THz (FWHM). Utilising tilted pulse-front optical rectification in a suitable $LiNbO_3$ nonlinear optical crystal [Heb02, Hir11], this enables the generation of terahertz pulses with spectral components between 0.2 THz and 2 THz ($\frac{1}{e}$-width). Immediately after generation, the terahertz pulses possess a pulse energy of approximately 400 pJ, resulting in peak field strengths of 5 kV/cm in a diffraction-limited far-field focus.
Femtosecond snapshot of a molecular orbital. To test whether light-wave driven tunnelling through a single molecule is possible with this setup, my colleagues deposited pentacene molecules on an insulating sodium chloride sheet. Using this STM and optical setup, they were able record the first femtosecond snapshot of a molecular orbital [Coc16]. Figure 2.6 depicts the HOMO of a pentacene molecule adsorbed on sodium chloride recorded in two different modes of operation. The steady-state STM image (see Fig. 2.6a) reveals the rich sub-molecular structure of the orbital with five nodal planes. Clearly, the lightwave-driven snapshot (Fig. 2.6b) reproduces these atomic-scale details proving its capabilities to image individual molecular orbitals by lightwave-controlled tunnelling of single electrons. The question arises how exactly the electron tunnelling was induced. In general, the lightwave-driven current can have two origins. On the one hand, it can originate from time-integrated rectification of the voltage pulse at a nonlinearity within the current-voltage characteristic

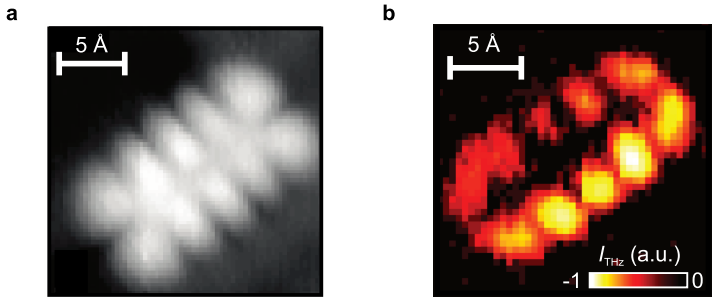


Figure 2.6 | DC imaging vs. LW-driven imaging. a, The HOMO of pentacene is recorded by means of steady-state constant-current STM imaging. The greyscale ranges from 0 to 2 Å. Scan paramters: bias voltage $V = -1.7\,\mathrm{V}$ and setpoint tunnelling current $I = 0.83\,\mathrm{pA}$. **b**, The terahertz-driven tunnelling image reproduces the orbital shape resolved with DC biasing (panel a) for a peak voltage of $V_{\mathrm{LW}}^{\mathrm{peak}} \approx -2.05\,\mathrm{V}$. Scan parameters: $V = 2.2\,\mathrm{mV}$ and $I = 2.2\,\mathrm{pA}$.

of the sample [Coc13]. On the other hand, if the voltage pulse exhibits a suitable waveform and peak voltage, unidirectional state-selective tunnelling current driven within a fraction of the pulse duration can occur. From the image in Fig. 2.6b alone, this cannot be inferred. Instead, the full temporal evolution of the terahertz pulses injected into the STM junction must be recorded in order to disclose whether the tunnelling window is indeed confined to timescales of 100 fs.

Field-resolved detection. In order to accurately characterise the waveform of the generated terahertz pulses, a field-resolved detection scheme utilising electro-optic sampling (EOS) is used. Notably, EOS serves as an "oscilloscope for light" by tracing the oscillating carrier wave with sub-cycle temporal resolution. The detection mechanism relies on the linear electro-optic effect, also known as Pockels effect [Boy08], where an electric field applied to a suitable material induces a field-dependent birefringence. An ultrashort gate pulse passing through the crystal experiences the birefringence, and the electric field strength is encoded onto the polarisation state of the gate pulse.

To record the evolution of the electric field of the the terahertz pulses, they are superimposed with the gate pulses and focused into an electro-optic detection crystal

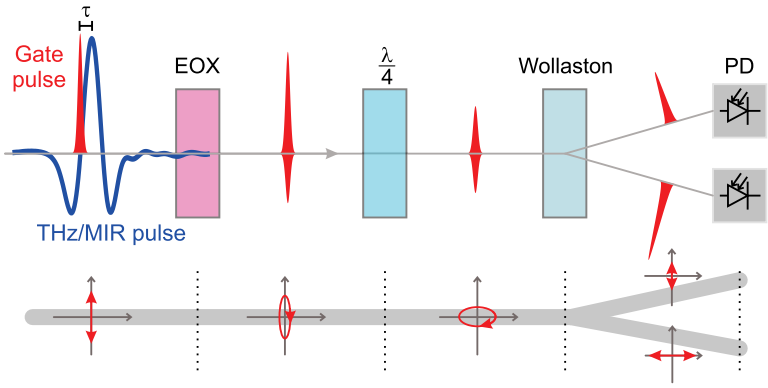


Figure 2.7 | Experimental setup of electro-optic sampling. A (multi-)terahertz pulse (dark blue curve) and a short linearly (here vertically) polarised gate pulse (red) are spatially overlapped in a suitable nonlinear optical crystal (EOX) with a variable delay time τ. The instantaneous electric field of the (multi-)terahertz pulse induces birefringence in the EOX by means of the Pockels effect, which imprints itself on the gate pulse by rotating its polarisation. The now elliptically polarised gate pulse passes through a quarter-wave plate ($\frac{\lambda}{4}$) and is split into its two perpendicular polarisation components by means of a Wollaston prim. These are subsequently detected by a pair of photodiodes (PD). The lower row depicts the polarisation state of the near-infrared pulse after each optical component.

(see Fig. 2.7, EOX) – in our case zinc telluride. Importantly, the oscillation period of the terahertz pulses is so slow compared to the pulse duration of the gate pulses that it can be approximated as quasi-static [Mit03]. The electric field of the terahertz pulse thus induces a quasi-static birefringence within the electro-optic crystal, which modifies the linearly polarised gate light (blue) by phase retardation and polarisation rotation, resulting in elliptically polarised light. After the interaction of the two pulses, the near-infrared gate pulse passes a quarter-wave plate ($\frac{\lambda}{4}$) and a Wollaston prism, and is focused into two photodiodes. The wave plate is rotated such that in the absence of the terahertz pulse, the linearly polarised incident gate pulse is transformed into purely circularly polarised light. As the Wollaston prism separates

horizontal and vertical polarisation components, this leads to equally strong near-infrared induced signals S in both photodiodes. The detected signal in the so-called balanced photodiode circuit is therefore zero, $\Delta S = 0$. However, when a terahertz field is present, the gate pulse becomes elliptically polarised after passing the electro-optic crystal and therefore is not circularly polarised after the $\frac{\lambda}{4}$-plate. The resulting small signal imbalance between the two photodiodes ΔS is linear in the instantaneous electric field of the terahertz pulse. Varying the delay time between gate and terahertz pulse, the full electric field waveform can be reconstructed stroboscopically with a time resolution limited only by the gate pulse duration of 15 fs. By making use of this sensitive scheme in combination with lock-in amplification, signals down to the shot noise limit are detectable.

The generated terahertz pulses within our setup are displayed in Fig. 2.8a. The recorded cosine-like half-cycle waveform features a large asymmetry ratio (relative strength of largest positive and negative excursion) of more than two. The spectral intensity (see Fig. 2.8b, dark blue), obtained by Fourier transform of the waveform, spans from 0.2 THz to 2 THz ($\frac{1}{e}$-width). Together with the flat spectral phase (Fig. 2.8b, light blue), this supports the cosine-like half-cycle waveform.

The EOS scheme can also be used for higher frequency radiation in the mid-infrared regime [Hub00a, Küb04], which becomes crucial in later chapters of this thesis (sec. 4.3). Recent studies suggest that even near-infrared transients can be electro-optically sampled utilising 4 fs visible gate pulses [Kei16].

Notably, EOS is only applicable if the temporal evolution of the electric field under the envelope is identical for subsequent pulses. Thus, the CEP-stability of the sub-cycle terahertz waveforms is confirmed and the lightwave-driven currents leading to the image in Fig. 2.6b can indeed be attributed to state-selective tunnelling out of the HOMO of pentacene into the tip when the peak field is tuned to the molecular resonance.

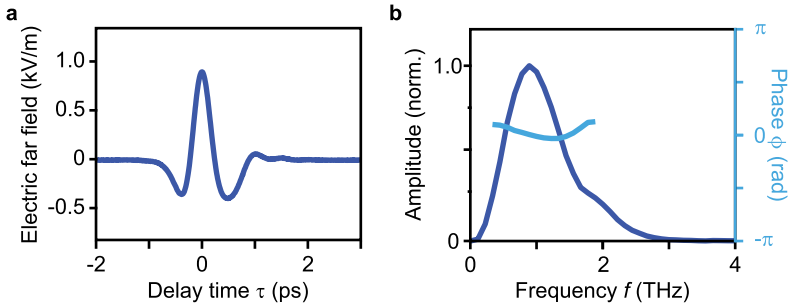


Figure 2.8 | Electro-optically detected terahertz waveform. a, The terahertz waveform detected using electro-optic sampling shows an almost perfect cosine shape under the carrier envelope with half an oscillation cycle within the FWHM and a large asymmetry ratio (relative strength of largest positive and negative excursion) of more than two. **b**, The terahertz spectrum generated by optical rectification in $LiNbO_3$ spans the region from around 0.2 THz to 2 THz with its peak at 0.9 THz. The spectral phase is close to 0 rad and only shows a slight positive curvature.

2.4 Shot-to-shot detection of single-molecule switching statistics

The seminal work on LW-STM with atomic resolution by Cocker *et al.* [Coc16] showcased that apart from the imaging capabilities, terahertz-driven STM can be used for the visualisation of dynamics. They presented the triggering and observation of vertical vibrations of a single pentacene molecule induced by charging through LW-driven single-electron tunnelling. Apart from vertical vibrations, the realm of molecular kinetics encompasses numerous more intriguing phenomena, including changes in atomic composition, and inter- and intramolecular vibrations, which merit further investigation. Especially, studying irreversible kinetics or atom-scale dynamics where the final state differs from the initial state would be highly compelling. By visualising chemical reactions and gaining control over them, our understanding of the microscopic world can be greatly enhanced, empowering us to manipulate reactions according to our desired outcomes. To embark on this exciting journey, we initiated a crucial first step by applying LW-STM to investigate a simple yet significant

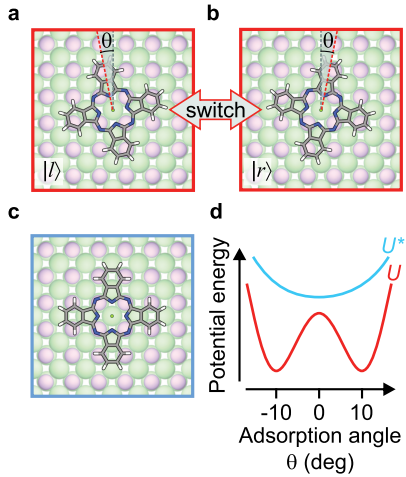


Figure 2.9 | Adsorption geometry and potentials of the molecular switch. **a**+**b**, The neutral MgPc molecule adsorbs in one of the two energetically degenerate ground states where the molecule's axes are rotated by $\theta = \pm 10°$ with respect to the salt lattice. The two states are labelled $|l\rangle$ (panel a) and $|r\rangle$ (panel b). **c**, The singly-charged MgPc molecule prefers to align with the underlying salt. **d**, The potential landscape of the neutral molecule U is described by a double-well potential with an energy barrier at 0° (red curve). The charged molecule, however, is subject to a parabolic potential energy landscape U^* centred at 0° (light blue curve).

reaction/transformation process involving a molecular switch. Magnesium phthalocyanine (MgPc) is a tetrapyrrole consisting of four isoindole-type units, where the connecting carbon atoms in the macrocycle are replaced by nitrogen atoms and a magnesium atom is bound at the centre (see Fig. 2.9). When MgPc is evaporated onto a multilayer NaCl substrate, three dominant adsorption geometries can be observed [Pat19]. The magnesium atom of the organic complex preferably comes to rest directly on top of a chlorine ion and the molecular axes are either oriented at an angle of $\theta = +10°$ (labelled $|r\rangle$) or $\theta = -10°$ ($|l\rangle$) with respect to the underlying cubic salt lattice (see Fig. 2.9a and b). A third configuration arises when MgPc adsorbs on top of a sodium ion. Other adsorption geometries occur at impurities, defects, and step edges of the substrate, but these instances are less common. For most MgPc molecules, which are localised on top of a chlorine ion, the adsorption potential in the ground state can be described as a double-well potential U with minima at $\theta = \pm 10°$ (Fig. 2.9d, red curve). However, the charged state (Fig. 2.9c) has a completely different preferred orientation in a potential U^* (Fig. 2.9d, blue curve) with a single minimum due to the altered Coulomb and van der Waals interactions. Therefore, by injecting a single electron into the LUMO, the molecule starts to rotate towards its new equilibrium position ($\theta = 0°$). Depending on the statistical time

spent in the charged state, which is governed by the tunnelling probability of the electron back into the tip or through the NaCl layers into the metallic substrate [Ste14], the molecule has had enough time to cross the zero crossing $\theta = 0°$ or not, before dropping back into the potential U. Thus, after discharging of the LUMO, there is a non-zero probability that the molecule ends up on the opposite side to where it started – i.e., it switched. This process is re-instigated after each electron injection, possibly switching the molecule back and forth between the different ground states. It is to be noted, that cases where the molecule oscillates across the zero-crossing twice are negligible due to the short lifetimes and strong damping in the charged potential.

In conventional STM imaging, when $V_{\mathrm{DC}} > V_{\mathrm{LUMO}}$, continuous electron injection causes constant switching back and forth. Therefore, a superposition of the two configurations is detected. With ultrashort pulses, each injecting only a single electron into the LUMO, switching back and forth is possible in a much more controlled fashion.

By taking low-bias constant-height scans, the instantaneous orientation of the molecule can be visualised. Panels a and b in Fig. 2.10 depict the same molecule, in the left $|l\rangle$ and in the right $|r\rangle$ adsorption minimum, respectively. In these $20\,\text{Å} \times 20\,\text{Å}$ images recorded in constant-height mode at a DC bias of 6 mV, the NaCl lattice axes are aligned with the edges of the frames. A comparison of the two depicted configurations reveals two clearly different orientations of the four isoindole units of the MgPc molecule.

Ultrafast action spectroscopy. Distinguishing between the two states is even possible without scanning across the whole molecule: By placing the STM tip close to the edge of the molecule, for example at the cross-hair position marked in Fig. 2.10a and b, the two states can be discriminated by the non-resonant in-gap current. Despite of not charging the molecule, the presence of a molecular orbital under the tip greatly influences the direct tip-substrate tunnelling current. As a result, at this particular tip position, the in-gap current $I_{|l\rangle}$ for the left configuration is larger than $I_{|r\rangle}$ for the right configuration (Fig. 2.10b, y-axis). Since the switching process is too fast for the electronics to record, switching between the two configurations is detectable as sharp step-like features in the in-gap detection current I_{det}. As switching is only possible by charging the LUMO with a sufficiently large bias voltage and under the

condition $V_{\mathrm{DC}} = 6\,\mathrm{mV} \ll V_{\mathrm{LUMO}}$, each switching event can only be caused by a sufficiently strong terahertz pulse. Thus, every suitable laser shot leads to a boolean value indicating whether the molecule either has or has not switched configuration. With the known laser repetition rate, from the data in such switching time traces the overall switching statistics of the MgPc molecule can be extracted as a series of boolean values. The enlarged section in the right upper corner of Fig. 2.10b shows ten laser shots and the resulting three switching events. The mean switching probability from, for example, left to right configuration can be calculated as the ratio of the number of switching events $|l\rangle \rightarrow |r\rangle$ divided by the number of laser pulses that impinged on the molecule in its $|l\rangle$ state. The switching probability in the other direction can be acquired accordingly.
This detection scheme builds upon action spectroscopy, as introduced independently by Böckmann *et al.* [Böc19] in 2019, and further enhances it by incorporating precisely controlled single-electron injection. By intentionally constraining these tunnelling events to timescales shorter than a single cycle of the applied field, we refer to this novel technique as “ultrafast action spectroscopy”.

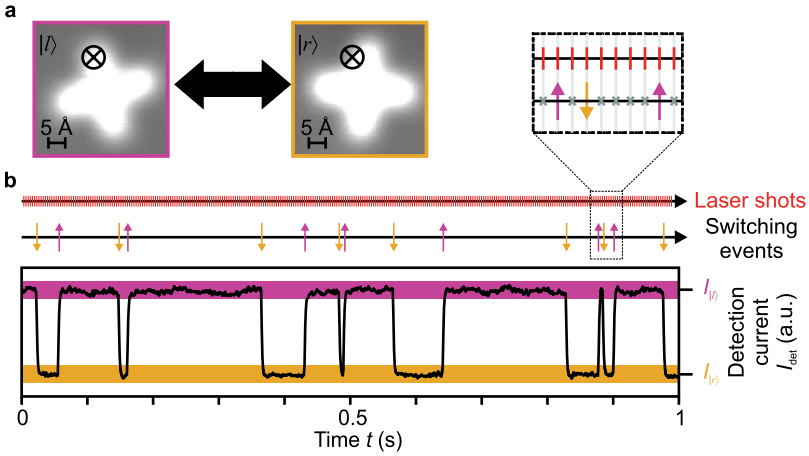


Figure 2.10 | Extraction of the switching statistics. a, Constant-height STM images reveal the orientation of the bistable molecule. The uncharged MgPc molecule can lie in either the left (pink frame) or right (orange frame) rotated configuration with respect to the underlying salt lattice. Voltage setpoint: $V = 6\,\mathrm{mV}$, linear greyscale $0{-}2\,\mathrm{pA}$. **b**, Injecting ultrashort laser pulses (red lines on the timeline) that can transiently charge the molecule leads to repeated toggling between the two stable ground states. Switching events from left to right and right to left are indicated by orange and purple arrows, respectively. The enlarged section in the top-right corner shows ten laser pulses and the resulting three switching events. By placing the STM tip at the cross-hair position marked in panel a and applying a weak DC bias ($V = 6\,\mathrm{mV}$), the molecule's configuration at each instance in time can be determined. Each switching event leads to a step-like change of the detection current I_{det}. Using this technique, which we coin ultrafast action spectroscopy, the outcome of every single laser shot can be evaluated, allowing for the extraction of the molecule's switching statistics.

Spatial variation of switching statistics. In order to derive information about local differences in switching behaviour, switching time traces were recorded for multiple tip positions. Figure 2.11a shows the mean position-dependent switching probability $|l\rangle \rightarrow |r\rangle$. Interestingly, a rich sub-molecular structure with locally increased or reduced switching probability emerges. Moreover, the observable switching rate pattern substantially differs from the pattern obtained for the $|r\rangle \rightarrow |l\rangle$ case (see Suppl. Fig. A.1.) This difference raises the question to which extent the spatial variations either stem from position-dependent charging or position-dependent switching probability once charged. To determine the dominant effect, we simulated the charging probability of the molecule in its left configuration $|l\rangle$. To this end, the orbital structure of the LUMOs was calculated using density-functional theory by our collaboration partner Nikolaj Moll. The overlap integral of a spherical tip-wave function with the numerically determined LUMO orbitals was computed for various tip positions utilising the Tersoff-Hamann approach (see eq. 2.4). A comparison of this map with the experimentally observed switching map reveals clear analogies (see Fig. 2.11b). Both maps feature eight regions of higher charging/switching probability and the same nodal planes. Even the slight asymmetry between lobes is reproduced and can be attributed to an internal torsion of the molecule (see Suppl. Fig. A.5). This leads to a smaller tip-sample distance and thus higher charging/switching probabilities for every other lobe. In conclusion, the effect dominating the spatial variations in the experiment is the position-dependent charging probability which differs on sub-Ångström length scales. The probability for an injected electron to trigger a switching event is almost position-independent. Hence, the overall position-dependent switching probability p can, to first approximation, be written as a product:

$$p(\boldsymbol{r}) = T(\boldsymbol{r}) \cdot \alpha \tag{2.11}$$

with the position-dependent charging probability $T(\boldsymbol{r})$ of the LUMO and the constant switching probability once an electron has been placed in the LUMO, α. By comparing the switching rate with the LW-driven tunnelling current, we can determine α to lie in the range of 18–24% when the molecule is adsorbed on an insulating sheet of bilayer NaCl. For a larger number of NaCl layers the statistical lifetime in the charged state increases which would lead to an expected increase of α. Indeed,

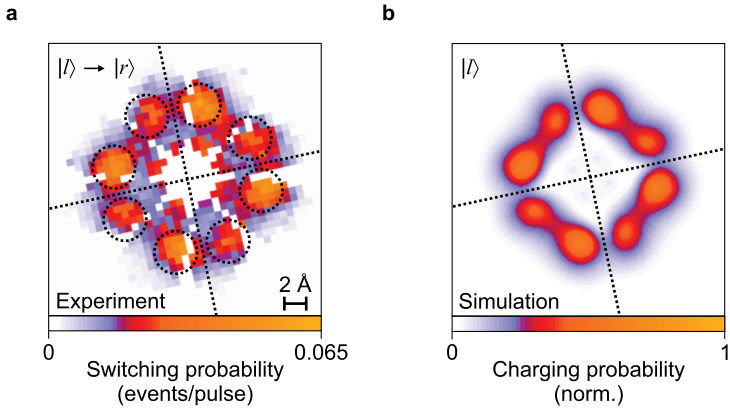


Figure 2.11 | Experimental spatial dependence of switching probability and simulated charging probability. a, The experimentally determined spatial map of the local switching probability for switching from left to right ($|l\rangle \rightarrow |r\rangle$) is depicted. Pixels where no switching could be detected are coloured in white. Eight regions of increased switching rate (dashed circles), of which every other is more prominent, are clearly visible. The symmetry axes of the molecule in its left configuration are illustrated with dashed lines. **b**, The charging probability is simulated by calculating the overlap integral of an s-type tip-wave function with the LUMO density of MgPc. The map obtained by repeating the calculation for different tip positions reveals eight areas of increased charging probability. Even the asymmetry in the brightness of these areas, as seen in the experimental data, is reproduced.

repeating the same experiments on a trilayer of NaCl a value for α of around 32% can be extracted.

Importantly, depending on the time the molecule spends in its charged state, the molecule has more or less time to rotate in its new potential. Hence, α is governed by the statistical lifetime of the charged state, which can be tuned by changing the number of NaCl layers between molecule and metallic substrate. Using a microscopy combining AFM and STM, Steurer *et al.* [Ste14] have identified the charged-state lifetime for various NaCl film thicknesses ranging from 5-11 monolayers. By extrapolating their data to the bi- and trilayer regime, the electron can be estimated to remain in the LUMO for several hundred femtoseconds or several picoseconds, re-

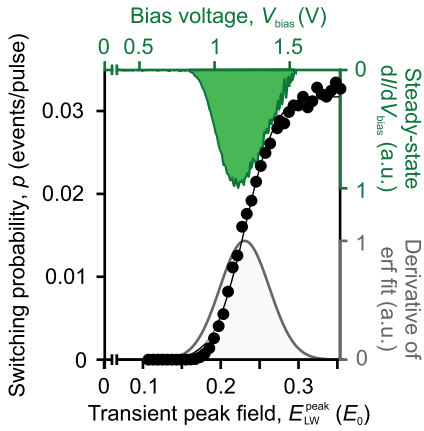


Figure 2.12 | Switching onset and voltage gauge. The switching probability p (black data points) nonlinearly depends on the peak electric field $E_{\mathrm{LW}}^{\mathrm{peak}}$ (in units of the unattenuated far field E_0). The derivative of the error-function fit to $p(E_{\mathrm{LW}}^{\mathrm{peak}})$ (black line), $\mathrm{d}p/\mathrm{d}E_{\mathrm{LW}}^{\mathrm{peak}}$ (grey curve and shaded area), reveals a Gaussian shape centred at $0.23\,E_0$. From DC spectroscopy (green curve and shaded area), the LUMO resonance of the molecule is determined to lie at 1.2 V. As the molecule must be charged for a switching event to occur, this gives us a quantitative voltage gauge allowing for the conversion from the switching probability to the tip-confined peak voltage.

spectively. All data presented up to this point were recorded on bilayer NaCl. For the remainder of this thesis we aimed towards higher switching yields, which was accomplished by increasing the insulator thickness. If not stated otherwise, the figures in this thesis from this point onwards show data from three layers of NaCl as insulating sheet.

Voltage-dependent switching rate. The extraction of path-selective statistics of the single-pulse switching probability with shot-to-shot fidelity and sub-Ångström precision now allows us to not only investigate the intrinsic switching statistics of a single molecular rotor but also to explore how it may be influenced on demand. Furthermore, the molecular switch can act as a local voltage gauge as is closely explained in the following section.

We have found that the switching probability does not only depend on the exact tip position, but the chance for the molecule to switch also critically hinges on the bias applied across the junction and thus the field strength of the terahertz pulses. The black data points in Fig. 2.12 show the influence of the far-field transient peak field $E_{\mathrm{LW}}^{\mathrm{peak}}$ in units of the unattenuated field strength E_0 on the mean switching probability $p(E_{\mathrm{LW}}^{\mathrm{peak}})$. A gradual rise from $0\,E_0$ up to $0.12\,E_0$ is followed by a steep increase of the switching probability. Above around $0.17\,E_0$, p saturates at a value of

3.26%. The switching onset strongly resembles an error function centred at $0.145\,E_0$. The error-function fit and its derivative are plotted as black line and shaded grey area, respectively. The peak of the MgPc molecule's LUMO resonance obtained from steady-state $\mathrm{d}I/\mathrm{d}V$ measurements (see Fig. 2.12, right y-axis in green) is centred at 1.2 V. As mentioned before, only terahertz peak voltages $V_{\mathrm{LW}}^{\mathrm{peak}}$ that are sufficiently strong to charge the molecule, hence reach the LUMO resonance, can induce a change of the adsorption geometry:

$$p(E_{\mathrm{LW}}^{\mathrm{peak}}) = \alpha \cdot T(E_{\mathrm{LW}}^{\mathrm{peak}}) \tag{2.12}$$

$$\frac{\mathrm{d}p}{\mathrm{d}E_{\mathrm{LW}}^{\mathrm{peak}}} = \alpha \cdot \frac{\mathrm{d}T}{\mathrm{d}E_{\mathrm{LW}}^{\mathrm{peak}}} \propto \frac{\mathrm{d}I}{\mathrm{d}V} \tag{2.13}$$

Thus, the peak of the derivative $\mathrm{d}p/\mathrm{d}E_{\mathrm{LW}}^{\mathrm{peak}}$ can be concluded to lie at a field strength $E_{\mathrm{LW}}^{\mathrm{peak}}$, which transforms into a voltage across the junction of $V_{\mathrm{LW}}^{\mathrm{peak}} = V_{\mathrm{LUMO}}$. This direct conversion from switching probability to applied tip-sample bias provides a voltage gauge for the peak of the transient voltage pulse in units of volts. In the case of the data at hand, an incident far field of $0.145\,E_0$ corresponds to a transient peak voltage across the tunnelling junction of 1.2 V. Consequently, the unattenuated far field leads to a transient peak voltage of 8.3 V. It is important to highlight, that for this experiment the terahertz transients traversing both interferometer arms (see sec. 2.3) were superimposed constructively to achieve the peak field of E_0.
The single-molecule switch can thus be used as an Ångström-sized voltage gauge, enabling the readout of local field distributions. Knowing the exact field strength necessary to induce dynamics of nanoscopic devices, we envision the possibility to precisely determine adsorption energies of molecules on surfaces or the binding energy of multiple molecules to another.

2.5 Sampling ultrastrong atomically confined light fields

The work presented in this section was accomplished in collaboration with Dr. Dominik Peller and Carmen Roelcke. Our manuscript was published in the journal Nature Photonics under the title "Quantitative sampling of atomic-scale electromagnetic waveforms" *[Pel21].*

Electromagnetic radiation has the interesting property that it is not pushed out of spaces that are smaller than its wavelength λ but instead can penetrate into them. Using nanostructures such as nanoantennas [Mer08, Liu12, Yos15], nanoparticles [Esm19], nanogaps [Seo09, Sav12], and tips [Krü11, Her12, För16], light confinement to volumes even smaller than the skin depth in metals [Sav12, Yos15, Ryb16] of several nanometres has been achieved. This confinement leads to a massive field enhancement by several orders of magnitude. Choosing suitable nanostructures, the strength of the near field can be custom-tailored, leading to advancements in light harvesting [Bar03], lightwave electronics [Krü11, Her12, Coc13, Lan14, För16, Ryb16, Coc16, Yos16, Jel17], nanoimaging [Eis14, Coc13, Coc16, Bas16, Yos16, Yos18], and spectroscopy [Bas16, Li17] even to sub-molecular resolution [Coc16, Lee19, Amm21]. While for some applications only the absolute field strength is of importance, in the majority of cases the atomic light-matter interaction additionally depends critically on the precise temporal shape of the oscillating carrier wave in the near field [Kub07, Krü11, Her12, Bar15, Ryb16, Rác17, Yos19, Jes19]. The exact shape of the nanostructure as well as quantum mechanical effects such as tunnelling [Bar15] strongly influence the exact near-field waveforms.

So far, quantitatively extracting the electromagnetic transients on the Ångström scale has been highly challenging. In the far field, on the other hand, multiple techniques have already been in use for several decades that can map out the oscillating carrier wave, like an oscilloscope for light [Küb04, Sed20]. The most prominent method involves electro-optic sampling (EOS), which has already been explained in a previous section (sec. 2.3). This field-resolved detection of waveforms in the terahertz, mid-infrared [Hub00a, Küb04], and even near-infrared spectral regions [Kei16] has allowed for the investigation of light-matter interaction faster than a single cycle of light.

In 2021, we demonstrated the first direct quantitative measurement of near-field waveforms within an atomically-sized nanogap [Pel21]. We built on lightwave-driven STM, utilising the carrier wave of terahertz pulses to drive controlled single-electron tunnelling from a sharp tungsten tip into the sample. The terahertz radiation was focused onto the tip, which acts as an antenna, and propagated along the shaft. The confinement of the electromagnetic field into the Ångström-scale tunnelling gap leads to a field enhancement by several orders of magnitude. Most effects, such as coupling to the metallic tip, plasmonic propagation into the near field, screening, as well as phase changes due to the antenna geometry can be classically calculated. However, classical electrodynamics fail to quantify the quantum effects of the nanoscale geometry of the junction such as electron tunnelling.

In order to realise sub-nanoscale and simultaneously calibrated detection of waveforms, a similarly-scaled field gauge must be inserted into the tip-sample junction. For this purpose, we make use of the molecular switch introduced in section 2.4. The switching probability of this single-molecule detector is strongly nonlinearly dependent on the tip-confined peak voltage (see Fig. 2.12 and Fig. 2.13a). Therefore, in proximity to the steepest region of the error-function-shaped onset, corresponding to the LUMO resonance at 1.2 V, small changes to the peak voltage (red horizontal bars) lead to strong variations in the switching probability (red vertical arrows), maximising the sensitivity of our detection scheme.

To not only read out the peak voltage across the junction, but the full waveform, we utilised the interference between two terahertz pulses. Pairs of terahertz pulses, which can be individually scaled in field strength and can be delayed with respect to one another, were created in the Michelson interferometer described in section 2.3. Two terahertz transients were superimposed and coupled into the STM junction with a variable delay time τ. The pulses shall henceforth be called test waveform $V_{\mathrm{NF}}(t)$ and gate waveform $V_{\mathrm{gate}}(t)$ (Fig. 2.13b, red transient), respectively. The far-field peak field of the gate waveform was chosen to be five times larger than that of the test waveform. The delay-time-dependent total voltage applied across the junction due to interference is $V_{\mathrm{sum}}(t) = V_{\mathrm{gate}}(t-\tau) + V_{\mathrm{NF}}(t)$. For a fixed delay time τ, illuminating the tip with the superimposed waveform will lead to a peak voltage of $V_{\mathrm{sum}}^{\mathrm{peak}}(\tau) = V_{\mathrm{gate}}^{\mathrm{peak}} + V_{\mathrm{NF}}(\tau)$ across the gap. Performing ultrafast action spectroscopy, we can extract the mean switching probability p per pulse pair for

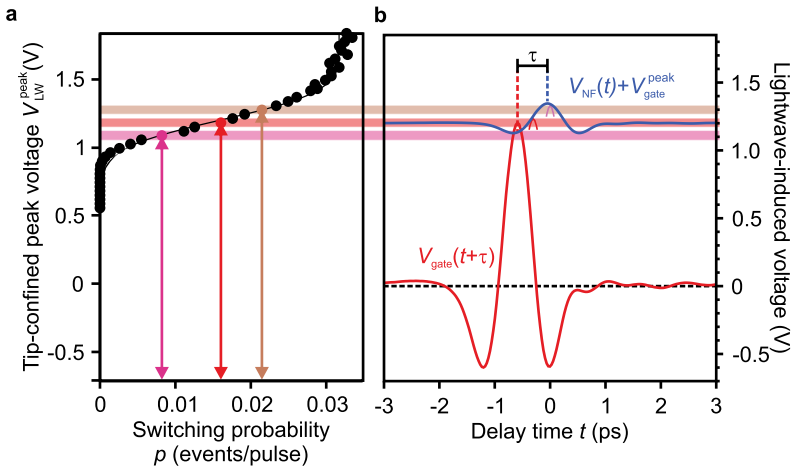


Figure 2.13 | Molecular switch as local voltage gauge. a, The switching onset (black data points) as a function of peak voltage follows an error-function shape. Due to the nonlinear dependence of p on $V_{\mathrm{LW}}^{\mathrm{peak}}$, small offsets to the tip-confined peak voltage (coloured horizontal bars) result in large modulations of the switching probability (arrows). **b**, A terahertz pulse with its peak field tuned to the steepest region of $p(V_{\mathrm{LW}}^{\mathrm{peak}})$ acts as a gate pulse (red transient). Superimposing it with a five times weaker test waveform $V_{\mathrm{NF}}(t)$ (not shown) leads to a delay-time-dependent total peak field (blue curve) that traces the shape of the test waveform: $V_{\mathrm{sum}}^{\mathrm{peak}}(\tau) = V_{\mathrm{gate}}^{\mathrm{peak}} + V_{\mathrm{NF}}(\tau)$.

this specific delay time. By varying τ, we obtain the quantity $p(\tau)$, which encodes the instantaneous peak voltage for each delay time τ. Using the calibration curve (see Fig. 2.12), $p(V_{\mathrm{peak}}^{\mathrm{sum}})$ can be converted into $V_{\mathrm{peak}}^{\mathrm{sum}}(\tau)$. Because the peak voltage of the interfered pulse traces the test waveform $V_{\mathrm{NF}}(t)$, we can retrieve the first time-resolved femtosecond voltage transient on the atomic scale.

The far-field waveform depicted in Fig. 2.14a, recorded using EOS, leads to a transient voltage burst shown in Fig. 2.14b (blue data points). The most prominent differences between the incident terahertz pulse and its near field are the change of the asymmetry, the additional positive cycle at around a delay time of +0.6 ps, and the overall elongation of the pulse duration. More quantitative information on the

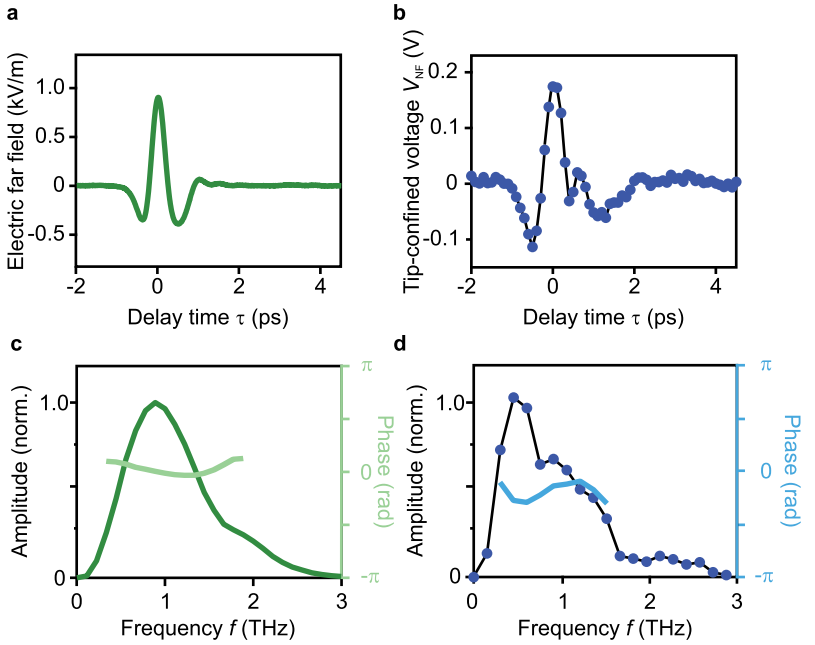


Figure 2.14 | Atomic-scale near-field waveform. a, The far-field waveform depicted in this panel was measured in a field-resolved manner using EOS. **b**, The near-field waveform in the Ångström-sized tunnelling gap is obtained using the molecular switch as atom-scale voltage gauge. Compared to the far field, the waveform is elongated and features an additional positive peak at positive delay times. The black line serves as a guide to the eye. **c**, Via a Fourier transform, the amplitude (dark green line) and phase (light green line) spectra of the far-field waveform are obtained. **d**, The Fourier transform of the near-field waveform shows a different spectral amplitude (blue data points) and phase (light blue line) than the far field. A clear red shift of the maximum of the amplitude spectrum and a CEP shift can be detected.

exact spectral changes can be extracted by analysing the Fourier transform of the two pulses. The amplitude and phase spectra for far and near field are shown in Fig. 2.14c and d, respectively. The tip-confined spectrum (panel d) peaks at a lower frequency of 0.5 THz compared to the far field (0.9 THz) and exhibits some additional oscillatory features. The spectral phases (panel c, light green and panel d, light blue) are both similarly flat, but the near-field pulse features an overall frequency-independent downshift of approximately $-\frac{\pi}{3}$ in comparison to the close-to-zero phase in the far field.

The sum of all effects that modify the incident electromagnetic radiation including coupling, propagation, as well as quantum phenomena can thus be visualised in a complex-valued and frequency-dependent far-to-near-field transfer function (see Suppl. Fig. A.2, blue data points and light blue solid line) assuming a locally homogeneous field distribution and a tip-sample distance of 10 Å. In order to determine which contributions to the transfer function arise from classical electrodynamics and which effects are of quantum nature, the classical field propagation was calculated numerically [Nee18]. In a finite-element calculation, Maxwell's equations were solved in the frequency domain. Taking into account the geometry of our real tip-sample junction (etched tip 10 Å above the planar surface) and the complex-valued dielectric functions of the materials (tungsten tip above a gold surface), the classically expected transfer function for the electric field was determined. The conversion from field strength obtained in the simulations to near-field voltage V_{NF} was done by multiplying the electric field at the centre of the junction by the tip-sample separation d: $V_{\mathrm{NF}} = E_{\mathrm{NF}} \cdot d$. In the regime where the field enhancement scales inversely with the gap size ($E_{\mathrm{NF}} \propto \frac{1}{\mathrm{d}}$), as envisaged by classical antenna theory [Kan18], the voltage drop between the tip and the substrate remains unaffected by the height of the tip. This establishes the absolute scale for the tip-confined voltage waveform $V_{\mathrm{NF}}(t)$. By employing this gauge, classical simulations accurately predict a tip enhancement of up to 2×10^5, generating atomically strong fields from an incident far-field strength of around 350 V/cm.

The frequency-dependent field enhancement and phase retardation are depicted in Suppl. Fig. A.2.

Applying the calculated transfer function to the far-field spectra results in the black dotted curve for the amplitude and the grey line for the phase depicted in Fig. 2.15a.

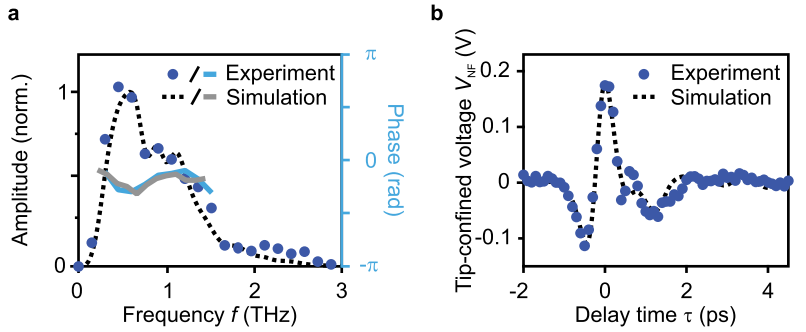


Figure 2.15 | Simulated vs. experimental near field. a, The measured (blue data points and bright blue line) and simulated (black dotted and grey lines) amplitude and phase spectra of the near field in the tunnelling junction exhibit a striking resemblance. **b**, The simulated pulse (black dotted line) acquired by Fourier transform and the measured near-field waveform (blue data points) agree likewise.

The agreement of the experimental data (blue points and blue line) with the simulated spectra is striking. Performing a Fourier transform into the time domain reveals the almost perfect overlap of measured and simulated voltage waveforms (see Fig. 2.15b).

Our approach of superimposing two waveforms and employing a molecular switch as an atomically precise voltage gauge can be used to quantitatively extract the tip-confined transient bias pulse without the need for any indirect observation techniques. However, within our experiments we implicitly made two assumptions. First, we considered that ultrafast tunnelling of electrons sets in at the same voltage as in the DC case. This rises the question whether equating the peak of the derivative $\mathrm{d}p/\mathrm{d}V^{\mathrm{peak}}$ with the LUMO resonance might be incorrect, as different local fields may be necessary to facilitate lightwave-driven opposed to steady-state tunnelling. Second, the process of tunnelling electrons may have a back-action on the instantaneous voltage applied across the junction that disturbs the quantity we aim to extract. Theoretical work investigating light-matter interaction at atomic scales has shown that indeed light-driven tunnelling currents can lead to retardation effects and thus modify the electric fields [Jes19]. Our theory collaborators in the team

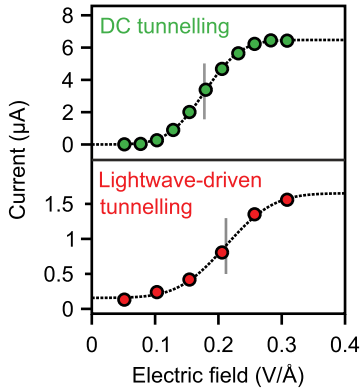


Figure 2.16 | Tunnelling onset: Static vs. ultrafast biasing. The tunnelling current into the molecule is determined as a function of the electric field within the tunnelling gap for static (green circles) and ultrafast (red circles) biasing. A clear error-function-shaped rise emerges in both cases. The tunnelling onset, identified as the inflection point of the fit (vertical grey line), lies at 0.18 V/Å and 0.21 V/Å, respectively.

of Prof. Dr. Ángel Rubio established a full quantum mechanical description of the system with time-dependent density-functional theory (TD-DFT).

As Fig. 2.16 illustrates, the TD-DFT calculations reveal that DC tunnelling under static biasing (green data points) and lightwave-driven tunnelling utilising ultrashort voltage pulses (red data points) sets in at approximately the same electric field strength in the tunnelling junction of 0.2 V/Å (0.18 V/Å for DC biasing and 0.21 V/Å for lightwave biasing). This validates the use of the molecular switch as a voltage gauge, confirming that our experimental calibration of the local transient voltage is accurate. Plotting the temporal evolution of the electric field, averaged over a small region on top of the molecule, unveils strong retardation effects (see Suppl. Fig. A.3) that modify the field distribution. However, this is only true if the number of electrons tunnelled per pulse is larger than one. When a single isolated electron tunnels from the tip into the molecule, it will be subject to the unaltered energy landscape. Our experiments thus directly rely on the quantum mechanical principle that a particle does not modify its own potential. The tunnelling process is quasi-instantaneous and only occurs at the peak of the transient. The test waveform only adds a small offset to the field crest of the gate pulse. Due to the fact that on average less than one electron tunnels within the small time window given by the peak duration of the near field, our sampling scheme is valid and allows for the quantitative extraction of the near-field waveform with atomic precision.

It deserves mention that these TD-DFT calculations were done for a centre frequency of 40 THz in order to reduce the computation time. Cross-checks comparing the result to calculations done with the true centre frequency of 0.9 THz revealed the applicability of this approach. Even at 40 THz, the Keldysh parameter (see eq. 2.8) remains below one and the photon energy well below any electronic excitation in the system. Notably, the fact that the tunnelling onset remains at the same peak field strength of 0.2 V/Å (see Fig. 2.16) even in the multi-terahertz and mid-infrared regime shows the scalability of the LW-STM scheme.
With the described technique, we have found a way that allows for spatio-temporal sampling of atomic-scale near fields. We anticipate that finely calibrated videos depicting sub-Ångström field distributions will not only push the boundaries of classical nano-optics but also offer a direct visualisation of the quantum nature of atomic-scale light-matter interactions. Furthermore, we can now experimentally validate state-of-the-art simulations that bridge the gap between macroscopic light and atomic waveforms. This development has the potential to revolutionise the design of nanotechnology, enabling the creation of novel metamaterials and atomic-scale devices that leverage precisely tailored coherent near fields, all within a time frame faster than a single cycle of light.

2.6 Controlled excitation of librational motion

The results presented in this section were obtained in close collaboration with Dr. Dominik Peller, Dr. Thomas Buchner, and Carmen Roelcke. We published our work in the journal Nature under the title "Sub-cycle atomic-scale forces coherently control a single-molecule switch" *[Pel20b].*

In the previous sections, we have established that the terahertz waveform coupled into an STM junction drastically changes in shape upon propagation into the near field. Utilising a single-molecule switch with a voltage-dependent switching probability and our novel detection scheme of ultrafast action spectroscopy, we were able to directly map out the temporal evolution of the ultrafast voltage pulse applied across the tip-sample junction. With this technique to measure the ultrafast bias pulse shape in the tunnelling gap, we can now custom-tailor our near field. Specifically, in our setup we can chirp the terahertz pulse or apply a spatial/spectral filter, record the far field using EOS (see sec. 2.3), and directly calculate the expected near-field waveform by applying the newly determined complex-valued transfer function (see Suppl. Fig. A.2).

But which effect would such a change of the waveform have in a lightwave-driven STM? For state-selective tunnelling with a transient bias, as has been established in section 2.2, the top 10% of the waveform are used to temporally confine the tunnelling window to sub-picosecond timescales. With the handles on the CEP of the terahertz pulse as well as its spectral composition and phase, we can effectively alter two properties that are relevant for the charging mechanism: the duration of the tunnelling window and the peak field. However, if the electric field directly interacts with the sample, for example by applying a Coulomb force, not only the field strength but also the frequencies and the relative phases within the pulse are crucial for the exact interaction with the sample. Applying forces to matter with various scanning probe techniques has been shown numerous times before to reorder atoms at surfaces into nanostructures [Hla00, Hei02, Str04, Pav17, Esa18]. In the following section, I will introduce how we are able to coherently control the reaction dynamics of a molecular switch using sub-cycle forces applied with Ångström spatial resolution.

To test the hypothesis whether the electric field of the terahertz pulses can excite ultrafast dynamics, we make use of a well-known molecule system. The molecular

switch MgPc, which has been introduced in section 2.4, proves to be a suitable candidate. As mentioned above, the potential landscape of the molecule's neutral and negatively charged states differ considerably. In the electronic ground state, the molecule is subject to a double-well potential symmetric to 0° with minima at around ±10°. In contrast, in the singly-negatively charged case, the potential changes to a singular minimum centred around 0° due to the changed Coulomb interaction of the molecule with the polar sodium-chloride surface (see Fig. 2.9). This allows the molecule to rotate azimuthally and to be switched with a certain probability once charged. A weak pump pulse with its peak voltage $V_{\mathrm{THz}}^{\mathrm{pump}}$ well below the LUMO resonance will not be able to induce switching. For clarity, please refer to Fig. 2.13 where the onset of switching as a function of the tip-confined peak voltage reveals negligible probabilities to switch the molecule below tip-confined voltages of around 0.8 V.

However, a terahertz pulse incident on the tip-sample junction with a peak voltage well below the resonance may still induce dynamics that can be revealed using a pump-probe scheme. Keeping its peak field strength well below any resonances ensures that any effects on the motion of the MgPc molecule must be due to its electric field and cannot be the result of a charging event. The switching probability resulting from a time-delayed electronic trigger pulse is then expected to encode information on the dynamics excited by the ultrafast atomic force.

The green markers in Fig. 2.17 illustrate the variation of the switching probability per pulse pair with the delay time τ between the excitation pulse and the electronic trigger pulse. Clearly, an oscillation pattern of the switching probability at a frequency of 0.3 THz arises. It shows no decay in amplitude up to several tens of picoseconds after excitation, pointing towards weakly-damped dynamics. This is surprising for a molecule adsorbed on top of a polar substrate. In comparison with calculations of phonon modes of the molecular frame using density-functional theory, a possible origin of the oscillation is found. The oscillation frequency of the probability matches a frustrated in-plane rotation of the molecular rotor [Pel20b].

Upon reversing the polarity of the pump pulse, both amplitude and frequency are barely changed (see Fig. 2.17, orange markers). Merely the phase is shifted by π, signifying a coherent excitation by the ultrafast pump pulse. Further measurements with varying pump field strength uncover that the mean switching probability $\overline{p}$

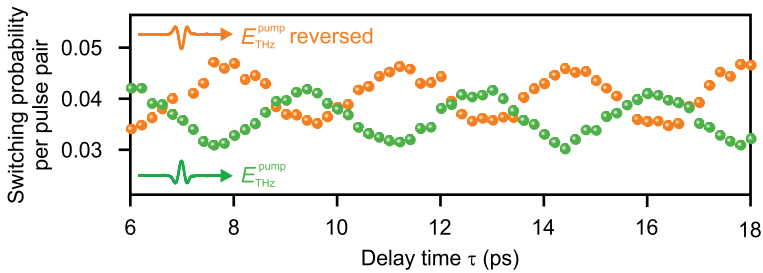


Figure 2.17 | Coherent modulation of switching statistics. Varying the pump-probe delay time reveals modulations of the switching probability per pulse pair with a period of 0.3 THz (green markers). Reversing the polarity of the pump pulse, whilst retaining all other experimental settings unchanged, exposes an oscillation with the same amplitude and frequency but a phase shift of π (orange markers). This indicates a coherent modulation excited by the force pulse that leads to variations of the switching probability.

stays constant as a function of the delay time for a wide range of peak fields (see Fig. 2.18a). Only once the transient voltage V_{LW}^{pump} of the pump pulse suffices to reach the LUMO, $\overline{p}$ starts to rise steeply. The peak-to-peak amplitude p_{pp} of the switching probability oscillation (see Fig. 2.18b), on the other hand, increases linearly with the pump field-strength over the entire measured field range. These results lead to the conclusion that the electric field of the pump pulse couples to the molecular frame, either by directly exerting a force on the charged atoms or via the polarisability of the sample. This interaction coherently excites an in-plane motion, which strongly influences the switching probability.

Repeating such pump-probe experiments, in which only the probe pulse is able to charge the molecule for varying tip positions, reveals strong spatial variations of $\overline{p}$ and p_{pp}, as demonstrated in Fig. 2.19a. Traces of the time-dependent switching probability were recorded at six different positions along the molecular switch, as indicated in Fig. 2.19b. At the positions A–C, no clear modulations of the switching probability p_{pp} are visible, whereas at the locations D–F, strong oscillations of p_{pp} are detected. These spots with maximal as well as the spots with negligible force-induced modulations are less than 5 Å apart. Interestingly, the mean switching probability

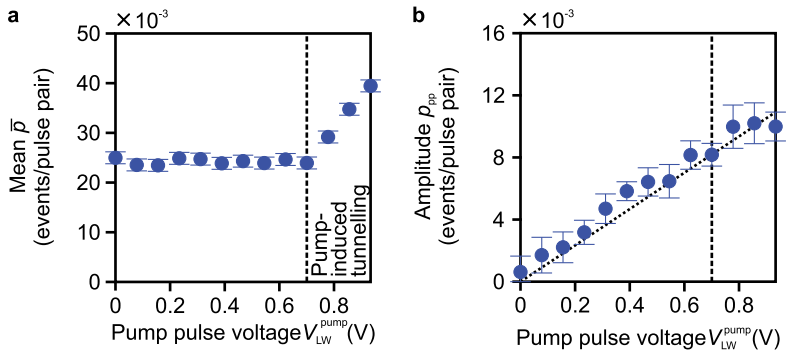


Figure 2.18 | Effect of pump peak voltage on mean and amplitude of switching oscillations. The THz pump – THz probe experiments reveal a clear harmonic oscillation of the switching probability. The mean $\overline{p}$ and peak-to-peak amplitude p_{pp} of these coherent vibrations (see definition in Fig. 2.19a) exhibit a completely different dependence on the peak pump voltage $V_{\mathrm{LW}}^{\mathrm{pump}}$. **a**, The mean switching probability $\overline{p}$ remains constant up to a voltage of around 0.7 V, where the pump pulse begins to be strong enough to tunnel electrons into the LUMO and thus trigger switching events. **b**, The amplitude p_{pp} of the oscillations rises linearly even for low peak voltages, indicating a completely different excitation mechanism that does not rely on charging of the molecule.

shows a reversed spatial dependence: Areas with the largest $\overline{p}$ possess diminishing p_{pp}, while areas with a small $\overline{p}$ display large delay-time-dependent variations of up to 39%. This proves that the excitation by the pump pulse acts independently from the electronic trigger. The electric field in the junction applied by the terahertz transient possesses a mostly vertical/out-of-plane component, as any transversal fields are efficiently screened by the polar salt layers and the free charges in the metal substrate. Nonetheless, the force pulse efficiently excites a mode that influences the switching probability, which is an in-plane/azimuthal rotation. Density-functional theory calculations performed by our collaborator Nikolaj Moll reveal that applying an out-of-plane field locally leads to a minuscule azimuthal displacement of the molecular frame from the potential minimum. These excursions critically depend on the exact position of the applied field. The calculated positions of largest field-induced rotations (see Suppl. Fig. A.1) strongly agree with the positions of high excited mod-

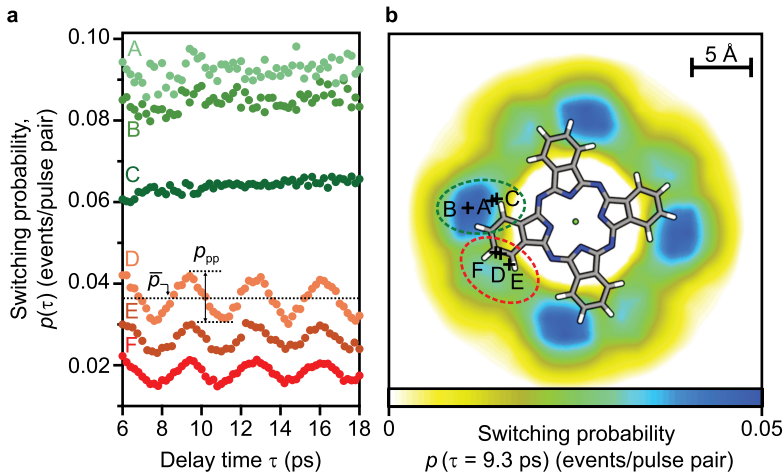


Figure 2.19 | Spatial variations of dynamics. a, Performing pump-probe measurements for different tip positions reveals locally varying dynamics. At the positions A–C, the mean switching probability $\overline{p}$ is larger but no clear modulations p_{pp} are discernible. At the locations D–F, in contrast, a low $\overline{p}$ but strong oscillations p_{pp} are detected. **b**, The map of the mean switching probability for a delay time $\tau = 9.3\,\mathrm{ps}$ shows the 6 spots marked A–F where the traces shown in panel a were recorded. The regions are separated by less than 5 Å but show completely different behaviours.

ulation depth of the switching probability shown in Fig. 2.19a. This conformity arises from the positions of specific atoms in the relaxed molecular geometry. The molecule does not lie completely flat on the NaCl substrate but instead, the isoindole units are slightly twisted (see Suppl. Fig. A.5), as alluded to in the discussion of the local single-pulse switching maps (see Fig. 2.11). A polar interaction of specific hydrogen atoms, which are slightly positively charged due to their smaller electronegativity (with respect to the carbon binding partners), with the anionic chlorine underneath leads to the molecule being relatively strongly bound to the surface at these points. On the one hand, as a result, these parts of the molecule are further away from the tip, leading to a reduced wave-function overlap, thus also a reduced tunnelling and switching probability (see Fig. 2.11). On the other hand, disrupting these lo-

cal hotspots of interaction using force pulses leads to the strongest changes in the adsorption potential and hence the largest mechanical oscillation amplitudes. This causes stronger fluctuations of the tunnelling probability and thus the switching probability can be more effectively modulated. This demonstrates the high spatial precision with which we can apply ultrafast forces and the strong local differences in the ultrafast modulation excited.
The coherent control of the molecule's switching dynamics with Ångström fidelity is a first step towards path-selective real-space reaction microscopy with combined sub-cycle and atomic precision. We believe that this approach will allow for the control of structural transitions in molecules and solids on their intrinsic length and timescales. Manipulating ultrafast dynamics will prove useful in countless applications including most prominently (bio-)chemical reactions [Sch17a] and phase transitions in strongly-correlated materials [Eic10].

2.7 Towards lightwave scanning tunnelling microscopy of non-classical quantum motion

The analyses presented in this section were performed in close collaboration with Dr. Yaroslav Gerasimenko and Vincent Eggers.
As demonstrated in the previous chapters, the ultrafast nanoscopy technique of LW-STM allows for glimpses into the intricate dynamics occurring in the nanoworld. For instance, this method has been successfully employed to record the first atomic-scale movie with femtosecond resolution, visualising the oscillations of a single pentacene molecule [Coc16]. As presented in previous sections of this thesis, we were able to recently record toggling events of a single-molecule switch (magnesium-phthalocyanine, short MgPc) with shot-to-shot fidelity using so-called ultrafast action spectroscopy. Moreover, we were not only able to observe the molecular switch but also to purposefully manipulate the process. Specifically, the switching probability could be influenced by up to 40% using an ultrafast atomic force pulse to excite a librational motion of the molecule within its electronic ground state (see sec. 2.6 and ref. [Pel20b]).
However, the motion of these single molecules – the vertical oscillation of pentacene

and the in-plane libration of MgPc – was so far considered to be purely classical. The size of these molecules investigated lies in the range of a few Ångström and their mass measures a few hundreds of atomic mass units u ($1\,u = 1.661 \times 10^{-27}\,\mathrm{kg}$). As these sizes and masses lie far in the nano-realm, the questions arises whether the molecular vibrations could be prepared in a non-classical state, only describable by quantum mechanics.

Non-classical oscillator states. Before going into detail on the librational motion of the MgPc molecule, a brief general overview of quantum harmonic oscillators in general is provided.

In the nano-realm, the effects of uncertainty and quantisation play crucial roles. Notably, the quantum mechanical analogue of a classical harmonic oscillator is a coherent state or so-called Glauber state [Gla63], where the uncertainties in the conjugated variables are equal and minimal. Such pairs of variables include position x and momentum p (as in the case of a linear oscillation) or angular displacement θ and angular momentum p_θ (in the case of a libration/torsion). Yet, the generic quantum harmonic oscillator can feature a variety of quasiprobability distributions – so-called squeezed states and Schrödinger cat states [Jan94] – none of which have a classical analogue of motion. In a quantum squeezed state, the uncertainty of one variable is reduced or "squeezed" at the expense of the uncertainty in the complementary non-commuting quadrature [Cav81]. However, the evolution of the mean of the variable would be indistinguishable from the classical trajectory. These seemingly exotic squeezed states naturally arise in quantum entanglement [Hal99] and are commonly used in the realm of quantum information. They have been both generated and observed in various systems, including optical and microwave modes [Yur90], trapped ions [Kie15], ultracold atoms [Hal99], molecular oscillators [Bur19], micrometre-scale mechanical resonators [Asp14, Wol15], and even in solids [Gar97].

For a visualisation and clarification of the differences between coherent and squeezed states, let us assume three differently prepared oscillator states and record the oscillator displacement A (be it x or θ) as a function of time. Figure 2.20a-c displays the temporal evolution of the oscillator displacement $A(t)$ of a coherent (blue), an amplitude-squeezed (red), and a phase-squeezed (green) state over one oscillation period. The occurrence histograms for a certain displacement at each point in time are displayed in Fig. 2.20d-f for the three cases. While the mean displacement at

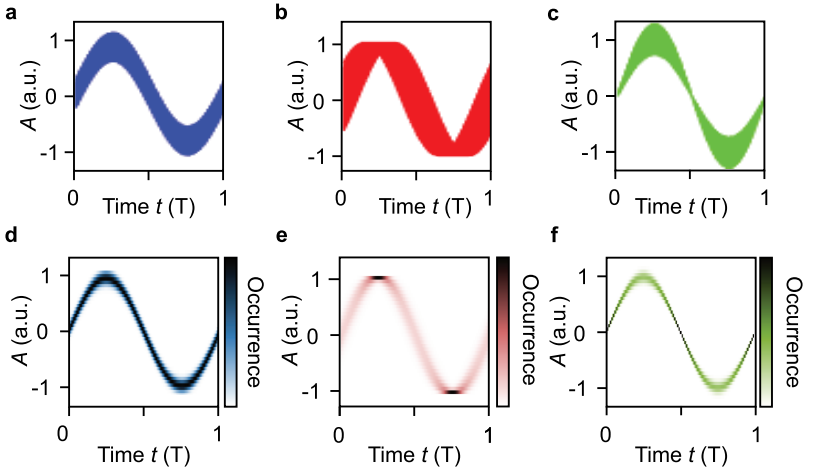


Figure 2.20 | Coherent and squeezed oscillator states. a-c, The time evolution of the displacement A of a coherent (panel a), an amplitude-squeezed (panel b), and a phase-squeezed state over a full oscillation period T are depicted. **d-f**, The histograms of the occurrence of a certain oscillator displacement A differ strongly for the three different states. The uncertainty is independent of the oscillation phase in the case of a coherent state (panel d), while it is minimised at the turning points for an amplitude-squeezed state (panel e) and at the zero-crossings for a phase-squeezed state (panel f).

each moment in time is the same for the three respective states, the uncertainties show a different behaviour: The width of the oscillation in case of a coherent state is constant over the whole period, whereas the histograms for the amplitude- and phase-squeezed states exhibit minima in their width at the turning points and the zero-crossings of the oscillation, respectively. Thus, the uncertainty in excursion is minimised twice within each oscillation period T for squeezed states, while for a coherent state it remains constant. This oscillation of the variance at the second harmonic of the mean oscillation frequency is an important indicator for non-classical oscillator states. This highlights that measurements of mean values alone are insufficient for describing quantum degrees of freedom. Thus, it is necessary to also obtain access to the uncertainty of the coordinate under study. We need to examine the

distribution of this variable and could potentially infer quantum properties from the mean and widths of these distributions. In the following, we discuss a few key considerations that could ultimately contribute to measuring the quantum motion of individual molecules in their natural environment using LW-STM.

Signatures of quantum motion in the switching statistics? As has been established in section 2.6, we can deflect the single-molecule oscillator MgPc from its potential minimum using an ultrashort force pulse. I have shown that the effect of the force pulse is extremely local (see Fig. 2.19). At locations where large oscillation amplitudes are excited and thus the harmonic estimation of the vibrational potential is no longer valid, intriguing librational motion could potentially be created. Moreover, the force pulse is not completely impulsive, but the full temporal evolution of the electric field could influence the precise motion induced. All these factors could, in principle, foster the creation of non-classical vibrational states in a single molecule [Dun93, Dun95].

However, the readout of the momentary position of the molecular frame proves challenging. In the previous sections, we used the switching probability of the librating molecular switch as an indicator for the mean angular displacement of the molecular frame from its potential minimum. The question arises whether the switching statistics of the vibrationally excited MgPc molecule encode more information on the vibrational state beyond the mean oscillator displacement. Studying statistical quantities beyond the first moment (the mean) may give first insights. As explained in section 2.4, the novel technique of ultrafast action spectroscopy is able to provide shot-to-shot information on the switching success. Thus, the outcome of every single switching attempt (laser pulse injection) can be registered and analysed.

To study the switching statistics as a function of the oscillator phase, time traces similar to the ones shown in Fig. 2.10b were recorded for various pump-probe delay times. Figure 2.21a depicts the occurrence histogram for $N = 2000$ switching attempts, colour-coded from low (green) to high (red) switching occurrence. The mean switching probability fluctuates periodically at a frequency of 0.3 THz (see Fig. 2.21b and sec. 2.6). Moreover, fine details are visible in the histogram width as a function of the delay time. To investigate the differences in the temporal evolution of the mean and the width of the switching histograms, we aim to magnify the intricate details in the histograms. To this end, we subtract the mean switching probability,

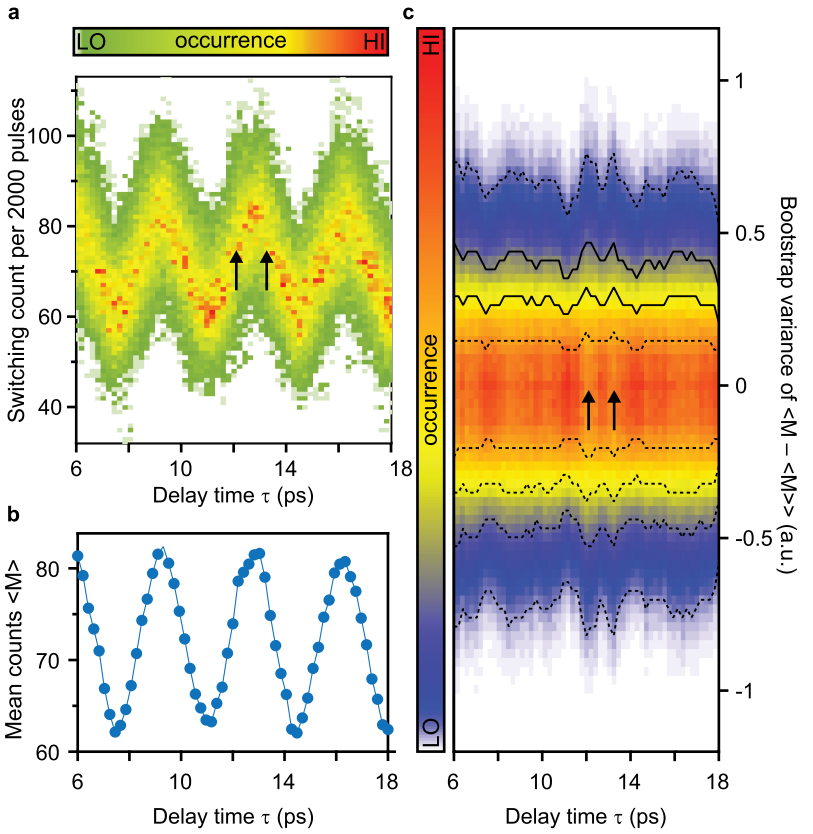


Figure 2.21 | Time-resolved counting statistics. a, We recorded the occurrence of the switching events M (for $N = 2000$) as a function of τ and plotted it in a histogram for each delay time. The black arrows indicate delay times with particularly broad distributions seen in panel c. **b**, Calculating the mean displacement for each delay time reveals oscillations of the mean number of switching events $\langle M \rangle$ at a frequency of 0.3 THz. **c**, To improve the statistical accuracy, the variance of the relative number of switching events $M - \langle M \rangle$ is bootstrapped. The oscillation frequency clearly differs from the one shown in panel b. Broad distributions occur at zero-crossings of the oscillation of the mean (black arrows). The dotted and solid black lines delineate the position of the 1%, 5%, 25%, as well as 50% occurrence.

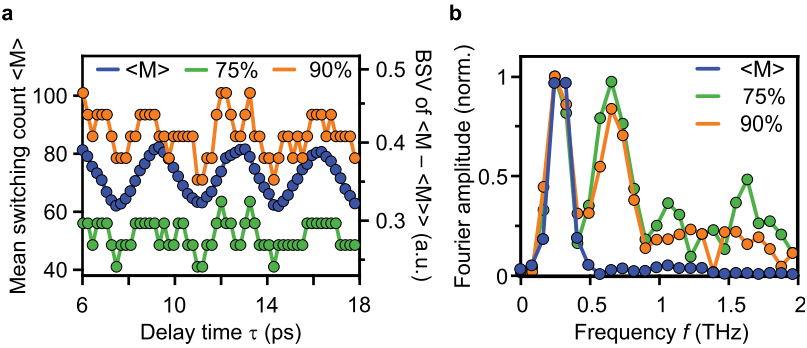


Figure 2.22 | Time evolution of the switching statistics. Closer analysis of the time evolution of mean and variance of switching provides further insight. **a**, The temporal evolution of the mean switching probability ⟨M⟩ (blue) and the line cuts through the bootstrap variance (BSV) depicted in Fig. 2.21c (solid black lines) at the 75% (green) and 90% (orange) occurrence level show different oscillation frequencies. **b**, Calculating the Fourier spectrum of the temporal evolution of mean switching count ⟨M⟩ (blue) as well as the line cuts through the BSV at the 75% (green) and 90% (orange) occurrence level depicted in panel a clarify the presence of higher harmonics of the mean oscillation frequency in the line cuts.

eliminating the dominant modulation at a frequency of 0.3 THz. To further improve the statistical power, a statistical resampling technique called bootstrapping (see appendix A.4.1) of the residue $\langle M - \langle M \rangle \rangle$ for each delay time is performed. Using this method, an estimate of the standard deviation of the switching probability $\sigma_{\langle M-\langle M \rangle \rangle}$ can be obtained without the need for any model of the underlying statistics. Strikingly, the width modulation (standard deviation) depicted in Fig. 2.21c oscillates faster than the oscillation of the mean switching probability and the most noticeable peaks appear at zero-crossings of the mean (Fig. 2.21c, black arrows). To assess the differences in the temporal evolution of the mean and the width of the switching occurrence histograms, two exemplary line cuts through the bootstrap variance (BSV) – the 75% (Fig. 2.21c, lower solid black line) and 90% occurrence levels (Fig. 2.21c, upper solid black line) – are compared to the mean switching probability. The conspicuous presence of higher-harmonic contributions detected in the time domain (Fig. 2.22a) becomes even clearer in the frequency domain (Fig. 2.22b).

The mean switching probability exhibits one dominating Fourier peak at 0.3 THz. In contrast, the line cut through the 90% occurrence level (orange curve) displays a second peak in the Fourier amplitude at 0.6−0.7 THz, while the cut at the 75% occurrence level (green curve) exhibits even higher additional frequencies (1.1 THz and 1.7 THz). This solidifies the observation that a clear high-frequency oscillation is present in the variance and readily demonstrates the new dimension of data contained in the switching statistics.

Notably, the mapping of uncertainties in the switching probability onto the specific oscillator state of MgPc is non-trivial. In order to switch the molecule, an electron must be transferred from the STM tip into the molecule. The probability of tunnelling one electron from the tip into the molecule critically depends on the dipole matrix element between the initial state (assumed s-wave of the tip) and the final state (new combined electronic and vibrational wave functions). It remains to be clarified to which extent the measurement (tunnelling) retains information of the vibrational state of the molecule. To solve this question, elaborate calculations involving the following steps are necessary: First, the combined vibrational and electronic wavefunction of such an excited molecule must be determined. The polar substrate (sodium chloride) may additionally influence the wavefunction. Second, it must be ascertained how exactly this wavefunction collapses upon measuring. Does a projection onto the real-space coordinate occur or does the molecule remain in a vibrational eigenstate? Third, the exact position of the electron injection is not arbitrarily precise but delocalised over a finite "area". This additional uncertainty combined with the aforementioned aspects renders making statements regarding the positional uncertainty of the oscillator from our measured data extremely difficult. Moreover, purely classical uncertainties additionally pollute our signal dramatically. In the counting statistics of single tunnelling/switching events, oscillations of the tip-sample distance as well as the position of the tip lead to large additional uncertainties, which must be considered in the analysis. Nonetheless, more detailed modelling in the future may help in disentangling the different contributions to the measured data.

To conclude, our work has introduced the counting statistics of single quantum events in LW-STM. Through the use of sub-cycle atomic forces directly acting on select atoms, we have excited librational motion in a single molecule. The switching probability as a function of pump-probe delay time shows intriguing features: Tracing the

mean switching probability visualises the librational motion of the molecule. Apart from that, fluctuations in the widths of the switching occurrence histograms are detected. The switching uncertainty is particularly broad at the zero-crossings of the oscillation. While the precise origin of these signatures is currently not clear, yet, the hypothesis that non-classical amplitude-squeezing of the libration state may play a role will be investigated with high priority. I have shown that ultrafast action spectroscopy within a lightwave-driven STM may illuminate the quantum realm of single-molecule oscillators and detect fluctuations at the quantum limit. In order to verify whether the width variations originate from quantum mechanical uncertainty and whether squeezed quantum states are detected, meticulous inspection and further statistical analyses are necessary.

Another path towards the visualisation and control of quantum states within LW-STM may involve the implementation of electroluminescence and the analysis of the statistics of the emitted photon energies. Previous studies have presented potential signatures for squeezed states using fluorescence experiments in the far field [Dun93]. Extending this idea to single nanoscopic entities, using the spatio-temporal capabilities of LW-STM may elucidate the position-dependent response of single molecules and the dynamics of different vibrational quantum states therein. A second possible route towards the detection of quantum oscillator states in single molecules lies in the use of a source of visible photons with tunable energies to excite the sample. Properties such as the absorption as well as the emission of an oscillating molecule may critically depend on the exact phase of the oscillation.

The foundational concept of counting statistics with sub-picosecond temporal and Ångström spatial resolution in LW-STM in combination with an additional information channel in form of the absorbed/emitted photons may pave the way towards the observation and even controlled engineering of quantum states of single molecules.

Chapter 3

Controlling phononic motion and ultrafast energy level shifts of atomic defect levels in monolayer tungsten diselenide

The data presented in this chapter were acquired and analysed in close collaboration with Carmen Roelcke, Dr. Yaroslav Gerasimenko, and Andreas Biereder. A manuscript has been submitted for publication. Some figures within this chapter are adapted from that manuscript.

Since the advent of graphene, the interest in atomically-thin crystals has been on a constant rise. Among these, monolayer transition-metal dichalcogenides (TMDCs) have garnered significant attention as a diverse class of materials, exhibiting a wide range of properties, especially with regards to their electric conductivity: The heterogeneous group of TMDCs features insulators, semiconductors, semimetals, and metals [Wil69, Cha20]. Diversifying this wide range even more, different atomically-thin TMDC crystals can be stacked on top of each other to create heterostructures. If two layers of crystals are combined and rotated with respect to one-another, so-called Moiré superstructures emerge, providing an additional level of nanoscopic complexity [Mak22]. These diverse and complex heterostructures offer an auspicious playing field for investigating and actively manipulating a plethora of fascinating mechanisms such as excitonic (electron–hole pair) [Wan18, Cia22, Nai22] and spin-related physics, including valley pseudospin and strong spin-orbit interactions [Xu14, Man15]. Mas-

tering such intricate properties is of fundamental interest as it offers an enticing opportunity for tailoring quantum materials [Zho21, Mak22]. Remarkable progress has been made through recent advancements in time- and angle-resolved photoelectron spectroscopy, enabling investigations of electron dynamics with momentum resolution on femtosecond timescales [Kar22, Sch22, Mad20, Wal21]. These advanced techniques not only provide fundamental insights into the physics of TMDCs, but also hold tremendous potential for the development of novel optoelectronic devices based on atomically-thin TMDCs.

All these electronic properties are intricately linked to the atomic lattice of the material. The nuclei form the potential landscape in which the charges propagate and thus strongly govern the dynamics of electronic excitation. Lattice vibrations thus not only alter the relative distance of the nuclei but affect the energies of electronic levels. Even the line width of exciton emission has been shown to sensitively depend on the strain of the lattice [Nie18]. The scattering times of excitons and thus the dynamics of these quasi-particles are strongly influenced by phonons [Hel21] reaffirming the strong interaction of electronic and phononic excitations [Nie18]. Hence, it is of utmost interest to study how phononic motion influences the energy of electronic levels in atomically-thin materials.

In this chapter, I will initially introduce the material class of transition metal dichalcogenides. We followed a multistep fabrication process to generate vacuum-compatible monolayer samples and we investigated mono-atomic vacancies therein. Following the controlled excitation of a vertical phonon mode (sec. 3.2) with a terahertz pulse, I will present how LW-STM is able to resolve the ultrafast changes to the tunnelling current as a function of delay time between excitation and readout. Atomic-scale differences in amplitude and baseline of these oscillations are detected. Extending the capabilities of LW-STM, which has so far only been employed for imaging, to achieve ultrafast energy resolution, I will introduce time-resolved light-wave driven scanning tunnelling spectroscopy. Using this novel approach, the first visualisation of ultrafast shifts of localised energy levels on sub-picosecond timescales (sec. 3.5) will be displayed.

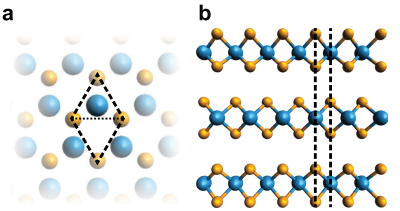


Figure 3.1 | Crystal structure of WSe_2. a, The top view of a tungsten diselenide monolayer visualises the hexagonal crystal structure. The unit cell (black dashed rhomb) consists of two prismatic cells (indicated by dotted line), one cell with a tungsten atom (blue sphere) and one without, surrounded by selenium atoms (yellow spheres). **b**, The side view of three WSe_2 monolayers illustrates the 2H-stacking order. Neighbouring layers are rotated by 60° with respect to each other and shifted such that a chalcogen comes to rest above a metal atom and vice versa (see dashed lines).

3.1 Monolayer transition-metal dichalcogenides

The material group of transition-metal dichalcogenides form the chemical structure MX_2, per definition consisting of a transition metal (M) from the d-group in the periodic table and two atoms from the chalcogen group (X). In the scope of this thesis, the semiconducting TMDC tungsten diselenide (WSe_2) was investigated closely. Its crystal structure is identical to its cousins' molybdenum diselenide ($MoSe_2$) and tungsten disulphide (WS_2). These materials form layered crystals with a trigonal prismatic structure where every transition metal in the centre is covalently bound to six surrounding chalcogen atoms. Within the layer, as illustrated in Fig. 3.1a, this reveals a hexagonal lattice structure similar to the honeycomb lattice of graphene. In the naturally occurring, so-called 2H-stacking order, neighbouring monolayers are rotated by 60° and shifted with respect to each other such that a tungsten atom comes to rest exactly between two selenium atoms of the neighbouring layers (see dashed lines in Fig. 3.1b). The van der Waals bound layers can be mechanically exfoliated down to the monolayer level.

As illustrated in Fig. 3.2, thinning down WSe_2 changes its electronic band structure. More precisely, the conduction band minimum shifts from a position in between the Γ- and K-points in the bulk material to the K-point in the monolayer, while the valence band maximum moves from the Γ-point to the K-point. This effectively

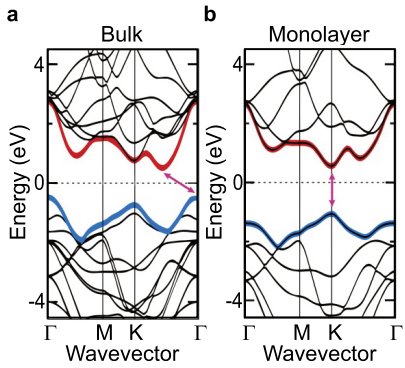


Figure 3.2 | Band structure of bulk and monolayer WSe_2. The valence (red) and conduction bands (blue) change drastically in shape when exfoliating WSe_2 down to the monolayer. **a**, The band structure of bulk WSe_2 when passing through high-symmetry points of the reciprocal lattice from Γ- to M-, K-, and back to the Γ-point is depicted. The indirect bandgap (pink arrow) measures 1.2 eV and lies between valence band maximum at the K-point and conduction band minimum located between the K- and the Γ-point. **b**, In the WSe_2 monolayer, the direct bandgap (pink arrow) lies at the K-point and measures 1.56 eV. The band structures are adapted from Kumar *et al.* [Kum12].

turns WSe_2 from an indirect- ($E_{gap} = 1.2$ eV) into a direct-bandgap ($E_{gap} = 1.56$ eV) semiconductor. This means that excited charge carriers (electrons and holes) after equilibration relax towards the same momentum/point in reciprocal space – the K-point – in a monolayer. This fundamentally changes exciton formation and their dynamics.

Sample fabrication. As mentioned above, the individual layers are weakly bound by van der Waals interaction, which allows for relatively simple mechanical exfoliation, reducing the crystal to one single layer. Traditional techniques such as polydimethylsiloxane or "sticky-tape" methods [Li22] were deemed unsuitable for our application due to the presence of polymer residues that not only would render the samples unsuitable for ultra-high vacuum environments but also would pose a threat to the scanning tip within the STM due to their insulating behaviour.

In our approach, a clean gold surface played a crucial role by providing a suitable platform for the transfer process [Vel18]. Prior to the transfer, the gold surface was cleaned during a thorough preparation procedure in a vacuum (approximately 10^{-9} mbar) environment. More precisely, the gold surface was sputtered with Argon ions followed by an annealing step to reconstruct the top layers and generate an atomically flat surface with a (111) surface crystal orientation. By gently pressing

a freshly cleaved bulk WSe_2 crystal (purchased from HQGraphene) onto the gold surface, the bottom-most monolayer adheres to the clean substrate. The precise nature of this interaction has been investigated in the context of gold-sulphur bonding but is anticipated to exhibit strong similarities in selenium. Due to a very weak hybridisation of orbitals [Zho16, Far16] between the topmost gold atoms and the lower chalcogen atoms of the crystal, this interaction can be identified as strong van der Waals [Vel18] in character. Because of the weak van der Waals interaction between individual WSe_2 layers, lifting off the bulk TMDC crystal leaves behind a single monolayer of WSe_2. Annealing the whole structure at 360°C in ultra-high vacuum for up to 12 hours until the pressure stayed on the order of 10^{-9} mbar ensured desorption of all remaining surface contaminants.
The described gold-assisted exfoliation approach not only eliminates the issues associated with polymer residues but also provides a reliable and effective method for the successful transfer of TMDC monolayers.

Moiré superlattice and atomic defects. As noted earlier, the stacking of monolayers of TMDCs gives rise to the emergence of a novel periodicity, which is determined by the rotation angle and lattice constants of its constituent layers. The intriguing surface corrugation can be imaged using STM, as Fig. 3.3a illustrates. The periodic maxima in the apparent height shall henceforth be referred to as Moiré peaks. Performing a Fourier transform of the entire image reveals six additional peaks at $(0.17 \pm 0.01)\,\text{Å}^{-1}$ from the centre. This corresponds to a Moiré periodicity of approximately $\frac{1}{(0.17\pm0.01)\,\text{Å}^{-1}} = (5.9 \pm 0.3)\,\text{Å}$ (see Fig. 3.3b). From the lattice mismatch between WSe_2 (lattice constant a=3.3 Å) and Au(111) (a=2.8 Å), the rotation angle between the monolayer and the gold surface can be estimated to lie at around 38°. A line cut along the white arrow in Fig. 3.3a reveals the Moiré height variations on the order of 2 pm (see Fig. 3.3c). Apart from the Moiré protrusions, local disruptions of the periodicity, as a result of atomic substitutions and vacancies, are visible in Fig. 3.3a. The observed local irregularities illustrate the well-known phenomenon that crystalline structures are never perfect. Instead, crystalline materials, such as the examined TMDC monolayer, feature ubiquitous point defects, substitutions, as well as line defects that strongly modify the physical properties of the sample. For instance, they can affect the thermal [Zha12] and electronic conductivity [Zho04] on the large scale. On the microscopic scale, they change the electronic structure

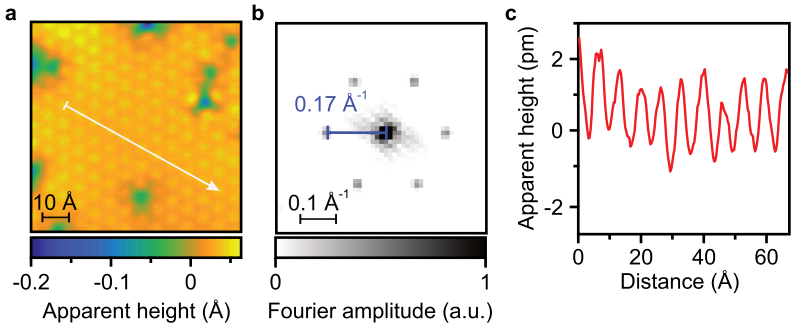


Figure 3.3 | Protrusion height of the Moiré superlattice. a, A DC-STM constant-current image of the WSe_2 monolayer reveals the hexagonal Moiré periodicity and additional local disruptions (blue areas). The white arrow denotes the direction of the line cut in panel c. Scan parameters: $V = 1.1\,\mathrm{V}$, $I = 48\,\mathrm{pA}$, field of view $79 \times 79\,\mathrm{Å}$. **b**, The two-dimensional Fourier transform of the DC-STM image exhibits six characteristic peaks in a distance of $(1.7 \pm 0.1)\,\mathrm{Å}^{-1}$ from the centre. This corresponds to a real-space periodicity of $(5.9 \pm 0.3)\,\mathrm{Å}$. **c**, The height profile along the white line indicated in panel a shows height variations in the topography of 2 pm.

[Sch19b] or phonon scattering [Wal63]. Intriguingly, point defects can be purposely engineered by variation of the parameters within the preparation procedure of the sample. For instance, the choice of a suitable annealing temperature and sputtering the TMDC surface with positively-charged Helium ions has been shown to strongly change the optical emission of MoS_2 monolayers [Mit21]. By annealing our sample to over 300°C for several hours, most substitutional atoms that had replaced selenium are dissociated, resulting in vacancies in the upper and lower selenium layer (Fig. 3.4a, red marker). The dangling bonds of the surrounding three tungsten atoms (only two are visible in sketch in Fig. 3.4a, the third lies within the image plane) hybridise and form three localised, so-called defect states. One of these lies within the valence band and is thus doubly occupied with electrons [Bar19]. The energies of the remaining two defect levels, which were identified as two spin-orbit split states [Sch19b], lie within the bandgap. The lower unoccupied defect state exhibits a threefold symmetry, which is revealed by a DC-STM image recorded at $V = 380\,\mathrm{mV}$ (see Fig. 3.4b). The three bright areas of increased apparent height in the constant-current scan shall be

termed “lobes” in the following descriptions.
The effect of a single missing atom on the local density of states becomes clear when comparing spectra obtained via conventional scanning tunnelling spectroscopy (DC-STS) recorded on a Moiré peak of the pristine monolayer surface (see Fig. 3.4b, green marker) with spectra acquired on the lobe (Fig. 3.4b, blue marker). At the Moiré peak, the semiconducting nature of WSe_2 is confirmed by the emergence of a conduction band at 1070 mV (Fig. 3.4c, green curve). Once this voltage is reached, the detected tunnelling current increases considerably, leading to a peak in the differential conductance $\mathrm{d}I/\mathrm{d}V$. Interestingly, these measurements reveal a small but finite density of states even within the bandgap (see non-zero $\mathrm{d}I/\mathrm{d}V$ curve below 850 mV), which is usually not observed in scans of monolayers that rest on top of an insulating substrate. This phenomenon results from the density of states of the gapless gold substrate that is still noticeable through the TMDC monolayer [Sør14].
The spectrum recorded at the defect location shows two additional peaks arising at around 380 meV and 569 meV, respectively (Fig. 3.4c, blue curve). These are the defect-states mentioned earlier and labelled D_1 and D_2 for future reference. Additionally, a small bandgap renormalisation is apparent, shifting the conduction band onset from around 1070 mV to 1100 mV.
These spectra recorded in conventional STM operation reveal differences in the local conductivity on Ångström length scales. The localised defect states arising at a mono-atomic selenium vacancy present themselves as intriguing entities for further study.

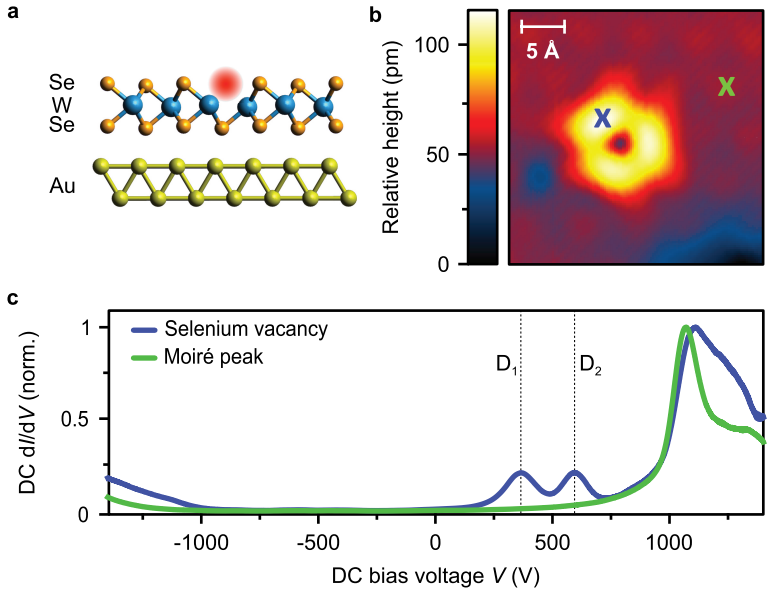


Figure 3.4 | Scanning tunnelling spectroscopy on selenium vacancies in tungsten diselenide monolayers. a, The side view of a WSe_2 monolayer adsorbed on gold displays a layer of tungsten (blue) covalently bound with two selenium (orange) layers. A selenium vacancy in the upper layer (red spot) considerably alters the electronic level structure locally. **b**, A threefold-symmetric structure clearly stands out against the background in a constant-current STM scan of a monolayer WSe_2 with a selenium vacancy. The background modulation results from a Moiré superlattice forming due to the lattice mismatch of gold and the monolayer WSe_2. The two coloured crosses mark the positions where the data shown in panel c were recorded. Scan parameters: $V = 380\,\mathrm{mV}$, $I = 49\,\mathrm{pA}$. **c**, Scanning tunnelling spectra obtained on the defect (blue curve) and on a Moiré peak of the surrounding unperturbed monolayer (green curve) reveal differences in the local conductance. At the positions of a selenium vacancy, additional levels between the Fermi energy and the conduction band onset arise (dotted lines labelled D_1 and D_2). Apart from the defect levels at around 380 mV and 569 mV, a small shift of the conduction band tunnelling onset to higher voltages is evident.

3.2 Exciting phononic motion in a monolayer WSe_2

As we have seen in the previous experiments presented in this thesis, a sub-cycle terahertz pulse impinging on the sample is able to excite ultrafast dynamics in molecular systems (see Chapter 2). In general, this method of inducing dynamics should be applicable to crystalline systems as well. To investigate this, we performed THz pump – THz probe experiments on the monolayer WSe_2, where the dynamics induced by the first pulse are detected by variations in the tunnelling currents driven by the time-delayed probe pulse.

The transient voltage these pulses possess in the near field is governed by the far-field strength. For the following experiments the field strengths of the pump and probe pulses were set to $E_{\mathrm{LW}}^{\mathrm{pump}} = 0.33\,E_0$ and $E_{\mathrm{LW}}^{\mathrm{probe}} = 0.6\,E_0$ in units of the unattenuated far field E_0. By varying the delay time τ between the two pulses and recording the lightwave-driven current $I_{\mathrm{LW}}^{\mathrm{pr}}$, ultrafast dynamics were uncovered, as becomes apparent in Fig. 3.5. A clear oscillation of the relative lightwave-driven current $I_{\mathrm{LW}}^{\mathrm{pr}} - \langle I_{\mathrm{LW}}^{\mathrm{pr}} \rangle$ with a period of approximately 3 ps, was observed on the WSe_2 monolayer (orange curve). Subtracting the quantity $\langle I_{\mathrm{LW}}^{\mathrm{pr}} \rangle$, the mean of the current in the whole measurement window, allowed us to focus on the oscillation dynamics. Importantly, such oscillations were not detected on a bare gold surface (red curve). The absence of the signal on bare gold rules out any measurement artefacts of the optics or electronic circuitry as the origin of the current oscillations. This points towards dynamics linked to the WSe_2 monolayer or even just the single defect excited by the pump pulse as the origin.

With the ultimate goal to determine the precise origin of the oscillatory pattern described above, we proceeded to scrutinise further characteristics of the selenium vacancy. For instance, we investigated the spatial dependence of amplitude and phase of the oscillation at positions around the selenium vacancy (tip locations marked in Fig. 3.6c). Figure 3.6a displays pump-probe data recorded at six different positions across the centre of the vacancy. Interestingly, the phase of the oscillations remains the same for all six positions. However, the amplitude critically depends on the exact location of excitation and/or readout (see Fig. 3.6b). Both amplitude and baseline of the oscillations show a marked maximum at the lobe position of

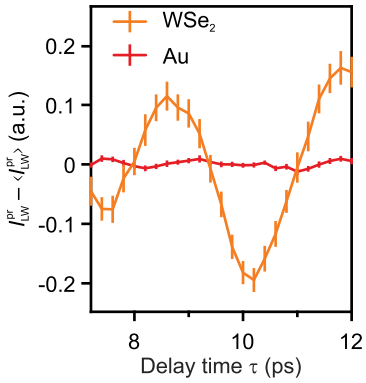


Figure 3.5 | Comparison of ultrafast current oscillations on a monolayer WSe_2 and on gold. We performed THz pump – THz-probe experiments at different positions on the WSe_2 sample. Measurements on the lobe of a single-atomic defect reveal clear oscillations of the lightwave-driven current $I_{LW}^{pr} - \langle I_{LW}^{pr} \rangle$ (orange curve). In contrast, on the gold surface (red curve), no oscillations are visible. Height setpoint: $V = 30\,\mathrm{mV}$, $I = 5\,\mathrm{pA}$, 2.9 Å approach.

the defect orbital (position 2) and decay towards positions further away from the vacancy. From the information in this panel alone, it is unclear whether the increased oscillation amplitude at the defect lobe only stems from a larger baseline current or indeed the mechanical displacement is larger there. As the line cut in Fig. 3.6d shows, the apparent height of the sample is increased dramatically at the lobe position. This larger current and thus the substantial increase in oscillation amplitude could arise from a smaller tip-sample distance, larger wave-function overlap, or higher local density of states. Additionally, possible origins of the larger oscillation amplitude could be an enhanced effect of the excitation pulse or stronger response to said pump pulse at this specific location.

As the oscillations in tunnelling current occur independent of the exact tip position, we expect the variation to predominantly result from periodic changes of tip-sample separation due to a mechanical motion excited by the pump pulse. In order to quantify the mechanical displacement necessary for the variation of the tunnelling current shown in Fig. 3.6a, we needed to extract the relation between tip-sample distance Δz and LW-driven tunnelling current I_{LW}.

As explained earlier, the tunnelling current decays exponentially with increasing tip-sample separation (see eq. 2.3), characterised by the decay constant κ. Supplementary Figure B.2 illustrates the experimental data and the analysis used to determine the decay constant. For two so-called retractions scans (I_{LW} vs. Δz), we obtained the two

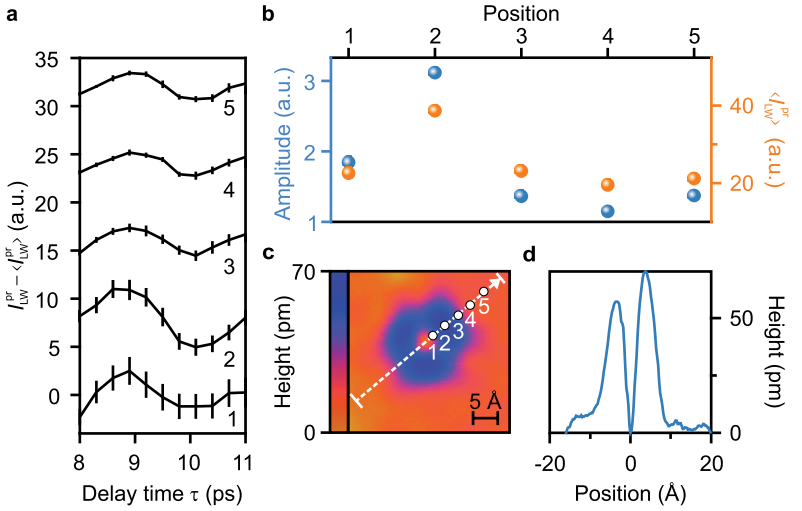


Figure 3.6 | Sub-nanometre positional dependence of current oscillation. THz pump – THz probe experiments were performed at five different positions on the monolayer WSe_2 sample (marked in panel c) with increasing distance to the vacancy centre. **a**, The tunnelling current $I_{\mathrm{LW}}^{\mathrm{pr}}$ driven by the probe pulse exhibits oscillations as a function of the pump-probe delay time at all positions but with strong variations depending on the exact tip location. The oscillation amplitude peaks at the lobe position (position 2) and then rapidly drops for positions further away from the defect. The averaged signal $\langle I_{\mathrm{LW}}^{\mathrm{pr}} \rangle$ of each trace is subtracted and the curves are vertically offset with respect to the trace below for clarity. The errorbars in panel a denote the standard deviation of the bootstrapped currents. **b**, The amplitude and mean signal critically depend on the tip position. A clear peak in both quantities appears on the lobe (position 2) followed by a rapid decrease upon moving further away from the vacancy. **c**, The conventional constant-current image of the defect ($V = 380\,\mathrm{mV}$, $I = 49\,\mathrm{pA}$) illustrates the tip positions during data acquisition for panel a, indicated by circular markers. **d**, A line cut through the image in panel c reveals two peaks of 57 Å and 67 Å in apparent height, respectively, symmetrical around the centre of the vacancy.

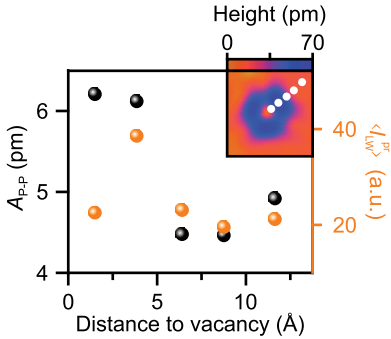


Figure 3.7 | Position-dependent oscillation amplitude. The amplitude of the oscillations observed in pump-probe experiments at five different positions (see inset) are converted into changes of tip-sample distance using the extracted tunnelling decay constant κ (see appendix B.3). The oscillation amplitude drops from 6 pm at the vacancy and the lobe of the defect state to below 5 pm for distances larger than 5 Å from the defect centre (black spheres, left y-axis). The mean current $\langle I_{\mathrm{LW}}^{\mathrm{pr}} \rangle$ over one oscillation period is plotted as a function of distance to the vacancy (orange spheres, right y-axis). The current $\langle I_{\mathrm{LW}}^{\mathrm{pr}} \rangle$ peaks at the lobe position and is approximately equally strong for all other positions. The inset depicts the DC constant-current image already shown in Fig. 3.6d. The measurement positions are marked by white points.

values for κ: 1.51 Å^{-1} and 1.59 Å^{-1}. Utilising the mean of the two decay constants $\kappa \approx 1.55$ Å^{-1} for further calculations, the oscillation amplitudes in Fig. 3.6 can be converted into picometres and plotted as a function of distance to the vacancy centre (Fig. 3.7, black spheres). Surprisingly, the amplitude variations do not simply scale with the mean tunnelling current driven at a each position (Fig. 3.7, orange spheres). Moreover, the oscillation amplitude of 6 pm extracted close to the vacancy is even larger than the height variations of the Moiré corrugation (see Fig. 3.3). Thus, it remains to be elucidated whether depending on the precise tip position the same pump pulse either leads to stronger excitation locally and/or the same excitation strength causes a larger mechanical deformation. The former could be due to a larger dipole formed by the increased electron density at the lobe position. The latter reasoning is supported by the fact that due to the missing selenium atom, the restoring force for mechanical deformations in the out-of-plane direction is reduced compared to the pristine semiconductor crystal.

To further encroach on possible mechanisms exciting the oscillation in the pump-

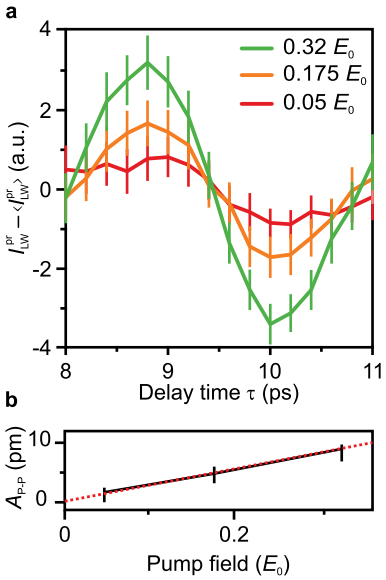


Figure 3.8 | Pump-field-dependent oscillation amplitude. a, We repeated pump-probe experiments for three different pump field strengths: $0.05\,E_0$ (red), $0.175\,E_0$ (orange), and $0.32\,E_0$ (green). The time-resolved lightwave-driven tunnelling current induced by the probe pulse reveals a clear trend of increasing oscillation amplitude, once a baseline in form of the time-averaged signal $\langle I_{LW}^{pr} \rangle$ is subtracted. The probe-pulse field strength was kept at $0.3\,E_0$. **b**, Utilising the known relation between the lightwave-driven current I_{LW} and the tip-sample distance Δz, the peak-to-peak amplitude $A_{P\text{-}P}$ of the oscillation for each field strength was obtained. The amplitude shows a linear dependence on the pump field, as indicated by a linear fit (red dashed line).

probe measurements, the field strength of the pump pulse was systematically varied. Figure 3.8a shows one period of oscillation between 8 ps and 11 ps for three different pump field strengths: $0.05\,E_0$ (red), $0.175\,E_0$ (orange), and $0.32\,E_0$ (green) when the tip is placed on the lobe (see Fig. 3.6c, position 2). The mean current $\langle I_{LW}^{pr} \rangle$ over the oscillation period was subtracted for simpler comparison of the oscillation amplitudes. With increasing pump field strength the amplitude clearly rises. Converting the current oscillations into changes of tip-sample distance leads to a linear increase of the amplitude $A_{P\text{-}P}$ with the pump field (see Fig. 3.8b). This points towards a field-driven excitation in contrast to a current-driven mechanism. The electric field of the pump pulse potentially couples to a dipole within the sample and acts as an ultrafast force displacing the monolayer from this equilibrium position and leading to the observed oscillations in tunnelling current.

It remained to be investigated whether this vertical displacement stems from a propagating or stationary displacement. To this end, we repeated the THz pump – THz probe experiments for different distances to a step edge within the sample. If the

excited atomic displacement propagated, it would be reflected off the step edge and interfere with the incident wave. We scanned for reflections or analogous features within the almost perfectly harmonic oscillation and observed no significant changes to the oscillation for different distances even for pump-probe delay times extending beyond 90 ps. In conclusion, we have established that a terahertz pump pulse is able to couple to the monolayer by means of its electric field to excite a vertical motion of the sample that does not propagate. This leads to an oscillatory behaviour of the monolayer as a function of pump-probe delay time, which is detected in the lightwave-driven tunnelling current. The extreme sub-nanometre-dependence of the amplitude and baseline of these oscillations on the exact tip position point towards either strong position-dependent coupling of the terahertz electric field to the sample or an increased excitation amplitude at the defect location.

3.3 Numerical calculations to elucidate the origin of current oscillations

In the pursuit to understand the underlying mechanisms leading to the spatial variations in pump-induced changes to the lightwave-driven tunnelling current, several proposals were discussed. The difference in electronegativity between the tungsten ($\chi = 2.36$) and selenium ($\chi = 2.55$) atoms of the monolayer expectantly leads to the formation of local partial charges, leaving the tungsten slightly positively charged. The incident terahertz electric field could indeed be able to couple to this dipole and excite a vertical oscillation. As we have found no propagation of the displacement along the surface of the monolayer and no damping of the vibration (see Suppl. Fig. B.3), the pump pulse most likely is exciting a standing-wave that is well-decoupled from the other collective excitations. However, the exact motion pattern cannot be extracted unequivocally from our measurements presented so far. In order to unravel the emergent phonon modes within the WSe_2/Au heterostructures, a comprehensive investigation was conducted employing first-principles calculations by our theory collaborators Maximilian Graml and Dr. Jan Wilhelm. The system was modelled as a pristine WSe_2 monolayer on a double layer of Au(111) acting as

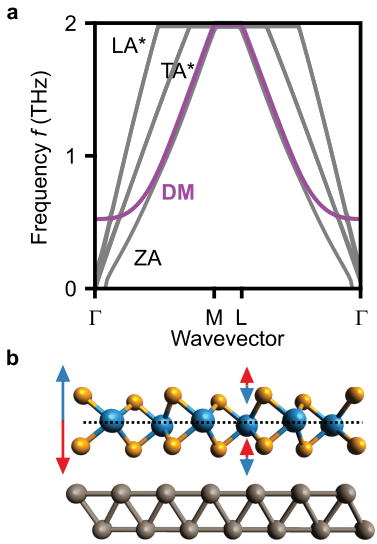


Figure 3.9 | Phonon dispersion of tungsten diselenide and drum mode. **a**, The phonon spectrum up to 2 THz shows the transverse (TA*, ZA) and longitudinal (LA*) acoustic branches as well as the drum mode (DM), which splits off from the ZA branch towards small crystal momenta. At the Γ-point, the DM exhibits a gap of approximately 0.5 THz. **b**, A real-space visualisation shows the atomic nuclei's motion in a pristine monolayer (selenium: yellow spheres, tungsten: blue spheres) when the DM is excited. Apart from a large centre-of-mass motion (long blue-red arrow) with respect to the fixed gold atoms (grey), there are some small out-of-phase intra-layer shifts of the selenium atoms (smaller blue-red arrows).

substrate. The tip was likewise represented as two layers of gold atoms. In order to take the lattice mismatch between Au(111) and WSe_2 and the resulting formation of a Moiré pattern into account, a large superstructure [Sar21] was constructed and arranged in a 4×4 supercell, comprising a total of 544 atoms. Employing density-functional theory with a Perdew-Burke-Ernzerhof exchange-correlation functional [Per96] and van der Waals interactions [Gri10], the total energies and forces on the atoms were calculated. The phonon band structure (Fig. 3.9a) was then determined using the "phonopy package" [Tog15]. For the frequency range of 0 THz to 2 THz, several acoustic modes are present in the plotted part of the Brillouin zone that is depicted in Fig. 3.9a ($\Gamma \rightarrow M \rightarrow L \rightarrow \Gamma$). Alongside the longitudinal (LA*) and transverse (TA*, ZA) acoustic branches, an unconventional phonon branch splits off the ZA branch at lower crystal momenta. This mode is gapped (approximately 0.5 THz) at the Γ-point of the WSe_2 monolayer Brillouin zone. This gapped mode, represented by the purple curve, corresponds to a drum-like collective vertical motion of the WSe_2 layer as a whole, in relation to the gold layers and thus shall be called drum mode (DM). The centre-of-mass motion (large arrow in Fig. 3.9b) is accom-

panied by an intra-layer vertical oscillation (small arrows in Fig. 3.9b), wherein the tungsten atoms exhibit an oscillation amplitude approximately 50% larger than that of the selenium atoms. The atomically strong electric field of the pump pulse may interact with the sample and excite DM through Coulomb interactions. For example due to charge transfer between the gold and WSe_2 monolayer, some charge imbalance, which can interact with the electric field of the pump pulse, is created. Besides that, the incident waveform is strong enough to induce polarisation by electrostatic induction [She22]. As a result, the terahertz field exerts forces on the atoms of the top atomic layers. The linear dependence of the amplitude on the excitation field, as shown in Fig. 3.8, suggests that the former mechanism dominates the interaction. Other phonon modes may also be excited due to the broadband excitation pulse (frequencies spanning 0.2−2 THz) but will decohere quickly due to dispersion. It must be noted that although the electric field of the pump pulse is predominantly oriented out-of-plane, weak in-plane vibrations may still be excited.

Thus, we have found that the electric field of the pump pulse could couple to charges in the monolayer, inducing a drum-like phonon mode. This periodic modulation of the tip-sample distance of a few picometres is detected by a time-delayed probe pulse driving ultrafast tunnelling currents. The strong positional dependence of the current baseline and the peak-to-peak amplitude shows that the exact atomically precise placement of the tip is critical. The single-atom vacancy disrupts the periodicity of the monolayer and the detected oscillations are strongest at the defect location. This raises the question whether the atoms surrounding the defect due to the reduced number of next-neighbours are more prone to excitation triggering the motion of the whole monolayer or whether the resulting oscillation amplitude is largest at this specific position again due to the lack of next-neighbours. Although the STM allows for atomic resolution in the tunnelling process, the electric field of the pump pulse is less strongly confined. As a result, the Coulomb interaction of the monolayer with the pump field is not restricted to the tip position. The readout mechanism however is strongly localised. Therefore, it is most likely that the positional variations of the detected signal stem from the differences in the response of the local nuclei to the drum mode.

Nonetheless, other effects could contribute to the variations in the current signal detected. Apart from changes to the tip-sample distance, energy levels within the

sample could be altered as a result of the pump pulse. Especially, the localised energy levels at the vacancy position (see Fig. 3.4, $\mathrm{d}I/\mathrm{d}V$ spectra) could respond to the field-induced vibration, shifting them in energy. In order to address this, a new ultrafast spectroscopy technique with combined sub-picosecond temporal, Ångström-spatial and few-millielectronvolt energy resolution must be developed.

3.4 Ultrafast snapshots of the local density of states

Alongside imaging on the atomic scale, accessing the position of energy levels is one of the fundamental elements of scanning tunnelling microscopy. The question arises whether similar information to conventional scanning tunnelling spectroscopy can be extracted utilising ultrafast biasing. This would allow for the study of changes to the local density of states of a sample as a function of time. As femtosecond imaging of select orbitals is possible with terahertz pulses, spectroscopy should be likewise. To test this hypothesis, we coupled terahertz pulses with varying field strength into the STM junction and recorded the resulting lightwave-driven tunnelling current I_{LW}. Figure 3.10a depicts the principle of operation: The transient voltage is tuned (red and orange curves) by varying the far-field strength of a terahertz pulse in units of the unattenuated far field E_0. Once the tip-confined voltage suffices to reach empty energy levels within the sample, state-selective tunnelling will set in. As the local spectrum extracted from DC-STS measurements differs considerably between the selenium vacancy (blue curve) and a Moiré peak on the pristine monolayer (green curve), the lightwave-driven current with its state selectivity is expected to display disparities as well. Indeed, as shown in Fig. 3.10b, for voltages within the semiconducting band gap, the lightwave-driven current remains meagre on the pristine WSe_2 spot (green curve, multiplied by a factor of 15) compared to the vacancy location (blue curve). This points towards an additional tunnelling channel at the defect location that is strongly localised and can be detected reliably by transient biasing of the junction. Numerically differentiating the ultrafast current I_{LW} by the peak field strength $E_{\mathrm{LW}}^{\mathrm{peak}}$ replicates the presence of two characteristic peaks (blue curve) on the defect lobe observed in DC-STS (see Fig. 3.10c). In stark contrast, minimal

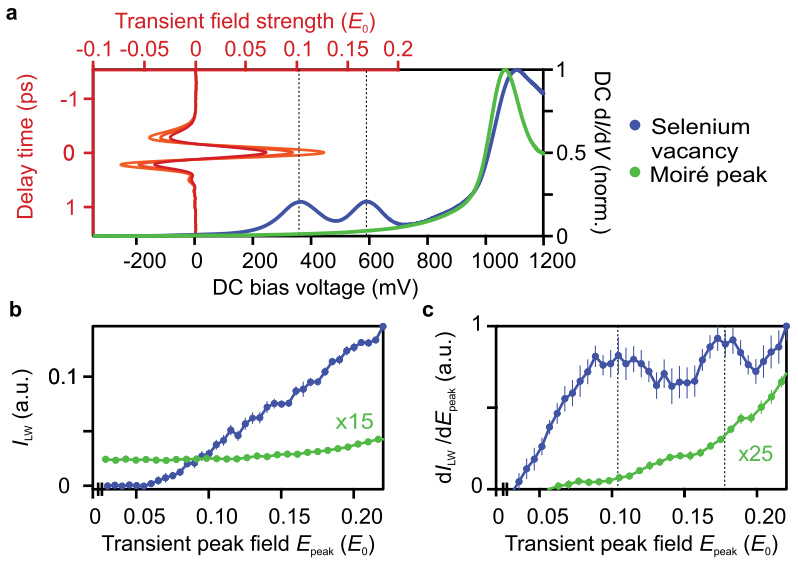


Figure 3.10 | Lightwave-driven scanning tunnelling spectroscopy. a, A sub-cycle terahertz pulse focused into the STM acts as an ultrafast bias voltage. Scaling the field strength of the far-field transient in units of the unattenuated far field E_0 (red to orange curves) will open state-selective tunnelling channels, depending on the field strength as well as the local density of states. The green and blue curves visualise the dI/dV spectra obtained from DC-STS for the tip positioned at the Moiré peak and the selenium vacancy, respectively (see Fig. 3.4b, crosses). The peaks of the defect levels are marked by dashed vertical lines. **b**, The lightwave-driven current varies strongly depending on the tip position. On the defect (blue), starting at around $0.06\,E_0$, the current rises strongly. In contrast, at the Moiré peak (green), the current is much lower and had to be multiplied by a factor of 15 for visibility on the same y-axis. **c**, Numerically calculating the derivative $\mathrm{d}I_{\mathrm{LW}}/\mathrm{d}E_{\mathrm{LW}}^{\mathrm{peak}}$ exposes two peaks on the selenium vacancy and no peaks on the undefected WSe_2 in spite of up-scaling the latter data by a factor of 25. Vertical dashed lines indicate the two hallmark peaks arising only at the defect location, which are reminiscent of the defect states seen in panel a. The error bars in panels b and c indicate the standard deviation of the bootstrap resampled current data and their $\mathrm{d}I_{\mathrm{LW}}/\mathrm{d}E_{\mathrm{LW}}^{\mathrm{peak}}$ spectra, respectively (see appendix A.4.1 on bootstrapping).

signal and no peaks are detected in the quantity $\mathrm{d}I_{\mathrm{LW}}/\mathrm{d}E_{\mathrm{LW}}^{\mathrm{peak}}$ on the Moiré spot up to the conduction band onset (Fig. 3.10c, green curve, data multiplied by 25). Thus, indeed, our lightwave-driven STS scheme (LW-STS) reproduces the two in-gap peaks seen in DC-STS and thus should enable the readout of the position of energy levels with sub-picosecond temporal resolution. To test this, it must be ensured that the data in Fig. 3.10c indeed represents the ultrafast equivalent of DC-STS.

Let us give some general thought to the temporal and energy resolution in LW-driven STS. The tunnelling of electrons is instantaneous, meaning that the applied instantaneous voltage of the waveform $V_{\mathrm{NF}}(t)$ governs the tunnelling probability $I_{\mathrm{LW}}^{\mathrm{inst}}(t)$ at each moment in time. This rapidly varying instantaneous current cannot be detected by the purely electronic STM current amplifier. Solely the time-integrated current acquired over multiple laser shots leads to a measurable signal. Notably, this total time-integrated current is not solely comprised of electrons tunnelling at $V_{\mathrm{LW}}^{\mathrm{peak}}$ because $V_{\mathrm{NF}}(t)$ ranges from its peak voltage $V_{\mathrm{LW}}^{\mathrm{peak}}$ down to negative voltages given by the near-field waveform. As a result, also lower-lying energy levels (lower than $V_{\mathrm{LW}}^{\mathrm{peak}}$) can contribute to the integrated current. This pollutes the signal and prevents unequivocal, straight-forward assignment of features detected in a I_{LW}–$V_{\mathrm{LW}}^{\mathrm{peak}}$ curve to true energy levels within the sample. Moreover, as this rectified current is driven by the entire waveform, the temporal resolution is limited to the full oscillating electric field, in our case almost 2 ps.

However, the tunnelling scenario is significantly different when considering energetically isolated levels or bands, where state-specific tunnelling can be actively induced [Coc16, Pel20b]. Thereby, the tunnelling window is confined in time to when the transient bias suffices to reach the sharp resonances of the sample and the tunnelling current is dominated by the current contribution where the instantaneous voltage is equal to the peak voltage $I_{\mathrm{LW}}^{\mathrm{inst}}(V(t) = V_{\mathrm{LW}}^{\mathrm{peak}})$. In the case of the examined WSe_2 monolayer sample, this means that the onset of tunnelling into the first defect level at the vacancy position and the onset into the conduction band on the pristine surface can indeed be attributed to temporally strongly confined tunnelling. Therefore, the derivative $\mathrm{d}I_{\mathrm{LW}}/\mathrm{d}E_{\mathrm{LW}}^{\mathrm{peak}}$ can be interpreted as an ultrafast readout of the local density of states when dealing with isolated, energetically sharp states.

Importantly, due to the asymmetrical shape of the near-field voltage $V_{\mathrm{NF}}(t)$ (asymmetry ratio of 2:1, see Fig. 2.14b) and the relative position of the Fermi level within

the band gap, the negative half-cycles of the near-field waveform prompt negligible tunnelling currents. Only once the voltages of the negative half-cycles reach the valence band below $-1\,\mathrm{V}$, a notable tunnelling current is driven from the sample towards the tip. Therefore, the detected tunnelling current in our scheme is solely dominated by currents driven into empty states of the sample as long as the peak voltage of the near-field waveform remains below $2\,\mathrm{V}$.
In light of this suitable energy level structure of defects in WSe_2, LW-STS may indeed provide information on the level structure on ultrafast timescales. As again the electronics are too slow to record the instantaneous current $I_{\mathrm{LW}}^{\mathrm{inst}}(t)$, a work-around has to be utilised in order to estimate the temporal resolution of our LW-STS scheme. A commonly used practice in ultrafast optics lies in performing an autocorrelation experiment. To this end, two weak terahertz waveforms with a field strength of $0.05\,E_0$ each are superimposed within the STM junction. Individually, the peak fields do not suffice to reach any resonances in the sample, however once constructively interfered, the peak voltage exceeds the onset of D_1. Figure 3.11a depicts the DC-STS scan up to the second defect level D_2 (blue line) and the two terahertz waveforms in red and orange. By varying the relative delay time τ and recording the terahertz-driven current, an autocorrelation curve is retrieved providing information on the temporal scales of the tunnelling events (see Fig. 3.11b). One can clearly detect an increase of tunnelling current once the two transients constructively interfere. The peak width of the autocorrelation curve, which is linked to the temporal resolution of our detection scheme, is well below a picosecond.
The width of the autocorrelation only provides a rough estimate for the temporal resolution. Instead, we are interested in the duration of the tunnelling window that a single terahertz pulse opens as this ultimately governs the temporal resolution of ultrafast LW-STS. To extract this value, as a first step a model simulating the ultrafast autocorrelation was set up. Subsequently, the instantaneous current driven by a single pulse can be simulated, providing insights in the temporal width of $I_{\mathrm{LW}}^{\mathrm{inst}}(t)$. Details on this work can be found in appendix B.1.
In order to model the lightwave-driven tunnelling current, two tunnelling contributions were taken into account. On the one hand, the tunnelling into the “background” consisting of pristine WSe_2 bound to the gold surface was assumed to be an instantaneous process without Coulomb blocking. This means that the electrons tunnelling

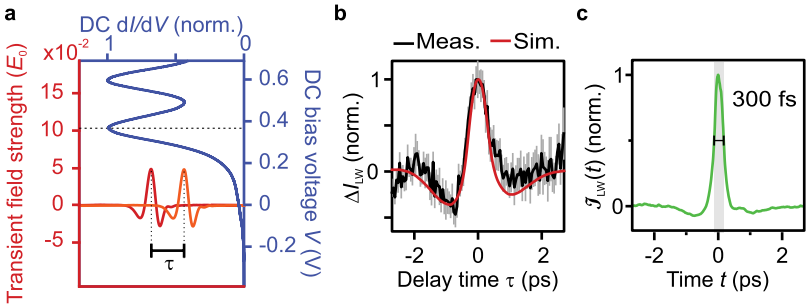


Figure 3.11 | Autocorrelation: experiment and simulation. To quantify the duration of the tunnelling window, we performed an autocorrelation experiment. **a**, Two terahertz pulses (red and orange waveforms) attenuated to 5% of the maximum peak field E_0 were focused into the STM junction with a mutual delay time τ. The DC-STS curve recorded at the defect position is plotted on the right y-axis (blue curve). The transient voltage applied by the individual pulses does not suffice to reach the defect level at around 350 mV. **b**, By interference of the two pulses, the effective field strengths may be sufficient to charge the defect, leading to a delay-time-dependent change of the tunnelling current ΔI_{LW}. The current baseline driven for non-interfering pulses was subtracted. The measured data (black with grey error bars) is closely reproduced by a rate equation model (red). The error bars indicate the standard deviation of the bootstrap resampled current data. **c**, The instantaneous current $\mathcal{I}_{LW}(t) = I(\mathcal{V}_{LW}(t))$ driven by a single transient voltage pulse that reaches the defect level $V_{\mathrm{LW}}^{\mathrm{peak}} = 260\,\mathrm{mV}$ is simulated. The FWHM of $\mathcal{I}_{LW}(t)$ (grey area) quantified as 300 fs is commonly identified as the temporal resolution of LW-STM at a certain peak field [Coc13, Coc16].

from the tip into the monolayer are mobile enough to immediately propagate away without influencing subsequent tunnelling processes. Electrons tunnelling into the defect states, on the other hand, remain there for a characteristic lifetime, effectively hindering the flow of further charge carriers into the respective defect state. This lifetime was included in the form of a rate equation model with an exponential decay time of the electron population in the defect states. Additionally, the shape of the voltage pulses had to be determined for the simulation. Using photo-assisted tunnelling (see appendix B.6 for details), the exact temporal evolution of the voltage transient was determined and the autocorrelation could be simulated.

Figure 3.11b illustrates the comparison of the autocorrelation simulation with the experimental data. Notably, the agreement between the autocorrelation measurement and the simulation including a lifetime of 5 fs in the rate equation model points towards unperturbed probing of the lifetime-broadened first defect level. The agreement implies that perturbations by multiple electrons tunnelling or temporary charging of the defect level can be disregarded.
Building on the similarity of simulated and measured autocorrelation curves, the tunnelling current driven by a singular terahertz pulse was simulated. The instantaneous current was determined by solving the rate equation with a lifetime of 5 fs and a single electron tunnelling per pulse for a terahertz pulse with its electric field tuned to the first defect level. At this particular peak voltage of $V_{\mathrm{LW}}^{\mathrm{peak}} = 260\,\mathrm{mV}$, the instantaneous current possesses an FWHM of 300 fs. The width of this trace is commonly associated with the temporal resolution in LW-STM [Coc13, Coc16]. Therefore, by tuning the transient voltage pulse to an isolated level, we are able to perform STS of a single defect level with unprecedented femtosecond resolution. The only limiting factor in time resolution is given by the duration of the voltage pulse and therefore the far-field waveform.
With the knowledge that lightwave-biasing is indeed able to drive tunnelling currents with ultrafast temporal resolution, the conversion factor between the far-field strength and the voltage across the gap must be determined. It must be confirmed that indeed the position of the peaks in the $\mathrm{d}I_{\mathrm{LW}}/\mathrm{d}E_{\mathrm{LW}}^{\mathrm{peak}}$ spectra be equated with the defect states detected via conventional STS. Recording femtosecond STM images at various transient voltages, we provide evidence that the characteristic features observed in the LW-STS spectra link to the orbital structure of the corresponding defect states (see Fig. 3.12). More precisely, we compared STM images of the selenium vacancy in WSe_2 attained under static conditions (Fig. 3.12, top row) with those acquired under similar terahertz biasing (Fig. 3.12, bottom row). For low bias voltages (below the first defect level), both the ultrafast STM image ($E_{\mathrm{LW}}^{\mathrm{peak}} = 0.07\,E_0$) and the static STM image recorded at a DC bias voltage ($\mathrm{V_{DC}}$) of 210 mV exhibit a nearly circular contrast (Fig. 3.12a). As $E_{\mathrm{LW}}^{\mathrm{peak}}$ is increased and exceeds $0.1\,E_0$, the LW-driven STM image distinctly reproduces the threefold symmetry observed in the defect levels (Fig. 3.12b). Above the conduction band edge, due to the missing surface atom, the vacancy causes a depression in the signal compared to the unde-

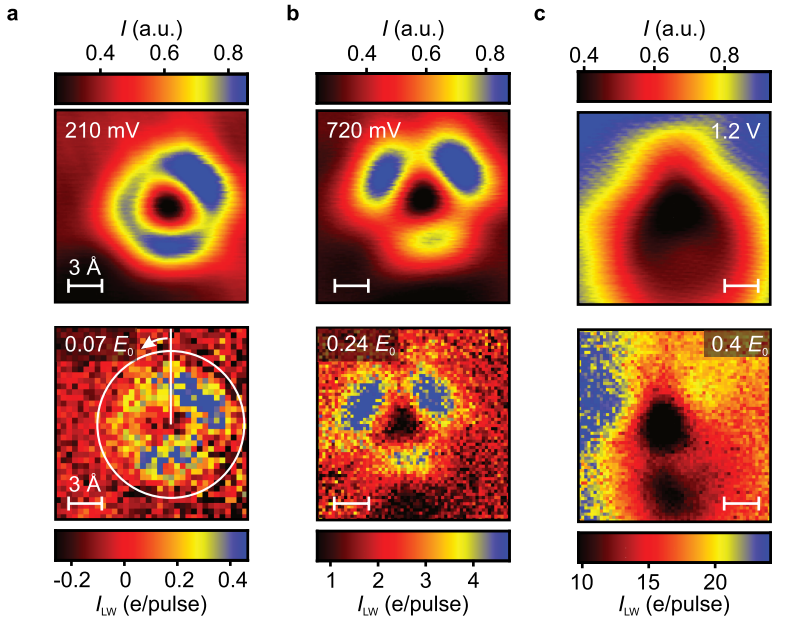


Figure 3.12 | Comparison of DC and ultrafast imaging. Tunnelling images of the selenium vacancy for different bias voltages reveal characteristic appearances. The DC constant-height images and lightwave-driven femtosecond snapshots are plotted in the top and bottom row, respectively. **a**, For bias voltages below the characteristic peak observed in STS (DC-STM: $V = 210\,\mathrm{mV}$; LW-STM: $E = 0.07\,E_0$), a circular shape with two slightly brighter areas is discernible. **b**, For higher voltages ($V = 720\,\mathrm{mV}$; $E = 0.24\,E_0$), a threefold-symmetric orbital structure emerges. **c**, At even higher bias voltages above the conduction band ($V = 1.2\,\mathrm{mV}$; $E = 0.24\,E_0$), the defect location is dominated by a dark area indicating a depression. In all panels, the white scale bars denote a length of 3 Å. Scanning parameters of DC constant-height images: $V = 700\,\mathrm{mV}$ (a) and $V = 1.2\,\mathrm{V}$ (b,c), $I = 49\,\mathrm{pA}$. Setpoint of lightwave-driven images ($V_{\mathrm{DC}} = 0\,\mathrm{V}$): $V = 30\,\mathrm{mV}$, $I = 49\,\mathrm{pA}$, 1.7 Å approach. The colour bars encompass the currents within the 3$^{\mathrm{rd}}$ and 97$^{\mathrm{th}}$ percentile of the current distribution. To improve the signal-to-noise ratio, the LW-STM images are averaged over two (panels a and b) and ten (panel c) repetitions. The white markers in the bottom row of panel a indicate the integration limits for the analysis plotted in Suppl. Fig. B.1.

fected WSe_2 (Fig. 3.12c). The change in contrast and apparent shape of the vacancy becomes even more evident when analysing the signal within the images in radial and azimuthal directions (see Suppl. Fig. B.1).
With the agreement between DC and LW-driven images as verification, we can claim that indeed the peaks arising in the transient scanning tunnelling spectra (see Fig. 3.10c) are linked to defect level tunnelling. Furthermore, we have obtained a calibration factor to accurately map the field strength to a transient voltage across the junction. By using the ratio between $E_{\mathrm{LW}}^{\mathrm{peak}}$ and V_{DC}, we arrived at a terahertz peak voltage for the maximum incident field E_0 of $V_0 = (3.0 \pm 0.1)\,\mathrm{V}$.
In summary, the autocorrelation simulation model confirms that tunnelling is confined to a short time window, allowing us to extract the local density of states in a purely lightwave-driven biasing scheme with a temporal resolution of 300 fs in proximity to the first defect level.
Additionally, now that a calibration factor is obtained for the conversion from the far-field strength to the peak voltage dropping across the tunnelling gap, the results obtained in section 3.2 can be analysed further. In those experiments, we saw that following a pump pulse with field strength $E_{\mathrm{LW}}^{\mathrm{pump}} = 0.33\,E_0$, ultrafast changes to the detected current driven by a probe pulse with field strength $E_{\mathrm{LW}}^{\mathrm{probe}} = 0.6\,E_0$ were registered. Using the conversion factor derived above, these field strengths convert to $V_{\mathrm{LW}}^{\mathrm{pump}} = 1\,\mathrm{V}$ and $V_{\mathrm{LW}}^{\mathrm{probe}} = 1.8\,\mathrm{V}$, respectively. This means the monolayer was excited close to and probed far above the conduction band onset. Even placing the STM tip above the defect lobe, the current detected was therefore most likely dominated by electrons tunnelling into the semiconducting conduction bands. By probing far above the conduction band onset, the oscillations in tunnelling current can indeed be attributed to variations in tip-sample distance as postulated in section 3.2. The question arises to which extent energy level shifts may be induced by the pump pulse. Specifically, it is compelling to understand how lower-lying energy levels, like the localised defect levels, may contribute to the total signal and how they are modified by the motion of the monolayer.
Thus, since LW-STS is able to retrieve the electronic spectrum of the energetically isolated lower defect level with a temporal resolution of 300 fs, we possess the necessary means to explore ultrafast changes to energy levels following an external stimulus.

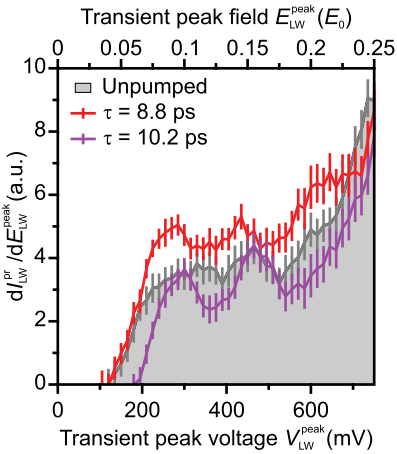


Figure 3.13 | Time-resolved LW-STS. Pumped ultrafast tunnelling spectra were recorded at delay times τ corresponding to a maximum ($\tau = 8.8\,\text{ps}$, red curve) and a minimum ($\tau = 10.2\,\text{ps}$, purple curve) of the pump-induced current oscillations shown in Fig. 3.5. The LW-STS data for an unpumped sample is illustrated as a grey area. Depending on the pump-probe delay time, the defect energy levels shift considerably in both the horizontal and vertical direction. The data reveals ultrafast changes to the local density of states. The pump field was set to $0.32\,E_0$ for all depicted measurements.

3.5 Exciting and monitoring ultrafast shifts of single-defect energy levels

The innovative ultrafast spectroscopy technique described above allows for the visualisation of the local density of states with sub-Ångström and sub-picosecond resolution. Moving even further from mere observation towards enabling the targeted manipulation of electronic levels on a sub-molecular level, we aimed to investigate how lattice vibrations may modulate the energy levels of the WSe_2 single-atomic defect.

In a previous section (see sec. 3.2), I have introduced a possible explantation for the oscillations in the probe-induced tunnelling current following a strong pump pulse. A drum mode excited by the first pulse periodically modulates the tunnelling gap. However, it remains unclear whether this mode predominantly simply varies the tip-sample distance or whether more intricate changes to the sample, such as the shifting of energy levels, occur. To disentangle this, we employ the novel concept of time-resolved light-wave driven scanning tunnelling spectroscopy.

To obtain data elucidating the energy spectrum, comparable to the data presented in Fig. 3.10, the probe pulse was scaled in field strength and the resulting tunnelling current detected for each field strength. Upon numerically calculating the derivative

$\mathrm{d}I_{\mathrm{LW}}/\mathrm{d}E_{\mathrm{LW}}^{\mathrm{peak}}$, we arrived at the ultrafast equivalent of DC scanning tunnelling spectroscopy, allowing for statements on the energy of specific states. In contrast to the data presented in Fig. 3.10, the probe pulse was preceded by an excitation pulse driving the DM. We repeated the recording of a LW-driven spectrum with a preceding pump pulse for two different delay times – 8.8 ps and 10.2 ps – corresponding to a maximum and minimum, respectively, of the current oscillation illustrated in Fig. 3.5. As we were most interested in shifts of the defect levels, we positioned the STM tip at the lobe position of the single-atom vacancy (see Fig. 3.13a, white marker). Stark changes between pumped and unpumped spectra become apparent (see Fig. 3.13b). Vertical shifts of the spectrum occur, which can be attributed to a change of wavefunction overlap of tip and sample due to mechanical motion. Crucially, in addition to an overall rescaling, the time-dependent local density of states associated with the vacancy exhibits noteworthy changes in its spectral characteristics. Specifically, the onset of tunnelling into the first defect level changes considerably with pump-probe delay time, as indicated by horizontal shifts of the rising edge of the defect level D_1. While the spectrum recorded at a delay time of 8.8 ps (red curve) closely resembles the unpumped spectrum (grey shaded area), the onset into D_1 for 10.2 ps (purple curve) is shifted to higher peak fields by approximately $0.015\,E_0$. With the known conversion factor from far-field strength to transient peak bias extracted previously, the shifts of the D_1 onset can be quantified as approximately 45 meV. As the LW-STS scheme is most precise for the investigation of stand-alone levels, we focused all quantitative conclusions on the onset into the D_1 level.

The similarity of the spectrum at a delay time of 8.8 ps and the unpumped case is highly intriguing. Examining the level onset for various excitation field strengths shed more light on the mechanisms. Fig. 3.14 displays the spectra taken on the defect lobe (marked spot in Fig. 3.14a) for three different pump field strengths and two delay times 8.8 ps and 10.2 ps. In the case of $\tau = 10.2$ ps, the spectrum at a low pump strength ($E_{\mathrm{LW}}^{\mathrm{peak}} = 0.05\,E_0$, below D_1) resembles the unpumped LW-driven spectrum. With increasing field strength ($0.175\,E_0$ and $0.32\,E_0$), the onset monotonically and almost linearly shifts to higher energies. For the other delay time of $\tau = 8.8$ ps, however, completely contrasting behaviour is detected. Moderately strong fields (0.05 and $0.175\,E_0$) shift the rising edge of D_1 to higher energies. Meanwhile, the pumped spectrum for the highest pump field of $0.32\,E_0$ is shifted back – resem-

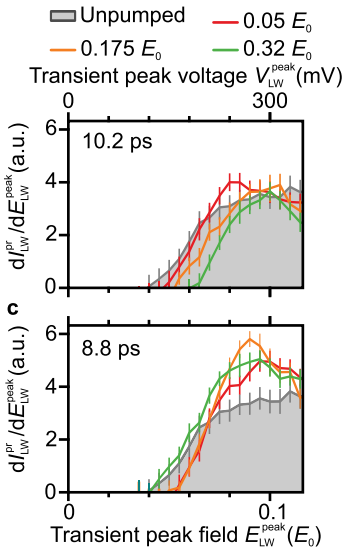


Figure 3.14 | Dependence of level-shift on pump-field strength. We investigated the effect of three different pump field strengths on the ultrafast spectra for the delay times 10.2 ps (upper panel) and 8.8 ps (lower panel). For a delay time of $\tau = 10.2$ ps at a low pump field strength ($E_{\mathrm{LW}}^{\mathrm{peak}} = 0.05\,E_0$, red curve), the tunnelling onset resembles the unpumped case (grey area). With increasing pump strengths ($E_{\mathrm{LW}}^{\mathrm{peak}} = 0.175\,E_0$, orange curve and $E_{\mathrm{LW}}^{\mathrm{peak}} = 0.32\,E_0$, green curve), the onset of D_1 shifts to higher probe fields (equivalent to voltages, see top x-axis). In contrast, for $\tau = 8.8$ ps the shift only occurs for low pump strengths ($0.05\,E_0$, red curve and $0.175\,E_0$, orange curve). The spectrum for a pump field of $0.32\,E_0$ (green curve) resembles the unpumped case.

bling the static one. This leads to the conclusion that not a single pump-induced effect is harmonically modulating the energy level position, but a complex interplay of several competing excitations is shifting the defect level energy periodically in time, depending on the field strength.

As shown earlier in this chapter, a drum-like vertical mode is most likely excited by the pump pulse. In first approximation this modifies the tip-sample distance. Notably, as the field-enhancement the far-field transient experiences during its propagation depends on the size of the gap. Thus, modifications of the junction size could influence the near-field peak voltage. However, the atomic displacements of 10 pm we determined from the current oscillations in Fig.3.8b) are a factor of 50 smaller than the equilibrium tip-sample distance of approximately $d \approx 5$ Å at which we performed these experiments. Such small variations barely influence the field enhancement, which scales with $\frac{1}{d}$. Moreover, variations in ultrafast peak voltage would also lead to delay-dependent shifts of the conduction band onset, which we do not observe.

Yet, the vertical displacement does indeed change the tunnelling probability due to

centre-of-mass motion of the whole monolayer. Besides, the adhesion of the bottom selenium layer to the gold substrate could lead to additional intralayer changes of atomic distances.

3.6 Numerical calculations on the energy level shifts

Revisiting the density-functional theory model of the sample, possible origins of the observed energy shifts can be perceived. Thus, the effect of atomic displacements following DM on the energy spectrum at the defect location must be closely examined. The most dominant motion, the centre-of-mass oscillation, periodically changes the position of the monolayer within the tunnelling gap by Δz_{Se} and thus also the Coulomb interaction of the defect with its image charge in the tip as well as the substrate (see appendix B.5 for details). With the peak-to-peak displacement of 10 pm extracted from measurements (see Fig. 3.8b), this leads to energy level shifts on the order of 10 meV. However, contrary to the experimental observations, this energy shift is symmetric about the equilibrium level position according to the DFT calculations, meaning that the energy shift for 8.8 ps and 10.2 ps should be equal in magnitude. The fact that the onset into D_1 for $\tau = 8.8$ ps resembles the unpumped case points towards a second competing effect. The observed phenomenon can potentially be explained by considering the effect of intra-cell distortions on the electronic level structure. Orbital hybridisation could be slightly affected by the variations in atomic separation of tungsten and selenium atoms. Moreover, alterations in intra-layer distance are anticipated to affect the dielectric function. In the simplistic picture of describing the defect as a hydrogen-like impurity, these changes may result in the renormalisation of energy levels. Besides that, as the selenium atoms of the top and bottom layer are displaced with different amplitude, slight modifications to spin-orbit coupling are expected to occur. As a result, the two spin-orbit split defect levels D_1 and D_2 are expected to shift in energy.
The combined influence of the mirror charge effect with modifications of the energy spectrum due to intra-unit cell motion could therefore provide a reasonable explanation for the observed shifts in defect levels.

In conclusion, we were able to extend the remarkable potential of LW-STM with its Ångström, single-electron, and femtosecond precision to include an energy resolution on the scale of 10 meV in time-resolved LW-STS. Leveraging this astounding precision that LW-STS can provide, opens up a wealth of opportunities for scientific exploration. In this study, we harnessed these capabilities to observe the dynamic evolution of energy levels within a single selenium vacancy following excitation. Our findings unveil a fascinating phenomenon: The excitation of an acoustic, predominantly vertical, drum-like mode induces an adiabatic shift in the first defect level. Astonishingly, this shift exhibited unipolar and non-monotonic behaviour, which can be attributed to intricate lattice distortions. This groundbreaking discovery paves the way for ultrafast spectroscopy at atomic length scales, equipping us with the essential tools to unravel the mysteries of local electron-phonon coupling. Our technique extends its reach to previously challenging nonlinearities and coherences [Ger17, Na19]. Additionally, with the capability of pulse shaping in our pump pulse, we have unlocked new avenues for exciting specific, atomically localised vibrations. This empowers us to potentially transiently control many-body electronic states by precisely adjusting their energy positions as needed.

3.7 Preliminary data on optical excitation of WSe_2 monolayers and homo-bilayers

In this and the previous chapters, I have shown that ultrafast excitation of single molecules and atomic defects using terahertz pulses can induce intriguing dynamics in various sample systems. Removal of a single electron from the HOMO of a pentacene molecule led to the alteration of the adsorption potential, activating a vertical motion of the molecule. More elaborate triggering of in-plane vibrations was achieved by directly interacting with select atoms of a MgPc molecule by means of an ultrafast atomic-scale force pulse. Shifting the focus from single molecules to crystalline materials, I showed that the electric field of a pump pulse is also able to excite a vertical drum-like mode in a WSe_2 monolayer. This non-propagating mode periodically modulated the tunnelling current driven by a probe pulse for more than

90 ps after the excitation. Introducing an energy resolution of 10 meV, the effect of the drum mode on the lower localised energy level of a single-atomic vacancy within the top layer was investigated. Through an intricate combination of several competing effects, the energy level shifts by up to 45 meV in a non-trivial manner on picosecond timescales.

In all these experiments, a terahertz pulse was used for ultrafast excitation. A natural consequential question is how the sample would react to excitation pulses in the visible spectrum of electromagnetic radiation. There, the Keldysh parameter (see eq. 2.8) is no longer below one, indicating a more particle-like behaviour of the laser light. The photon energies of visible light can exceed the bandgap in a semiconductor or the HOMO-LUMO separation in a molecule, which strongly perturbs the sample's electron population once absorbed. Numerous relaxation pathways of the excited electrons may include complex electron-phonon or electron-electron interaction. Previous work has shown that in TMDC heterobilayers the twist angle provides a strong handle on the Coulomb correlations of excitons [Mer20]. Alongside defect engineering [Kas19, Bar19] and strain manipulation [Ye19], the ability to locally tailor the properties of quasiparticles opens up avenues for exploring uncharted phases of matter within a broad range of van der Waals heterostructures.

In Chapter 3, I have so far demonstrated that by injecting a terahertz pulse into the tip-sample junction, the field of the pulse itself is able to excite a WSe_2 monolayer, which alters the vibrational state and affects the electronic system. Taking the inverse approach, one could excite the electronic structure of a sample and investigate the effect on the motion of atomic nuclei. To do so in LW-STM, an ultrashort pump pulse with sufficient photon energy is necessary to promote electrons to higher-lying energy levels followed by a probe pulse to inspect the local conductivity of monolayer and bilayer WSe_2 samples.

To this end, a source of laser pulses in the visible regime was set up. Figure 3.15a depicts the broadband spectrum created by supercontinuum generation in sapphire crystal with a thickness of 5 mm. Subsequently, the spectrum is compressed and cut, resulting in pump pulses with a spectrum centred around 750 nm (see yellow curve in Fig. 3.15a). The pump photon energy of 1.6 eV suffices to optically excite electrons into the conduction band of a WSe_2 bilayer. The unbound electrons and holes are expected to subsequently relax toward their respective band minima and

potentially form excitons near the fundamental energy gap [Wan12, Uge14]. In few-layer TMDCs, the strong Coulomb interaction and confinement to two dimensions should assist fast exciton formation. This has indeed been reported on in pump-probe studies, which have shown exciton formation times on sub-picosecond timescales in monolayer WSe_2 [Ceb16]. Apart from creating a hot electron–hole plasma, we can resonantly excite the A-exciton. As Fig. 3.15b shows, while the A-exciton peak in absorbance measurements blueshifts for a decreasing number of layers (1.62 eV in the quintuple layer to 1.67 eV in the monolayer) [Zha13], it remains well within our pump spectrum (grey background).

Previous work has closely examined the relative population of free hot charge carriers and bound excitons over time for various pump photon energies and extracted exciton decay times on the order of several picoseconds [Ste17]. In Steinleitner and colleagues' work [Ste17], the samples were created by mechanical exfoliation of bulk crystals and subsequent transfer onto a diamond substrate. As the TMDC layers in the work I am presenting, are directly placed on a Au(111) surface, we can expect shorter carrier lifetimes due to the high conductivity of the metal.

The pulse energy of 1.3 nJ used to optically excite the sample corresponds to 5.2×10^9 photons per pulse. The average effective optical absorption of a monolayer WSe_2 over the pump spectrum amounts to approximately 7% (16% at resonance) [He14, Aro15]. In combination with the spot size of 50 µm, this leads to an injected carrier density of $1.85 \times 10^{13} \frac{1}{\mathrm{cm}^2}$. The excited electron–hole plasma as well as the potential exciton population is expected to modify the local conductance measurably. In order to disentangle effects caused by photoemission and photo-assisted tunnelling of carriers out of the tip from true pump-induced changes of the tunnelling current, the same measurements were repeated for different tip-sample distances: with retracted tip, which corresponds to a distance of around 130 nm, and in tunnelling contact (Ångström-scale gap). The normalised results for the overlap region of optical pump and terahertz probe are plotted in Fig. 3.16a. Negative delay times correspond to the terahertz pulses arriving before the optical excitation. In the pure emission regime (tip-sample distance > 100 nm), the time-dependent current traces the near-field waveform of the terahertz transient (see appendix B.6) as the photoemitted charge carriers are accelerated by the instantaneous terahertz field. This leads to a time-overlap dependent photocurrent. The pump-probe signal in the tunnelling

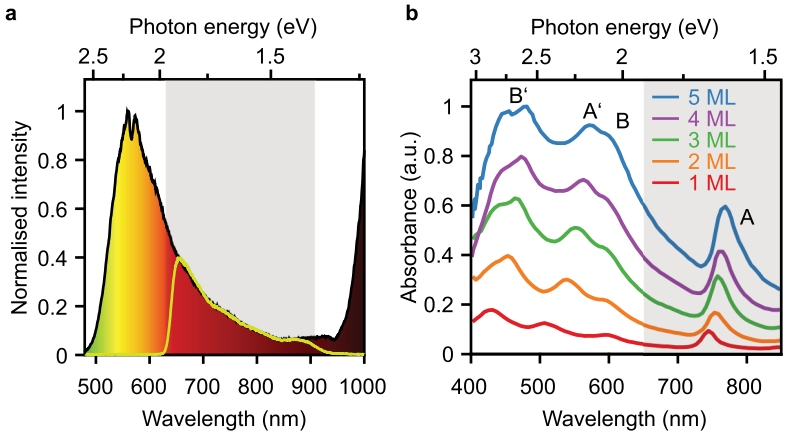


Figure 3.15 | Pump spectrum and TMDC absorbance. a, For optical excitation, a broad spectrum is at our disposal. This so-called supercontinuum generated in sapphire spans the visible spectrum from 530 nm up to the fundamental wavelength (1030 nm). To optically pump the system, only a fraction of the full spectrum was cut out and compressed in a prism compressor (yellow curve). The photon energies in this pump pulse span approximately 1.33 eV to 1.9 eV (grey shaded area). **b**, The absorbance of WSe_2 few-layer samples changes gradually from the mono- to the quintuple-layer case. Specific exciton resonances are labelled A, B, A', and B'. Our pump spectrum is visualised as a grey background, indicating the possibility for resonant excitation of the A-exciton. The absorbance spectra are reproduced from Zhao *et al.* [Zha13].

regime displays some interesting additional feature, which is absent in the emission regime. Even for delay times larger than 3 ps, where the emission current has returned to below the noise level, a negative signal persists in tunnelling current (see Fig. 3.16b). This long-lived current signal decays exponentially with a decay constant of 13 ps.

In order to encroach on possible origins of the long-living negative current, the experiment was repeated on a sample consisting of a monolayer WSe_2 exfoliated on gold. As Fig. 3.17 illustrates, on this sample the optical pump – THz probe data immediately returns to zero after the last oscillation cycle of the terahertz pulse

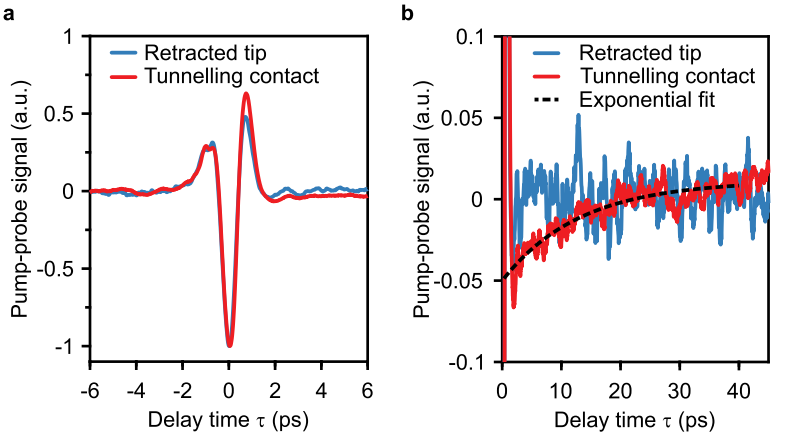


Figure 3.16 | Delay-time-dependent change of the current in emission and tunnelling regime. a, The photocurrent signal as a function of the pump-probe delay time is plotted from −6 ps to 6 ps. The two curves recorded with a retracted tip (blue curve) and in tunnelling contact (red curve) differ for positive delay times. **b**, We performed the same experiment for larger delay times up to 45 ps. In the case of a retracted tip (blue curve), the current immediately decays to zero after the ultrafast bias has returned to zero ($\tau > 3$ ps). In tunnelling contact, a negative current up to more than 30 ps is observed (red curve). An exponential fit to the data reveals a decay constant of 13 ps (black dashed curve).

has passed. Unfortunately, the signal-to-noise ratio is considerably lower in the data on the monolayer. Possible reasons include an unstable STM tip or a higher noise coupled into the signal due to mechanical fluctuation of the STM scan head.
Considering the gained insights regarding the signal decays, perhaps the visible pump pulses create a population of free charges, which after relaxation to the band extrema may form intra- and interlayer excitons. However, the large decay constant of 13 ps observed on the bilayer sample is inscrutable. Prior research has indicated that exciton radiative decay times in monolayers of TMDCs span from a few hundred femtoseconds to one picosecond [Wan18]. The lifetime of intra-layer excitons in bilayer WSe_2 is expected to be considerably longer due to the larger separation leading to higher binding energies compared to intra-layer excitons in monolayers

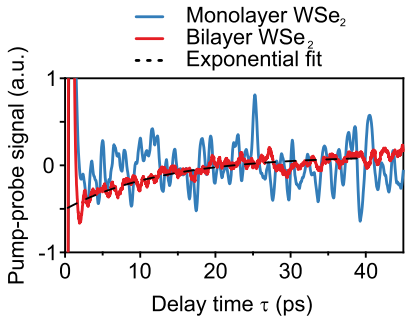


Figure 3.17 | Comparison of monolayer and bilayer WSe_2 band structures. We performed pump-probe experiments with a visible femtosecond pulse on a monolayer and bilayer WSe_2 sample. On the monolayer, no clear signal is visible immediately after the ultrafast transient voltage has returned to zero ($\tau > 3\,\mathrm{ps}$, see Fig. 3.16a), while a long-living signal is observed on the bilayer sample.

[Mer19a, Kip23]. The fact, that the negative current signal does not appear on a monolayer could point towards an involvement of intra-layer excitons in the long-living signal observed on the bilayer WSe_2 sample. Previous studies have suggested that the creation of excitons in atomically-thin TMDCs through optical pulse excitation induces a significant band gap renormalisation, resulting in a reduction of the band gap energy by shifting the valence and conduction bands [Che15]. Such shifts of the WSe_2 valence could lead to measurable changes in the ultrafast currents and explain the observed signal decay.

To arrive at more definitive conclusions, additional experimental data is required. If valence band shifts are the main contribution to the signal we observe, variations of the terahertz field strength could provide information on their magnitude. Furthermore, shifts of the conduction band edge would be similarly detectable by flipping the polarity of the terahertz probe pulse. Gradually changing the photon energy of the pump pulses would clarify at which energies these signals arise, further illuminating the possible origin.

The experiments presented in this section demonstrate that utilising optical pump pulses in LW-STM, we are able to detect light-induced changes of the tunnelling probability that persist for up to 20 ps after the pump. These first experiments illustrate the capabilities of LW-STM in resolving dynamics of charge carriers and potentially long-living intra-layer excitons in two-dimensional semiconductors. This opens the door towards mapping changes in exciton decay times and decay paths depending on local variations. Local potential minima such as defects or Moiré corrugations [RA18]

are expected to facilitate exciton trapping. Besides that, twisted heterostructures of TMDCs open an exciting playground for investigation of exciton dynamics at the nanometre and picosecond scale.
The possibility to excite the system from equilibrium using either single-electron charging, sub-cycle forces, or femtosecond optical pulses paves the way for the investigation of a plethora of materials and their response to external stimuli.

Chapter 4

Femtosecond LW-STM: towards the visualisation of intramolecular motion

Combining STM with ultrashort laser pulses in lightwave-driven STM has allowed for the extremely precise characterisation of the nanoscopic dynamics of single molecules (Chapter 2) and mono-atomic defects in semiconductors (Chapter 3). As described above, librations of a molecular rotator and the collective motion of a WSe_2 monolayer with periods in the picosecond regime could be excited and monitored. The low-frequency phonon we excited is only one of many modes that are present in the monolayer. As Fig. 4.1 displays, a multitude of phonon modes with varying energies and momenta can exist within the atomically-thin crystal. The calculated rich phonon band structure spanning from a frequency of 0 THz to 10 THz reveals the limitation of our LW-STM. The higher-frequency phonons in TMDCs cannot be studied within our LW-STM setup due to the temporal resolution limited to a few hundred femtoseconds. Thus, the LW-STM technique based on terahertz pulses as ultrafast bias voltage described in the previous chapters can only provide insights on phonons and their effects slower than 2 THz (see orange shaded area in Fig. 4.1). We have shown that these collective vibrations of the lattice strongly change the electronic levels within the sample. Aside from that, they also interact with the charge carriers through electron–phonon interaction, thus governing electron and exciton dynamics. However, phonons are only the tip of the iceberg of the plethora of fascinating physical phenomena that transpire within the temporal realm of only a

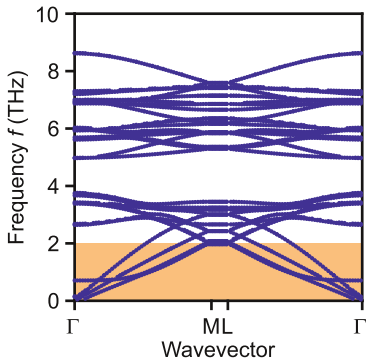


Figure 4.1 | Phonon spectrum of monolayer WSe_2. Calculated phonon spectrum of a monolayer WSe_2 on top of gold through several high-symmetry points of the Brillouin zone (Γ, M, and L). The phonon branches accessible with our existing LW-STM are marked by the orange shaded area ranging from a frequency of 0 THz to 2 THz.

few femtoseconds, such as intramolecular vibrations or electronic wavepacket motion. The limitations of the setup used for data acquisition in the previous chapters become especially clear in the pump-probe curves displayed in Fig. 3.16. There, all dynamics occurring within the first 500 fs of pump-probe delay time cannot be resolved due to the long pulse duration of the terahertz pulse. Additionally, as in state-selective LW-STM the electron can tunnel at any time the field-strength of the transient reaches a resonance in the sample, the temporal resolution is governed by the peak width of the terahertz transient. This has so far limited the tunnelling window to a few hundred femtoseconds. This timing jitter of the tunnelling process leads to an additional noise source and uncertainties in the signal.

All constituents of matter, not only the ones bound in periodic lattices, are constantly in motion. Molecules exhibit inter- and intramolecular vibrations that strongly alter their properties. Numerous research groups have investigated the importance of intramolecular vibrations on the functionality of molecules. It has been shown that for example in light harvesting, vibrations play a critical role for the efficiency of excitation energy transfer [Kun22], singlet fission [Dua20], and even photosynthesis [Ful14]. Moreover, the properties of small and ubiquitous molecules such as carbon dioxide [Joh95] or water [Per16] are strongly governed by their intramolecular vibrations. Investigating such vibrations in real-time and on atomic length scales would by highly appealing.

To overcome the deficiency of limited temporal resolution in the existing terahertz

LW-STM, one possible solution is centred around reducing the pulse duration, confining the tunnelling event to timescales below 20 fs or even 10 fs. Additionally, if one is interested in the time evolution of a system only a few femtoseconds after excitation, the pump and probe pulses have to be incident on the sample in immediate succession. This is only possible if the pre- and post-oscillations of the pulses are weak and/or decay quickly. For LW-STM, this would mean that a hypothetical light pulse shaped like a Dirac-delta distribution possesses the optimal temporal shape. However, a delta-spike-shaped electric field contradicts Maxwell's equations for a freely propagating electromagnetic wave, where the time integral over the electric field must vanish. Therefore, a sub-cycle pulse where the field of a strong positive half-cycle exceeds the feeble negative excursions by a factor of 1.5 to 2 is a favourable and still realistic pulse shape. The goal of obtaining a shorter tunnelling window can be achieved by increasing the frequency of the carrier wave from 0.9 THz into the multi-terahertz or mid-infrared regime [Hub00a, Dhi17, Lei23]. Due to the shorter oscillation period, the tunnelling window and the minimal time delay between two pulses before they begin to interfere can be reduced simultaneously. Thereby, measurements of the fast local dynamics only a few tens of femtoseconds after the excitation pulse can be investigated, as the post-oscillations of the mid-infrared pulse decay on this timescale.

Criteria for suitable mid-infrared pulses. In order to achieve an even faster temporal resolution in a next-generation LW-STM, the laser pulses must fulfil several requirements. Apart from the high stability and constant carrier-envelope phase (CEP), which has already been discussed in the introduction of LW-STM (section 2.3), the centre frequency must be considerably increased. This must be achieved while maintaining the high field strengths necessary to drive tunnelling currents within the Ångström-sized gap.

As no laser sources exist that directly emit suitable mid-infrared pulses, a wide range of techniques have been developed for the generation of CEP-stable multi-terahertz pulses. These include photon mixing in gas jets [Coo00], spintronic emitters [Sei16], difference-frequency generation in nonlinear crystals [Hub00a], and a novel source termed quantum well emitters [Mei22]. The reader is referred to section 4.3 for details. After careful consideration of tunability, efficiency, and bandwidth of the table-top solutions, terahertz generation through the nonlinear optical effect of difference-

frequency generation (DFG, see explanations following eq. 2.10) within the broad spectrum of a single pulse [Hub00a, Sel08, Jun10] proves the most fitting for our applications and the desired laser parameters. In this mixing process, the difference frequencies of spectral components present in the incident pulse are created.
In order to estimate the power of the pump laser necessary for the generation of the intense mid-infrared pulses via DFG, the following concepts were considered:
After generation, the multi-terahertz pulses are focused into the STM junction, acting as an ultrafast bias. Thus, the field strength of the pulses in the near field has to be large enough to result in bias voltages of several volts across the Ångström-scale tunnelling gap. As most interesting semiconducting crystals and molecules exhibit energy levels in the range of $-2\,\mathrm{V}$ to $2\,\mathrm{V}$ from the Fermi level, such transient voltages should be preferably reached with the laser pulses. Assuming a gap size of 10 Å, this amounts to extreme field strengths of $20\,\mathrm{MV/cm}$ across the tip-sample junction. To estimate the necessary pulse energy and far-field strength, a comparison between the pulses employed in the first generation terahertz LW-STM and the second generation mid-infrared LW-STM provides insights. First, the higher-frequency radiation ($25\,\mathrm{THz}$ versus $0.9\,\mathrm{THz}$) of the new source would allow for tighter focusing in the tunnelling junction. With the short focal length implemented in the new STM scan head ($f = 7.5\,\mathrm{mm}$, see sec. 4.6), a focus with an FWHM of $40\,\mu\mathrm{m}$ for a centre frequency of $25\,\mathrm{THz}$ can be easily achieved. Notably, this is a factor of 50 smaller than the focal spot in the first-generation terahertz-driven STM (approximately $2\,\mathrm{mm}$ FWHM). As a result, assuming that the employed pulse energy remains the same, this would lead to a 50^2 times larger intensity and 50 times enhanced electric field. Second, as the mid-infrared pulses are shorter by a factor of approximately 25, the peak intensity is also scaled by this value. The peak field is thus stronger by a factor of $\sqrt{25}$.
Moreover, when the radiation propagates along the tip towards the nanoscopic gap, its field strength is enhanced by several orders of magnitude (see sec. 2.5). This near-field enhancement factor is roughly inversely proportional to the frequency of radiation. The field enhancement of $2{\times}10^5$ for a centre frequency of $0.9\,\mathrm{THz}$ determined in section 2.5 is therefore expected to be reduced to approximately $3{-}5{\times}10^3$ for a centre frequency of $25\,\mathrm{THz}$. In total, compared to the terahertz-range, a mid-infrared pulse with the same pulse energy is expected to lead to an approximately

six times larger peak near field. Since the terahertz pulse energies of 400 pJ (at a centre frequency of 0.9 THz) in the first-generation STM (see sec. 2.2) suffice for the application in LW-STM, in the second-generation mid-infrared-driven STM, a mid-infrared pulse energy of $400\,\mathrm{pJ}/\sqrt{6} \approx 160\,\mathrm{pJ}$ should be adequate.

Prerequisites for the laser pulses. In the previous paragraph, I have established that a mid-infrared pulse energy of 160 pJ or preferably more must be available. Taking an additional buffer of a factor of ten into account, this leads to a desired mid-infrared pulse energy of around 1.6 nJ. As we want to generate the mid-infrared pulses by means of DFG, a suitable pump laser with sufficiently large near-infrared pulse energies must be employed. Reported conversion efficiencies of DFG using near-infrared pulses to generate multi-terahertz pulses lay in the range of a few percent [Jun10]. Sell *et al.* [Sel08] even reported on quantum conversion efficiencies of up to 18% in periodically poled lithium niobate as nonlinear crystal when pumped by near-infrared pulses with a centre wavelength of 1550 nm. Aiming towards a mid-infrared pulse energy of 1.6 nJ for a spectrum centred around 25 THz and cautiously assuming a quantum conversion efficiency of 1%, near-infrared pulses with an energy of around 2 µJ are needed. For this estimation, near-infrared pulses centred around a wavelength of 1028 nm (frequency of 292 THz) were assumed.

In order to maximise the conversion efficiencies in nonlinear optical processes, such as DFG, the peak intensity must be maximised. Thus, the pulse energy must be compressed in space and time. The former can be achieved by tight focusing of the beam, while for the latter, more elaborate compression techniques are necessary. Moreover, for the generation of multi-terahertz radiation based on DFG, the efficiency is also indirectly dependent on the pulse duration. Specifically, in order to generate a certain frequency with DFG, the pump spectrum has to be broad enough to support said difference frequency. For the generation of mid-infrared radiation centred at 25 THz exhibiting a pulse duration shorter than a single cycle of light, near-infrared laser pulses with a broad bandwidth spanning roughly 15−40 THz are necessary. For a laser source with a centre wavelength of 1028 nm and thus centre frequency of 292 THz, this corresponds to a quarter-octave spanning spectrum (wavelengths between 950 nm and 1130 nm) and a Fourier-transform-limited pulse duration of around 12 fs (for details on Fourier-transform limits see sec. 4.2). Achieving such ultrashort near-infrared pulses is therefore an essential first step for the realisation

of ultrashort mid-infrared pulses.
Such broadband, high-power, and high-repetition laser sources are not readily commercially available. Instead, we opted for the spectral broadening and temporal compression of near-infrared laser pulses within our laboratory. Therefore, when choosing a suitable pump laser amplifier, we must consider the non-unity efficiency of spectral broadening schemes. Hence, an additional factor of two in the pulse energy of the laser amplifier should be taken into account, leading to a desired minimal near-infrared pulse energy of 4 µJ.
Moreover, if the pulse duration is reduced, so is the duration of the transient bias applied across the STM junction. This decreases the tunnelling window and therefore the number of electrons tunnelling per pulse. In the interest of maintaining similar tunnelling currents (electrons tunnelled per second) and keeping the data acquisition times short, the repetition rate of the laser source has to be increased from 610 kHz in the first-generation THz-STM into the few megahertz (MHz) range. In combination with the minimum pulse energy of 4 µJ, this necessitates a near-infrared source emitting an average power of several Watts. In summary, the near-infrared pulses used for the generation of multi-terahertz pulses that are LW-STM-compatible have to fulfil a multitude of requirements including a broad spectrum, short pulse duration, and high pulse energies.

4.1 Optimisation of ultrashort near-infrared laser pulses

The myriad of requirements imposed on the laser source reduce the selection to fibre-based laser amplifiers. Only these are capable of generating high-energy pulses on the order of tens of microjoules while maintaining a high repetition rate of several megahertz. As a foundation of the optical setup for the new mid-infrared driven LW-STM (MIR-STM), a high-power diode-pumped Ytterbium-doped-fibre (Yb:fibre) amplifier is used (commercially available from Active Fiber Systems). It consists of a laser oscillator and a chirped pulse amplification (CPA) scheme [Nob18] with two laser amplifiers (see Fig. 4.2a). The oscillator output is amplified in a first Yb:doped, diode-pumped fibre (Fig. 4.2a, Amp1). The pulses are stretched by means of highly-

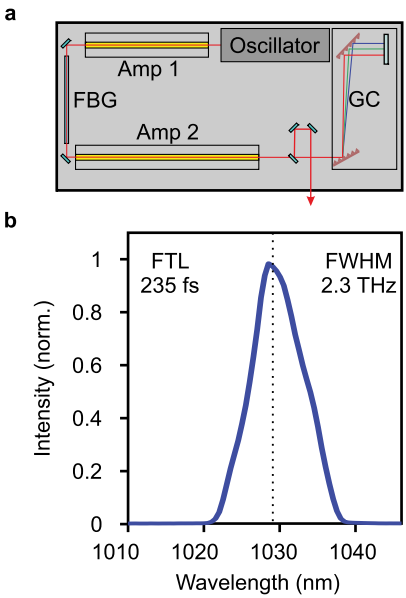


Figure 4.2 | Sketch of laser amplifier and its measured output spectrum. **a**, The output pulses of the oscillator are amplified in a first diode-pumped Yb:fibre amplifier (Amp1) before being temporally stretched in a fibre Bragg grating (FBG). The CPA scheme allows for the amplification of the laser power up to 100 W inside a second amplifier (Amp2). The pulses are recompressed afterwards using a grating compressor (GC). **b**, The intensity spectrum of the output pulses of the laser amplifier exhibits a centre frequency of 291 THz (centre wavelength 1029 nm, marked with dotted line). The full width at half maximum (FWHM) amounts to 2.3 THz and the Fourier-transform-limited (FTL) pulse duration is 235 fs (for details on Fourier-transform limits see sec. 4.2).

nonlinear fibre Bragg grating (Fig. 4.2a, FBG) prior to the main amplification fibre (Fig. 4.2a, Amp2) and recompressed afterwards using a grating compressor. The output pulses of the laser system exhibit an FWHM duration of 267 fs with an intensity spectrum centred at 1030 nm (Fig. 4.2b). The repetition rate is tunable between 50 kHz and the oscillator repetition rate of 18.57 MHz. The output power reaches 100 W at a capped pulse energy of 200 µJ. The pulse duration can be stretched up to several nanoseconds using the FBG.

Illustration of the optical setup. Before presenting the characterisation of the Yb:fibre amplifier, I will provide a brief overview of the different steps necessary to achieve the desired phase-stable single-cycle mid-infrared pulses as well as a source of tunable visible pulses for the application in the novel LW-STM. Figure 4.3 illustrates the main tools I utilised to arrive at this goal. Each step will be explained in detail in later sections within this chapter.

The output of the high-power Yb:fibre laser amplifier is guided through a beam-

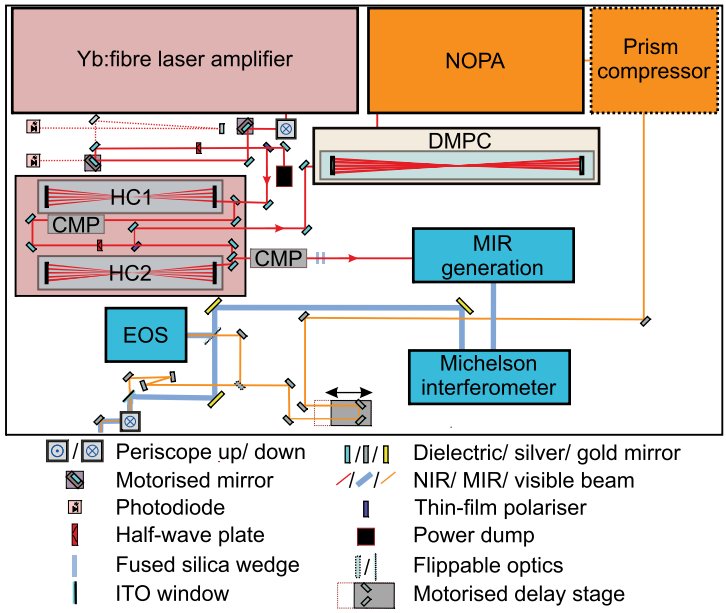


Figure 4.3 | Sketch of the optical setup. The near-infrared pulses emitted by a high-power Yb:fibre laser are guided through an active beam-stabilisation scheme prior to spectral broadening in a first Herriott-type multipass cell (HC1). After passing through HC1 and a chirped-mirror compressor (CMP), the compressed pulse is split into two parts, allowing for the separate generation of ultrashort mid-infrared and visible pulses. The largest fraction of the HC1 output is guided through HC2 and a second CMP to be then used for the generation of broadband phase-stable mid-infrared pulses via intrapulse DFG (see sec. 4.3). These mid-infrared pulses can be either directly focused into the tip-sample junction of the STM or pairs of pulses can be created in a Michelson interferometer (see sec. 4.4) before they are guided into the STM. A small fraction of the HC1 output-pulse energy is guided through an empty multipass cell ("delay multipass cell", DMPC, see sec. 4.5) to establish a beam path that is equally long as the one in HC2. In a non-collinear optical parametric amplifier (NOPA), tunable laser pulses in the visible domain are created (see sec. 4.5) and then compressed in a prism compressor to ensure ultrashort pulse durations. The visible and mid-infrared pulses are recombined using an ITO window before being focused onto the STM tip. Alternatively, the mid-infrared and visible pulses can be guided into an electro-optic detection scheme (EOS, see sec. 2.3). The prism compressor is not yet set up at the time of writing this thesis, which is indicated by the dotted outline of the corresponding box.

stabilisation scheme that utilises two motorised mirrors and four-quadrant photodiodes to actively stabilise the beam pointing. Subsequently, to spectrally broaden and temporally compress the near-infrared pulses, the near-infrared power is focused into the first of two Herriott-type pulse-compression cells (HC1). Their mode of operation will be explained in detail in section 4.2. By means of a broadband half-wave plate and a thin-film polariser (TFP) placed in between the two Herriott-cell stages, the pulse energy can be split into two parts such that two different goals can be pursued: the generation of phase-locked mid-infrared as well as tunable visible laser pulses.
The larger fraction of the pulse energy propagates through a second Herriott cell (HC2), broadening the spectra of the pulses even further. After compression in a chirped-mirror compressor and two fused silica wedges, the near-infrared pulses are focused into a gallium selenide crystal to generate pulses in the mid-infrared range by means of intrapulse DFG (see sec. 4.3). The resulting ultrashort, inherently phase-stable pulses can be split into two parts within a Michelson interferometer (see sec. 4.4), creating pairs of mid-infrared pulses. These pairs of pulses are then focused onto the STM tip to act as transient bias voltage. The pulse pairs, which can be delayed with respect to each other, will be used to perform MIR pump – MIR probe experiments. Alternatively, the mid-infrared pulses can also be guided into an electro-optic detection scheme (see sec. 2.3) for characterisation.
Instead of perturbing the system with a mid-infrared pulse, it is often desirable to investigate the response of the sample to excitations of the electronic degree of freedom. This is done by focusing pulses with sufficiently large photon energies onto the sample. To this end, laser pulses in the visible domain must be generated. The small part of near-infrared power split off after HC1 passes through an empty delay multipass cell (DMPC, see sec. 4.5) to generate a beam path which is similar in length to HC2. Subsequently, the near-infrared pulses are converted into the visible domain and amplified within a non-collinear optical parametric amplifier (NOPA, see sec. 4.5). This provides a tunable source for visible pulses that must be compressed temporally within a prism compressor prior to temporal and spatial overlap with the mid-infrared pulses using a glass slide covered with indium tin oxide (ITO). Subsequently, the visible and mid-infrared pulses are focused into the tip-sample junction.

The multitude of nonlinear optical processes mentioned in this paragraph and ex-

plained in more detail in later sections relies on a high-power source of ultrashort laser pulses. Only then can we achieve the desired laser parameters for an MIR-STM. In addition to the requirements of simultaneously high-energy and ultrashort near-infrared pulses, further prerequisites on the stability of the laser source and the pulse characteristics must be imposed. First, the optical beam path from the laser output to the incoupling of the generated mid-infrared radiation into the measurement device may span several to tens of metres. For this reason, an ultimately stable beam pointing and beam position are necessary. Second, to ensure a constant pulse energy in the generated mid-infrared range, the power of the laser amplifier must be highly stable as well. Third, for optimal focusing and interaction between different spectral components of the initial pulse within the nonlinear crystal, all frequencies should be homogeneously distributed over the beam area. The extent of existing inhomogeneity is quantified by the so-called spatial chirp (see sec. 4.1). Finally, the majority of the pulse energy must be confined temporally within the ultrashort pulse. However, the average power that is detected does not necessarily originate exclusively from the pulses. For example, a continuous wave background or even just an extended pedestal, which the short pulse rests upon, will distribute the laser power over a larger time window and thus drastically reduce the peak intensity (see sec. 4.1). This in turn will diminish the conversion efficiencies in nonlinear processes.
In order to fulfil all these criteria, the laser source must be characterised and optimised in the optical lab to ensure stable operation.

Beam-pointing stability. On the way from the laser output to the actual measurement apparatus, a long path of tens of metres with a myriad of delicately aligned optical components has to be traversed. As these components can be very sensitive to variations of the beam position, hence such fluctuations have to be minimised. Moreover, the focal size and position should not vary more than a few microns once the laser is guided into the tip-sample junction. Therefore, beam-pointing stability is of utmost importance for our application and must be closely monitored. To quantify the beam pointing, the output of the Yb:fibre amplifier was guided over several metres of beam path and then visualised using a beam profiler. Over the course of two hours, the beam position was registered each second. The pointing fluctuations of the laser source escalate with increasing amplifier pumping up to 33 µrad and 18.5 µrad root-mean-square (RMS) angular deviation in the x- and y-direction, respectively,

for 100 W at 5 MHz. Importantly, over the measured time, the beam continuously drifted in the horizontal direction. These unpredictable fluctuations of beam position necessitate an active beam-stabilisation scheme. Using two piezo-controlled mirror mounts and two four-quadrant diodes, the instantaneous beam centre-of-mass is monitored and actively controlled, reducing beam-pointing fluctuations to below 5 µrad RMS. Air flow, which also strongly influences beam propagation, is minimised by guiding the beam through tubes and housing the whole optical table.

Noise spectrum and power stability. In addition to ensuring positional stability, maintaining a stable laser output power is crucial during operation due to the sensitivity of nonlinear optical effects to even minor intensity fluctuations. Moreover, to minimise the measurement time, the signal-to-noise ratio needs to be maximised, approaching the shot-noise limit.

Monitoring the variation of the laser shot-to-shot signal on a photodiode over a time window of seven seconds reveals an RMS variation from the mean of 0.6%, which is acceptable for such a high-power laser amplifier. The amplifier was operated at a repetition rate of 5 MHz and output power of 100 W. Calculating the Fourier transform of the time trace provides the amplitude noise spectral density (see Fig. 4.4a), showing the overall drop of excess noise towards higher frequencies and several local maxima in the 100−1000 Hz range. For our measurement scheme of lock-in detection, it is essential to choose a modulation frequency where the noise is as small as possible. For realistic mechanical chopping frequencies – several hundreds to a thousand hertz –, the most desirable frequency windows are around 760 Hz or 1010 Hz (marked as dotted lines in Fig. 4.4a). By choosing a chopper with a larger number of spokes, even a chopping frequency of 2 kHz may be attainable.

Possible tuning knobs to change the noise spectrum lie in the temperature of the deionised water in the laser cooling circuit, the chiller compressor, and the laser pumping diodes as well as optimised electronic circuitry.

In spite of all attempts to minimise the noise, it is necessary to collect measurement data over several hours to maximise the signal-to-noise ratio. To this end, the output power of the laser must be suitably stable. After letting the amplifier system and the whole optical setup thermally equilibrate, the power fluctuations of around 0.1% root-mean-square over 150 minutes (see Fig. 4.4b) are remarkably low ensuring very stable measurement conditions over long periods of time.

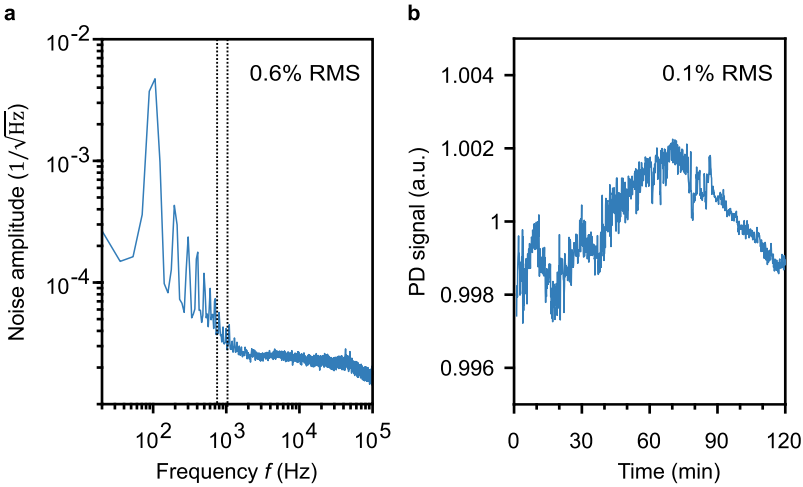


Figure 4.4 | Shot-to-shot and long-term stability. a, Monitoring the pulse train over 7 s reveals an RMS noise of 0.6%. Calculating the Fourier transform uncovers the dominating noise peak at around 100 Hz. Local minima in noise and thus good choices for modulation frequencies in lock-in detection are marked by dotted lines at 760 Hz and 1010 Hz. **b**, The long-term power stability monitored using a photodiode (PD) over 120 minutes possesses an RMS noise of 0.1%.

Beam quality. For optimal propagation and focusing conditions, it is useful to work with a beam profile that resembles a Gaussian beam as closely as possible. The M^2 parameter, also called beam propagation ratio or beam quality factor, quantifies how close the beam profile is to an ideal Gaussian profile. A value of $M^2 = 1$ represents a perfect Gaussian beam, while larger values may lead to increased beam divergence, larger foci, irregular beam profiles, or other imperfections.

The optimal laser source outputs light pulses whose spatio-temporal shape can be modelled by a Gaussian. For laser pulses propagating in the z-direction, the intensity distribution can be written as:

$$I(r,z,t) = I_0\left(\frac{w_0}{w(z)}\right)^2 \exp\left(\frac{-2r^2}{w(z)^2}\right) \exp\left(\frac{-2t^2}{\tau^2}\right) \tag{4.1}$$

with the intensity at the beam centre I_0, the beam waist radius w_0, and the momentary Gaussian beam radius $w(z)$ defined as the radial distance at which the intensity has dropped to $\frac{1}{e^2}$ of its value at the centre. Accordingly, the temporal envelope is quantified by τ, which indicates a drop to $\frac{1}{e^2} \times I_0$ in time.
At each position z along the beam propagation axis, the beam radius is given by:

$$w(z) = w_0 \sqrt{1 + \left(\frac{z}{z_\mathrm{R}}\right)} \tag{4.2}$$

with the Rayleigh length:

$$z_\mathrm{R} = \frac{\pi w_0^2 n}{\lambda M^2}, \tag{4.3}$$

which quantifies the propagation distance over which the beam area doubles from the beam-waist position.
The M^2 parameter of a given beam is determined by measuring the beam diameter at multiple positions within the Rayleigh length and fitting the beam divergence:

$$w^2(z) = w_0^2 + \mathrm{M}^4 \left(\frac{\lambda}{\pi w_0}\right)^2 (z - z_0)^2 \tag{4.4}$$

For the Yb:fibre amplifier used for this project, the average M^2 values are 1.2 for both the x- and y-direction. These values are close to the optimal case of $M^2 = 1$, indicating good focusing conditions.
Spatial homogeneity. The intensity spectrum of the laser output is not necessarily the same at different positions across the beam profile. This spatial inhomogeneity or spatial chirp leads to less efficient nonlinear optical processes (see sec. 4.2) and should thus be minimised.
Measuring the intensity spectrum indeed reveals strong inhomogeneities in the spectral composition across the beam diameter. Supplementary Figures C.1a and b illustrate the striking differences in the spectra measured at different positions within the beam area. A gradual change along the vertical axis of the beam cross section is revealed. More precisely, at the upper portion of the beam, a larger spectral weight of the long-wavelength side can be observed, while the lower portion features stronger short-wavelength components (see Suppl. Fig. C.1a). Similarly, albeit less pronounced, in the horizontal direction, the short-wavelength spectral weight

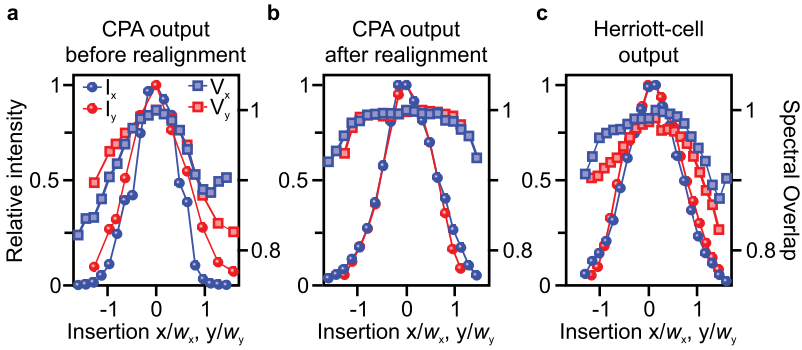


Figure 4.5 | Relative intensity and V-value. The spatial homogeneity of the laser beam is investigated at the CPA output before realignment (**a**), after realignment (**b**), and at the Herriott-cell output (**c**) by recording the intensity spectrum at various positions across the beam profile. The relative intensity (left axis), showing a Gaussian shape for all three panels, is obtained by integrating the spectra and normalising them to the spectrum on the beam axis ($x = y = 0$). A quantification of the homogeneity is given by the spectral overlap or V-value (right axis). By realignment of the compressor after the CPA, the spectral overlap in both radial directions V_x and V_y could evidently be improved from a weighted average of 97.7%/96.5% to 99.7%/99.5%. The beam at the output of the Herriott cells possesses slightly inferior spatial homogeneity than the realigned CPA output.

increases from left to right (Suppl. Fig. C.1b). This information can be analysed further in order to quantify the spectral changes. A regularly used figure of merit for spatial homogeneity is the spectral overlap or so-called V-value:

$$V = \frac{\left[\int \sqrt{I(\lambda) I_0(\lambda)}\, d\lambda\right]^2}{\int I(\lambda)\, d\lambda \cdot \int I_0(\lambda)\, d\lambda} \tag{4.5}$$

This V-value quantifies how much the intensity spectrum at a certain position $I(\lambda)$ differs from the spectrum on the beam axis $I_0(\lambda)$. In Fig. 4.5a, the relative intensity $I_{\mathrm{rel}} = \left(\frac{\int I(\lambda) d\lambda}{\int I_0(\lambda) d\lambda}\right)$ and V-values are depicted as a function of the transverse coordinate x/w_x or y/w_y, where w_x and (w_y) denote the beam waist in the x- and y-direction, respectively. By weighting the spectral overlap with the relative intensity, a single

scalar can be used to describe the spatial homogeneity of a laser beam in the x- and y-direction, separately:

$$\overline{V_x} = \frac{\int I_{\mathrm{rel}}(x) \cdot V_x(x) dx}{\int I_{\mathrm{rel}}(x) dx} \tag{4.6}$$

$$\overline{V_y} = \frac{\int I_{\mathrm{rel}}(y) \cdot V_y(y) dy}{\int I_{\mathrm{rel}}(y) dy} \tag{4.7}$$

The weighted V-values extracted for the CPA output, $\overline{V_x} = 97.7\%$ and $\overline{V_y} = 96.5\%$, are considerably worse than for comparable laser sources.
After trouble-shooting, the issue was identified to be a misalignment of the grating compressor after the CPA within the laser housing. Correction of the grating and end-mirror alignment greatly improved the spatial homogeneity (see Suppl. Fig. C.1c and d). The spatial variation has become barely visible in the y-direction and only minor changes in the x-direction are discernible. The V-values have correspondingly greatly improved to $\overline{V_x} = 99.7\%$ and $\overline{V_y} = 99.5\%$ (see Fig. 4.5b).
For completeness, the V-values of the output of the dual-stage Herriott-type pulse-compression cell (short HC) are depicted in Fig. 4.5c. The mode of operation of the HC is detailed in section 4.2, which the reader may kindly be referred to for further information. Measuring the spectra across the HC output beam profile reveals that the relative intensity curve possesses a Gaussian shape. The spectral overlap is slightly worse than at the CPA output with $\overline{V_x} = 98.1\%$ and $\overline{V_y} = 96.9\%$. Minor changes in the spectral width are expected from the slight inhomogeneities in nonlinear processes within the HC. However, shifts in the centroid position are also detected. This is most likely a remnant of the spatial inhomogeneities of the pump laser (see Suppl. Fig. C.1c and d).
Pulse contrast. In addition to optimising the laser characteristics in the spatial domain, modifications in the temporal domain offer another possibility to enhance the performance of the employed nonlinear effects. Notably, the efficiency of nonlinear processes is not determined by the pulse energy alone but rather the peak intensity of the pulse. Thus, the precise temporal shape of the pulses, known as pulse contrast, must first be closely investigated.
To this end, a custom-built non-collinear SHG autocorrelator was employed (see appendix C.1.2). In SHG-autocorrelation measurement schemes, the laser pulse is

split into two equally strong copies that are delayed with respect to each other prior to superposition in a nonlinear crystal. Detection of the generated second-harmonic and sum-frequency components as a function of the delay time provides information about the initial pulse's duration. Figure 4.6a depicts the autocorrelation trace collected within the relevant delay range from −75 ps to +75 ps for a laser repetition rate of 488 kHz on a semi-logarithmic scale. As expected, the emerging trace reveals a main peak at a delay time of 0 ps, corresponding to perfect temporal overlap of the two autocorrelator arms. However, interestingly, the main peak resides upon a relatively large pedestal that even features several smaller peaks, for instance at a delay time of ±12 ps (Fig. 4.6b). Performing a low-noise measurement for a smaller time window (Fig. 4.6c) uncovers even more remarkable characteristics of the pulse pedestal. At time delays of −1 ps and +1 ps, large peaks with intensities of almost 10% of the main peak appear. Especially these latter side peaks lead to smearing of the pulse energy over a larger time window than the pulse duration (Gaussian fit in Fig. 4.6d). As a result, the accessible peak intensity for nonlinear processes is lower than initially estimated. As an approximate measure of the accessible intensity, I quantified the power within the FWHM of the pulse. However, since in the measurement process of SHG autocorrelation, the information about the precise temporal shape of the pulse is lost, a simulation had to be set up. To this end, the incident near-infrared pulse shape was modelled by overlaying a short Gaussian intensity envelope, simulating the femtosecond pulse, with multiple broad Gaussians and squared hyperbolic secants to mimic the pulse pedestal. By numerically autocorrelating and fitting the simulation to the measured data, an estimate for the true intensity envelope was obtained. Integration of the resulting intensity traces reveals that 37% of the pulse energy is confined within the FWHM. Assuming the FWHM as an estimate for the accessible pulse energy for nonlinear processes, this simulation suggests that only 37% of the total pulse energy can be harnessed for the essential nonlinear optical effects employed in our setup. Thus, it is of great importance to investigate and ideally eliminate the causes of the broad pulse pedestal.
A potential factor may lie in the characteristics of the employed laser source. To test this hypothesis, I systematically varied numerous laser-amplifier parameters and indirectly observed the ensuing effect on the pulse contrast by assessing the performance of the down-stream nonlinear optical processes. Since the employed multipass

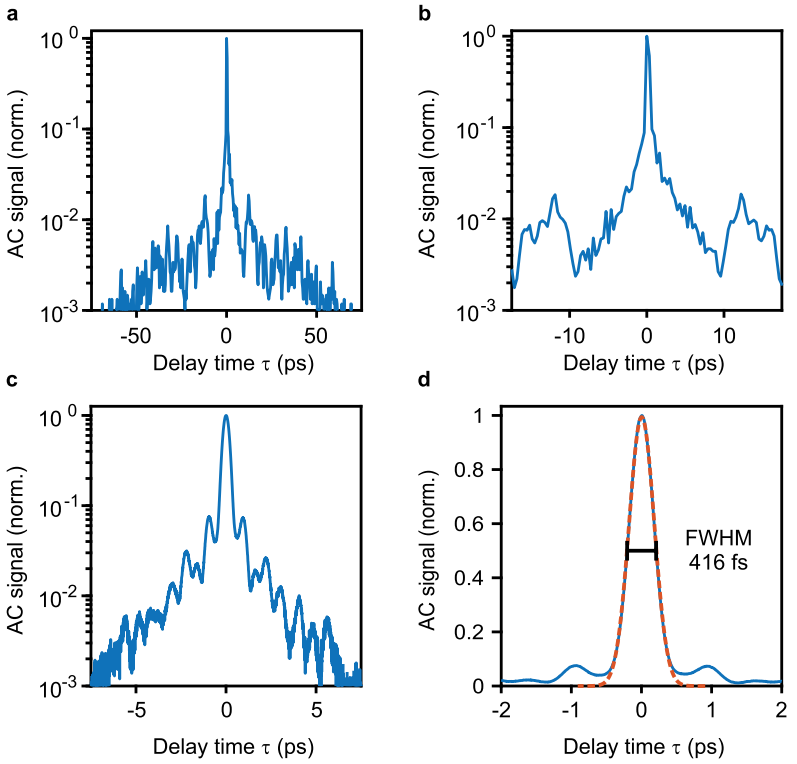


Figure 4.6 | Autocorrelation to characterise the pulse pedestal. a, The natural logarithm of the sum-frequency autocorrelation signal over a large delay-time window (±75 ps) reveals side peaks. **b**, In this enlarged section of the −17.5 ps to +17.5 ps range, the side peaks at ±12 ps reaching 2% of the main peak signal (at 0 ps) are even more discernible. **c**, A high-resolution low-noise autocorrelation trace for ±7.5 ps shows even more critical peaks at ±1 ps reaching 7% of the main peak signal. **d**, The autocorrelation signal for ±2 ps (blue curve) and a Gaussian fit to the data (red dashed curve) are depicted. The FWHM of the fit amounts to 416 fs, leading to an estimated pulse duration of 295 fs. These data were recorded with a repetition rate of 488 kHz and an average power of 100 W.

or Herriott cell (HC1) is the first nonlinear component in the optical path of the laser beam (see sec. 4.2), I observed the efficiency of this component. As HC1 was implemented to spectrally broaden the near-infrared pulse using self-phase modulation, the width of the spectrum and the Fourier-transform-limited pulse duration served as suitable measures of efficiency for this analysis. A more detailed explanation of Fourier-transform limits is provided in section 4.2.
An interesting outcome was observed upon variation of the diode pumping currents in the amplifier, i.e., the amplified output energies. Importantly, the total intensity guided through the Herriott cell for spectral broadening was held constant for different diode currents to ensure comparability. Specifically, a pulse energy of 40 µJ was split off by means of a half-wave plate and a polarising beamsplitter. As Fig. 4.7 illustrates, although the shapes of the output spectra closely resemble each other for three different pulse energies, the output spectrum of HC1 is widest for the lowest pulse energy of 41 µJ (red curve) and narrowest for the highest pulse energy of 200 µJ (blue curve), with an intermediate width laying in between (82 µJ, red curve). Remarkably, despite using the same pulse energy for broadening, different pumping currents applied to the laser amplifier diodes result in significant variations in spectral broadening. This phenomenon highlights the impact of the pumping current on the nonlinear processes occurring within the laser amplifier and identifies the pumping current as the key source of the above-mentioned wide pedestal in the pulse contrast.
Having established the cause for the unfavourable smeared shape of the pulse, the question arises how a larger pump current leads to a more narrow spectrum after the HC1 and vice versa. Most likely, an amplified spontaneous emission within the laser's gain medium plays a key role. To understand this, it is important to revisit the fundamental principles facilitating laser light: Normally, spontaneous emission produces the seed pulse that by positive feedback within the oscillator cavity leads to lasing once the lasing threshold is reached. The pulses leaving the oscillator, working at a repetition rate of 18.57 MHz, are pulse-picked down to the desired repetition rate (pulse picker of 39 in the case of 488 kHz) prior to injection into the dual-stage amplifier fibres. By changing the diode pumping current in the amplifier, the population inversion and thus amplification factor of the weak oscillator pulses can be varied. However, excess spontaneous emission is an unwanted effect as it consists of incoherent radiation which noticeably reduces the maximally possible amplifier

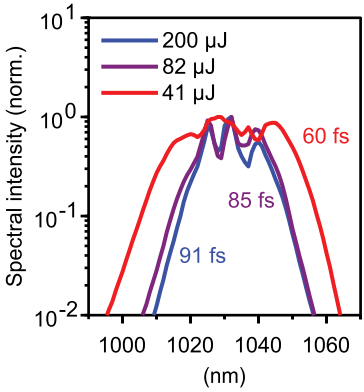


Figure 4.7 | Optimising the pulse contrast: influence of amplifier output power. The output-pulse energy of the laser amplifier was continuously tuned using the pumping current of the diodes. The three output spectra of the first HC were all recorded at 488 kHz and with the identical input-pulse energy of 40 µJ, which was controlled using a half-wave plate and a polarising beamsplitter. The spectral broadening increases with decreasing laser output energy, leading to Fourier-transform-limited pulse durations (intensity envelope) of 60 fs, 85 fs, and 91 fs for laser amplifier output energies of 41 µJ, 82 µJ, and 200 µJ, respectively.

gain. The contribution of amplified spontaneous emission to the overall signal self-evidently increases when population inversion increases. This boosts the incoherent background and depletes energy from the femtosecond pulse. Finally, this strongly impairs the efficiency of nonlinear processes, in the case of our setup the self-phase modulation within the Herriott cell, as depicted in Fig. 4.7.
In conclusion, for optimal spectral broadening within the Herriott cell, the laser amplifier should thus be pumped only just enough to generate the desired output-pulse energy. This minimises the loss of accessible pulse intensity and thereby maximises the efficiency of the essential nonlinear conversion mechanisms.

4.2 Pulse compression in Herriott-type multi-pass cells

In the last section, the origin of the unfavourable pulse pedestal was identified and the effect minimised. The optimised ultrashort near-infrared pulses, which shall be labelled “pump” pulses henceforth, can now be used for the generation of radiation in the terahertz spectral region. In general, in the multitude of generation techniques

(see sec. 4.3), the terahertz bandwidth critically depends on the pulse duration of the incident radiation. If we were to use the previously mentioned nonlinear conversion process of difference-frequency generation (DFG), as described in the explanations following equation 2.10, the highest frequency that can be generated is linked to the largest difference frequency in the near-infrared spectrum. Since we are interested in pulses in the multi-terahertz to mid-infrared spectral region, the incident spectrum must support these difference frequencies. The laser output pulses exhibit merely a bandwidth of 7.9 THz (wavelengths ranging from 1028 nm to 1032 nm). A terahertz pulse generated through OR using this spectrum would merely increase the temporal resolution of state-selective tunnelling in LW-STM by a factor of four. For the realisation of the intended ultrafast LW-STM that is able to visualise intramolecular vibrations or electronic dynamics occurring on the few-femtosecond scale, even shorter mid-infrared waveforms are necessary. Thus, a more broadband near-infrared pump spectrum is required.

Spectral phase. Before moving on to how this spectral broadening scheme is achieved in my setup, another prerequisite for achieving ultrashort pulses must be introduced.

The temporal evolution of the electric field $E(t)$ of a light pulse and its spectral composition $\tilde{E}(\omega)$ are linked by a Fourier transform:

$$\tilde{E}(\omega) = \mathfrak{F}[E(t)] = E_0(\omega) \cdot e^{-i\Phi(\omega)} \tag{4.8}$$

wherein $\Phi(\omega)$ denotes the spectral phase. It is thus immediately clear that optimising only the spectral intensity given by $I(\omega) = |\tilde{E}(\omega)|^2$ does not suffice. The effects of the amplitude $E_0(\omega)$ as well as the phase $\Phi(\omega)$ must be considered. The shortest possible pulse duration achievable for a given spectrum is obtained when $\Phi(\omega)$ is constant over all relevant frequencies ω. Only then do all spectral components optimally constructively interfere at $t = 0$, generating the so-called Fourier-transform-limited (FTL) pulse.

The phase $\Phi(\omega)$ can be expanded in a Taylor series:

$$\Phi(\omega) = \mathrm{D}_0 + \mathrm{D}_1(\omega - \omega_0) + \frac{1}{2} \cdot \mathrm{D}_2(\omega - \omega_0)^2 + \frac{1}{6} \cdot \mathrm{D}_3(\omega - \omega_0)^3 + ... \tag{4.9}$$

around the centre frequency ω_0. The term D_0 corresponds to the carrier-envelope phase (CEP) and describes the phase difference between the carrier wave and its intensity envelope. The second term proportional to the group delay D_1 relates to a linear phase shift and therefore shifts the full pulse in time. For the pulse duration, only higher orders in $(\omega - \omega_0)^n$ play a role. The group-delay dispersion (GDD) D_2 and third-order dispersion (TOD) D_3 stretch the pulse and in the case of TOD can lead to pre- and post-oscillations. As a result, the higher-order dispersion must be closely monitored and adjusted throughout the beam path to ensure short pulses. Only then is the most efficient interaction within the myriad of nonlinear effects utilised in the setup warranted.

Self-phase modulation. In order to obtain the first necessary ingredient for ultra-short pulses – the broad spectrum –, the near-infrared pump pulses are spectrally broadened within two Herriott-type multi-pass cells (short HC).

The broadening scheme within the Herriott cells relies on a third-order nonlinear optical effect within Argon gas. In contrast to second-order effects like DFG and SHG (see eq. 2.10), third-order phenomena impose no restrictions on the symmetry of the medium. The polarisation within an inversion-symmetric medium can be expanded as a power series of odd-order terms:

$$\mathbf{P} = \varepsilon_0 \Big(\chi^{(1)} \mathbf{E} + \chi^{(3)} \mathbf{E}^3 + ... \Big) \tag{4.10}$$

with the dielectric constant ε_0 and the material-specific linear as well as third-order susceptibilities $\chi^{(1)}$ and $\chi^{(3)}$, respectively.

The term proportional to $\mathbf{E}^3$ leads to a multitude of interesting effects such as self-phase modulation and self-focusing. Assuming an oscillating electric field given by $\mathbf{E} = \mathbf{E}_\omega \cos(\omega t)$ and neglecting terms beyond the third order, equation 4.10 simplifies to:

$$\mathbf{P} = \varepsilon_0 \Big(\chi^{(1)} \mathbf{E} + \chi^{(3)} \mathbf{E}^3 \Big) = \varepsilon_0 \Big(\chi^{(1)} + \frac{3}{4} \chi^{(3)} |\mathbf{E}_\omega|^2 \Big) \mathbf{E}_\omega \cos(\omega t) \tag{4.11}$$

Here, the trigonometric identity:

$$\cos^3(x) = \frac{1}{4} \cos(3x) + \frac{3}{4} \cos(x)$$

was used and only terms oscillating at ω were included, as only these are relevant for the two nonlinear effects to be explained here.
The final equation (eq. 4.11) strongly resembles the linear polarisation solely with an additional nonlinear term in the susceptibility:

$$\chi = \chi^{(1)} + \chi_{\mathrm{NL}} \coloneqq \chi^{(1)} + \frac{3}{4}\chi^{(3)}|\mathbf{E}_\omega|^2 \tag{4.12}$$

As the susceptibility is linked to the refractive index, the following equation is obtained after further simplifications (see ref. [Boy08]):

$$n = n_0 + \frac{2\chi^{(3)}}{8n_0}E^2 = n_0 + n_2 I \tag{4.13}$$

Thus, the refractive index becomes dependent on the intensity of the light, which leads to intriguing nonlinear optical phenomena.
Now, a linearly polarised electromagnetic wave propagating along the z-axis shall be considered:

$$E(z,t) = A(z,t) \cdot e^{ikz - i\omega_0 t} \tag{4.14}$$

with a slowly varying envelope $A(z,t)$, wavevector k, and frequency ω_0. The intensity-dependent refractive index leads to an intensity-dependent wavevector:

$$k = \frac{\omega_0 \cdot n(I)}{c} = \frac{\omega_0}{c}\Big(n_0 + n_2(I)\Big) \tag{4.15}$$

This results in an intensity-dependent nonlinear phase ϕ_{NL}:

$$\phi_{\mathrm{NL}}(t) = \frac{\omega_0}{c} n_2 I \cdot z \tag{4.16}$$

The instantaneous frequency ω_{inst} is related to the derivative of the phase:

$$\omega_{\mathrm{inst}} = \frac{\mathrm{d}\phi}{\mathrm{d}t} \tag{4.17}$$

Due to the nonlinear refractive index, this leads to a shift of the instantaneous frequency:

$$\delta\omega_{\mathrm{inst}} = \frac{\partial \phi_{NL}}{\partial t} \tag{4.18}$$

Importantly, ω_{inst} does not necessarily have to be a frequency that was previously present in the incident spectrum. As a result, new frequency components can be generated, effectively broadening the initial frequency spectrum. This third-order nonlinear process is labelled self-phase modulation (SPM). Figure 4.8a displays the SPM created by a Gaussian-shaped intensity envelope with the temporal width τ (see eq. 4.1) and a peak height of I_0. The nonlinear phase $\phi_{\text{NL}}(t)$ follows the same temporal profile as the intensity envelope (see Fig. 4.8a, centre). The derivative of the nonlinear phase yields the change in instantaneous frequency due to the nonlinear refractive index (Fig. 4.8a, right). It leads to a maximum increase and reduction of the instantaneous frequency $\omega_{\text{inst}} - \omega_0$ are given by ω^+ and ω^-, respectively. To maximise the spectral broadening, two main factors are important. First, the effect can be optimised by using a medium with a high nonlinear refractive index n_2 (see eq. 4.13). Second, a larger SPM can be achieved by maximising the distance travelled by the pulse within the medium. However, the latter strategy is not feasible due to a second $\chi^{(3)}$-phenomenon happening simultaneously to SPM.

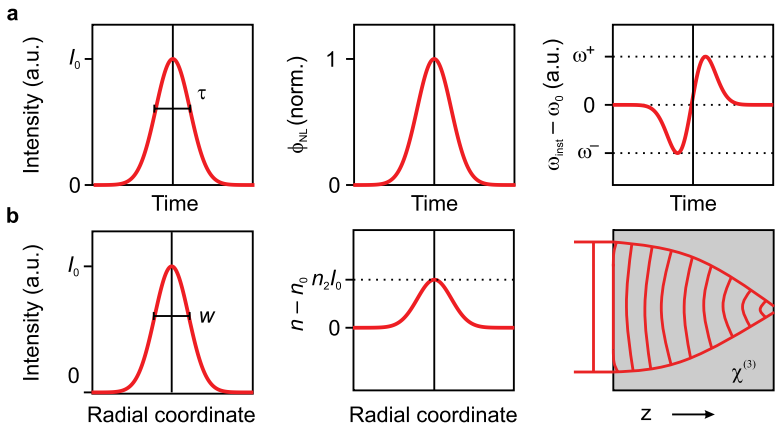


Figure 4.8 | Self-phase modulation and self-focusing. Due to the nonlinear refractive index, two intriguing phenomena arise. **a**, A Gaussian-shaped temporal intensity envelope with width τ and maximum I_0 (left panel) leads to a nonlinear phase due to self-phase modulation that follows the same shape (middle panel). The resulting nonlinear change to the instantaneous frequency $\omega_{\text{inst}} - \omega_0$ is given by the temporal derivative (right panel). The maximum increase and reduction of the instantaneous frequency are indicated by ω^+ and ω^-, respectively. The vertical black line in all three panels signifies the time at which the intensity reaches its maximum I_0. **b**, In case of a Gaussian intensity profile in the cross-sectional direction of the pulse with a width w and maximum I_0, the nonlinear change of the refractive index $n - n_0 = n_2 I$ becomes dependent on the radial coordinate (middle panel). Thus, upon entering a medium with a non-zero third-order susceptibility $\chi^{(3)}$, the wavefronts are bent, leading to reduction of the beam size as the pulse propagates in the z-direction (right panel). The vertical black line in the left and middle panels signifies the position at which the maximum intensity $I = I_0$ is reached.

Self-focusing. In addition to a temporal change of the intensity within the pulse, the intensity typically also varies spatially across the cross section of the beam. In the case of our setup, the laser pulses emitted by the Yb:fibre amplifier exhibit a nearly Gaussian spatial intensity profile (see eq. 4.1) with a beam waist radius w (see Fig. 4.8b, left) and a peak intensity I_0. This results in a radially decreasing refractive index experienced by the beam (Fig. 4.8b, centre) $n = n_0 + n_2 I$ and thus a lower propagation velocity v in the centre of the beam profile given by the simple relation:

$$v(r) = \frac{c}{n_0 + n_2 I(r)}. \tag{4.19}$$

This self-generated lens bends the wavefronts inwards, causing the beam diameter to shrink, thereby focusing the beam (see Fig. 4.8b, right). Due to this mechanism known as self-focusing (SF), the intensity of the laser beam increases. As this in turn will affect the spatial intensity profile, this is a self-amplifying effect. However, the naturally occurring effect of diffraction seeks to increase the beam diameter. Up to a certain threshold power, labelled the critical power, diffraction is able to effectively counteract the ever further reduction of the beam diameter. The critical power P_c is calculated from the known linear and nonlinear refractive indices, the vacuum wavelength of the light λ_0, as well as an analytical parameter that is dependent on the initial spatial intensity distribution of the beam. For a Gaussian mode profile, the constant possesses a value of approximately $\alpha \approx 1.9$ [Fib00], leading to the equation:

$$P_\mathrm{c} = \alpha \frac{\lambda_0^2}{4\pi n_0 n_2} \approx \frac{0.37\pi\lambda_0^2}{8 n_0 n_2} \tag{4.20}$$

For powers above P_c, so-called catastrophic self-focusing occurs, leading to ever-shrinking beam diameters until a plasma is created within the material. In some applications, this is even desired (e.g., supercontinuum generation). If the power is increased further, ultimately the nonlinear optical medium is damaged irreparably. Thus, when pulses with a power above P_c are employed and plasma generation is not desired (as is the case in controlled spectral broadening), the upper limit on the material length is given by the propagation length until the beam diameter has shrunk so much that it creates a plasma.

Herriott cell design. In the previous paragraph, it has been established that for optimal spectral broadening, the peak intensity should be preferably high and the propagation length in the nonlinear medium should be sufficiently long. However, as SPM and SF inadvertently occur simultaneously in any nonlinear medium, spectral broadening (facilitated by the former) cannot be harnessed without some extent of concomitant beam focusing (caused by the latter). Given that a high peak intensity of the incident pulse further increases the amount of self-focusing and thus leads to the formation of a plasma filament in the medium, employing a thick medium for spectral broadening is only possible for relatively low peak powers on the order of 1 MW. Even for near-infrared pulses longer than 1 ps, this limits the suitable pulse energies to below 10 µJ. This limitation can be overcome by using gas-filled capillaries as the medium, since due to low gas pressures smaller refractive indices than in solid state media are achieved. This increases the critical power P_c (see eq. 4.20), allowing for larger input-pulse energies. However, noticeable SPM comparable to solid state media only sets in for pulse energies above 100 µJ. This leaves a window between pulse energies of approximately 10 µJ and 100 µJ where these conventional techniques are not applicable for the spectral broadening and temporal compression of femtosecond near-infrared pulses.
In order to achieve large spectral broadening, while avoiding material and beam-profile deterioration, a new design was presented. In this scheme coined by Herriott *et al.* [Her64, Her65], the near-infrared beam repeatedly passes through either a thin bulk nonlinear medium or a pressurised compartment in a resonator-like setup. In each cycle, the beam acquires a nonlinear phase, leading to spectral broadening via SPM. In the case of a solid-state Herriott-type cell (HC), the SF is kept small as the propagation length within the medium is short. Additionally, the cell optics (focusing mirrors) in solid-state HC and gas-filled HC are designed such that they are able to compensate for the SF and create roughly equal-sized foci for subsequent passes [Sch16, Wei17b, Wei17a] through the resonator.
As mentioned above, the incident beam diameter and the focusing optics have to be chosen such that the peak intensity within the medium suffices to achieve considerable SPM while remaining below the threshold of plasma generation. In addition, different design parameters have to be adjusted depending on the specifications of the incident laser beam, the used nonlinear medium, and the desired output spectrum.

For example, the number of passes through the medium needs to be adapted, as this parameter determines the total nonlinear phase amassed by the near-infrared pulse. In principle, maximising the number of passes would be favourable, as this increases the exploited nonlinear effect of self-phase modulation. However, the maximal number of passes a resonator can maintain is limited due to geometrical reasons, which will be detailed in the following. As the beam must be coupled into and out of the resonator, all reflections apart from the first and last one must occur at different spots on the cell end mirrors. In a closed, off-axis ray path, in- and outcoupling can be achieved by placing a small pick-off mirror in front of one of the focusing mirrors or by directly employing a focusing mirror with a slit cut-out. In order not to clip parts of the beam, this pick-off mirror/cut-out must be larger than the beam diameter. After the pulse is coupled into the Herriott cell, two conditions need to be met throughout the numerous internal reflections: First, all reflection spots should be located at the same radial distance from the cell mirrors' centres. Second, neighbouring reflections on the resonator mirrors must be separated sufficiently far that they do not overlap. This provides a constraint on the beam diameter and number of passes through the resonator. After the final internal reflection, the beam is coupled out of the cell by passage through the original entry point. Thus, a closed optical path must be ensured. These prerequisites can be fulfilled by choosing a suitable geometry of the Herriott cell. The exact reflection pattern on the end mirrors of a Herriott cell is governed by two parameters: the number of full round trips through the cell n and the number of 360°-rotations the ray performs around the mirror axes m. Given the two quantities n and m, the angle θ_{nm} that subsequent reflections on one cell mirror are separated by must follow the relation:

$$n \cdot \theta_{nm} = 2\pi \cdot m, \text{ where } m = 1, 2, ..., n \tag{4.21}$$

Per round trip, the beam passes the nonlinear medium twice, leading to spectral broadening each time. For a closed path, the distance d_{nm} between the cell mirrors with the focal length f has to be chosen accordingly:

$$d_{nm} = 2f \cdot \left(1 - \cos\left(\frac{\theta_{nm}}{2}\right)\right) \tag{4.22}$$

A detailed mathematical derivation is provided in work by Herriott *et al.* [Her64]. As mentioned above, to maximise the nonlinear phase acquired by the pulse during propagation through the Herriott cell, it is generally desirable to choose a large number of full round trips through the medium, i.e., a large n. However, the quantity n is limited by several factors. First, due to the geometrical aspects mentioned above, only a limited number of reflections find room on the resonator end mirrors. Second, during the propagation through the medium and reflection off mirrors, the beam not only undergoes spectral broadening but also experiences dispersion. This temporally stretches the pulse leading to a reduced peak intensity, which in turn diminishes the broadening efficiency. Analogous to the above-mentioned restrictions on the number of round trips n and at a given focal length f, various restrictions also apply to the number of 360°-rotations m. Namely, n and m must be selected to be coprime numbers. This criterion results from the necessity that the beam does not reflect off the same position on the mirror more than once, which is necessary for the in- and outcoupling of the pulses. Furthermore, in order to achieve a self-imaging resonator design, where the focal spot at the centre does not change in size for subsequent passes through the cavity, for a fixed value m the distance between the resonator mirrors must be adjusted according to equation 4.22. Importantly, each mirror separation d_{nm} is linked to a different beam waist at the focus. Therefore, the intensity within the nonlinear medium can be varied by simply changing the resonator length d_{nm} and the value of m. This in turn affects the SPM, making the total nonlinear phase shift accumulated in the cell, given by the so-called B-integral, easily tunable:

$$B = \frac{2\pi}{\lambda_0} \int n_2(z) \cdot I(z) dz = \frac{2\pi}{\lambda_0} \int n_2(z) \cdot \frac{\varepsilon}{\tau \pi w^2(z)} dz \tag{4.23}$$

In this equation, λ, ε, and τ denote the central wavelength, energy, and duration of the near-infrared pulses in the medium with the nonlinear refractive index n_2. The variable z indicates the position along the direction of beam propagation.

Dual-stage Herriott cell. Due to all the constraints mentioned above, the pulse compression typically achievable in a single-stage Herriott cell is limited to a factor of approximately nine [Sei22]. Since we wish to reduce the output pulse duration and thus increase the spectral width considerably more, a dual-stage Herriott cell is necessary. By implementing two Herriott cells in series, a reduction of the near-

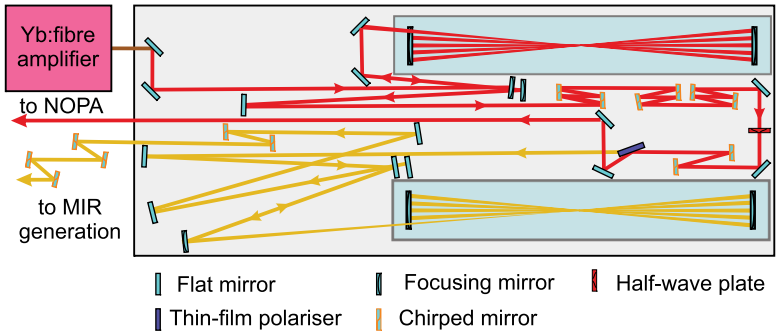


Figure 4.9 | Schematic of multipass cells in the setup. The laser pulses from the amplifier are focused into a first Argon-filled Herriott cell under a pressure of 16 bar (light grey box). After spectral broadening by SPM in 32 passes through the resonator ($n = 16$), the pulses are temporally compressed in a chirped-mirror compressor. By passing through a half-wave plate (red) and a thin-film polariser (dark blue) acting as a beamsplitter, a small fraction of the power is directed into the NOPA (see sec. 4.5). The largest portion of the pulse energy is focused into a second HC with a pressure of 10.5 bar (dark grey box, $n = 16$). Subsequently, using a second chirped-mirror compressor, the pulses are prepared for mid-infrared generation (see sec. 4.3).

infrared laser pulses from a duration of around 270 fs to below 80 fs in a first and to 13.5 fs in a second stage is achieved.

Figure 4.9 displays the final design of the dual-stage compression cell. The incident laser pulse from the amplifier is reflected off a focusing mirror ($f = 750$ mm) and focused into the first (top) pressurised resonator filled with 16 bar of Argon through a fused silica window and a 15° sector cut out in one of the resonator end mirrors. Within the resonator, the pulse propagates back and forth 16 times ($n = 16$), while travelling in azimuthal circles 13 times ($m = 13$). A camera image of the end mirror reveals the 16 round trips as spots on the end mirror (opposite to the cut out) of the first cell and the angular advance of 146° between subsequent reflections (see Fig. 4.10a). After these 16 round trips, the pulse leaves the resonator through the same cut-out but under a slightly upward tilted angle, such that it can be separated from the incident beam using a D-cut mirror. After pulse compression using suitable

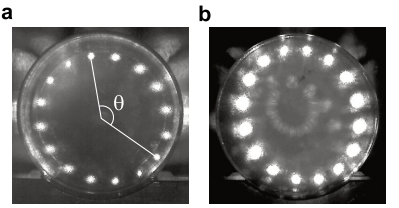


Figure 4.10 | Camera view of the multipass cell end mirrors. A camera is directed towards the unmodified end mirrors in HC1 and HC2 to monitor the cell performance during operation. **a**, A camera snapshot of the spot pattern on the end mirror of HC1 is shown for $n = 16, m = 13$. The so-called angular advance between two subsequent reflections $\theta = 146°$ is sketched for two exemplary spots. **b**, A snapshot of the spot pattern on the end mirror of HC2 reveals the 16 reflections the beam experiences within the resonator.

chirped mirrors, the pulse energy is split into two parts using a waveplate and a thin-film polariser. A small fraction of the pulse energy is used for the operation of a NOPA (sec. 4.5). The majority of the pulse energy is guided through a second HC (orange beam path) that is operated at a pressure of 10.5 mbar of Argon. Again, the pulse travels between the resonator mirrors 16 times ($n = 16$) before being compressed in a second chirped-mirror compressor and being guided towards the mid-infrared generation setup (see sec. 4.3). The individual steps are described in detail in the following.

The pulses coupled out after HC1 are characterised in detail to determine their suitability for our application. Figure 4.11a depicts the intensity spectrum of the HC1 output pulses as a function of the input-pulse energy. The FTL pulse duration drops from 260 fs at low input energy to around 90 fs for 29µJ (see Fig. 4.11b). However, the pulses emerging from HC1 are far from the FTL duration due to the dispersion of the medium and the self-phase modulation. The duration of these pulses can be accurately measured using autocorrelators or frequency-resolved gating (FROG), which are described in section 4.2.

By guiding the pulses through a chirped-mirror compressor with 14 reflections, a GDD of $-4900\,\mathrm{fs}^2$ ($-350\,\mathrm{fs}^2$ per reflection) can be applied, bringing the pulse duration close to FTL.

The overall transmission of HC1 including the chirped-mirror compressor amounts to 85%, resulting in an output energy of 24.6 µJ. Notably, the reduction of pulse

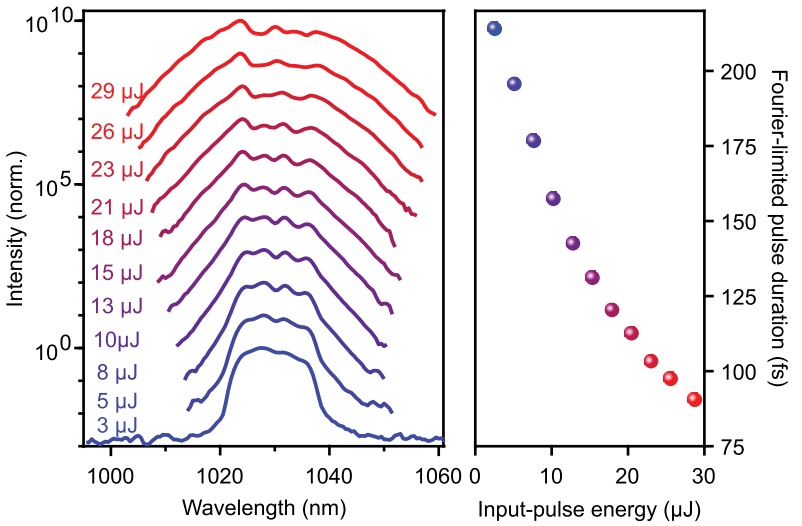


Figure 4.11 | Energy-dependent spectral broadening of the first Herriott cell. a, The output spectra of HC1 for increasing input-pulse energies are plotted using a semi-logarithmic scale. Curves are offset by one order of magnitude and values below 10^{-3} of the maximum are omitted for clarity. **b**, As a result of the spectral broadening, the FTL pulse duration drastically decreases with rising input-pulse energy, dropping below 90 fs for an input-pulse energy of 29 µJ.

duration from an FTL of 235 fs to 90 fs is not the best compression achievable with this setup. In fact, by optimising the chirp of the incident pulse, the gas pressure, and the pulse energy, the spectral broadening can be considerably amplified, reducing the output pulse duration to 75 fs. However, this was only attainable for low repetition rates up to 500 kHz. At higher repetition rates (1−2 MHz), I found that instabilities in spectral broadening begin to arise at input energies of 20 µJ. Further increasing the pulse energy, these instabilities drastically deteriorate the beam profile, rendering the pulses unsuitable for further operation. A possible reason was found in the instabilities of the gas. At high input energies, the peak intensity in the focal region at the centre of the cell is large enough to ionise the Argon gas and create a plasma. At low repetition rates, the gas has sufficient time to relax or be replaced between

subsequent laser pulses. As the repetition rate increases ever more, atoms are still ionised when a subsequent pulse arrives, leading to a build-up of plasma. The plasma leads to an alteration of the refractive index and the focusing conditions as well as to a deterioration of the beam profile. Further analyses and improvements must be made, if the desire persists to push the compression of HC1 to 75 fs or lower. After a second compression stage, the difference in output pulse duration of 11 fs compared to 13.5 fs between the optimised (low repetition rate) and stabilised (high repetition rate) case becomes irrelevant for our application.

Between the two HC stages, using a broadband half-wave plate and a thin-film polarising beamsplitter (TFP), the power can be variably separated into two parts. The power reflected off the TFP is used to operate a non-collinear optical parametric amplifier, which is closely described in section 4.5. The majority of pulse energy is transmitted through the TFP and focused into a second gas-filled Herriott-type multipass cell. As in the first multipass cell, the beam propagates through the resonator 32 times. However, in contrast to HC1, chirped end mirrors are implemented in the HC2 resonator. At each reflection, the chirped end mirrors imprint a GDD of $-50\,\mathrm{fs}^2$ onto the pulse. This is necessary as the spectra entering HC2 are already very broad. Thus, the dispersion in the medium stretches the pulses, which must be compensated for after each pass through the cell. Without the negative GDD of the mirrors, broadening in HC2 would only take place during the first few passes through the resonator and render the remaining passes futile. Notably, a delicate balance between spectral bandwidth and pulse duration must be maintained. If the chirped mirrors compensated for both the nonlinear phase and the dispersion of the gas-filled cell, an effect known as self-steepening [DeM67] could occur that uncontrollably compresses the pulse and damages the optics as a result. In our setup, the Argon gas at a pressure of 10.5 bar in HC2 leads to a material-caused GDD of $+53\,\mathrm{fs}^2$ per pass, which is almost compensated for by the chirped mirrors ($-50\,\mathrm{fs}^2$). The compensation of the material dispersion and not the nonlinear phase ensures a gradual broadening of the spectrum over all 32 passes (see camera image of the end mirror in Fig. 4.10b). After spectral broadening in the dual-stage HC, the spectrum exhibits spectral components from to 910 nm to 1100 nm at an input-pulse energy of around 29 µJ. The gradual broadening of the spectra after passing through both multipass cells is visualised in Fig. 4.12a. The FTL pulse duration drops from initially 155 fs

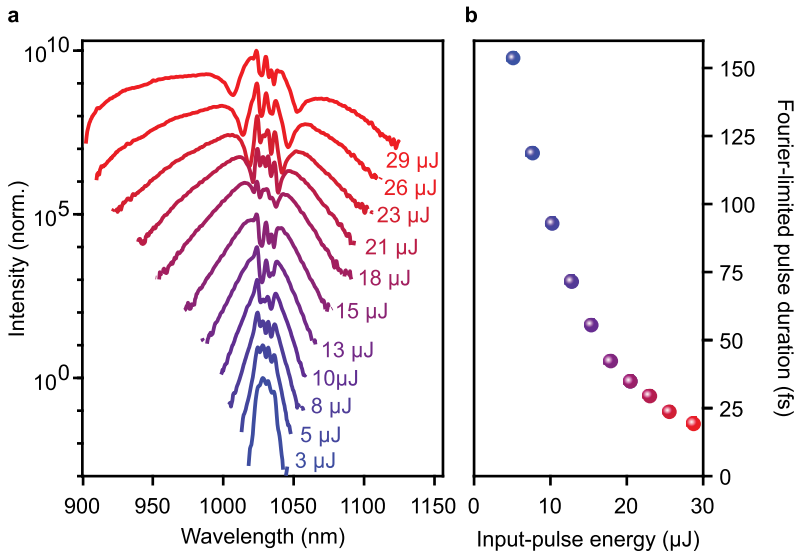


Figure 4.12 | Spectra and Fourier-limited pulse duration of the dual-stage Herriott cell. a, The output spectra of the dual-stage HC for increasing input-pulse energies are plotted on a semi-logarithmic scale. The curves are offset by one order of magnitude and measurement values below 10^{-3} of the peak are omitted for clarity. **b**, As a result of the spectral broadening, the FTL pulse duration drastically decreases with a rising input-pulse energy, dropping below 20 fs for an input-pulse energy of 29 µJ.

to below 20 fs (see Fig. 4.12b).

In addition to achieving broadband spectra, which has been detailed above, the generation of femtosecond near-infrared laser pulses necessitates a second requirement: All frequency components must interfere constructively. This is equivalent to demanding the relative phase of all spectral components to be brought as close to zero as possible. As a first step towards this goal, the pulse duration is determined in order to ensure that the pulses are as close to the FTL as possible.

Pulse diagnostics and compression. In order to temporally compress pulses to their FTL, the group-delay dispersion has to be first determined and then compen-

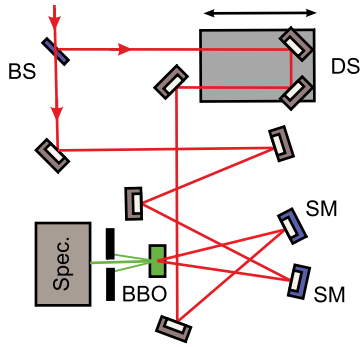


Figure 4.13 | Frequency-resolved optical gating. An incident transient is split into two identical copies using either a 50/50 beamsplitter or a D-cut mirror (BS). The beam path in one part can be varied using a motorised delay stage (DS). Subsequently, both beams are focused into the same β-barium borate (BBO) crystal under a small acute angle using two spherical mirrors (SM). When both beampaths temporally and spatially overlap, on top of the second harmonic of each arm separately, a substantial sum-frequency signal will arise. The spectral composition of the sum-frequency signal is recorded for each delay time using a spectrometer (Spec.), leading to a so-called spectrogram.

sated using suitable optical components.

As all electronic measurement techniques operate far too slowly to resolve the oscillation frequency of near-infrared light, one must employ optical nonlinearities in creative ways to measure pulse durations and spectral phases.

The most common method to determine the phase of femtosecond pulses is frequency-resolved optical gating (FROG). A customary variant called SHG-FROG is utilised here to assess the near-infrared pulses after the Herriott cells.

Figure 4.13 depicts the general mode of operation of an SHG-FROG. The pulse to be analysed is separated into two equal parts, which are spatially as well as temporally superimposed in a χ^2 nonlinear optical crystal (here β-barium borate, short BBO) under a small acute angle. The generated sum frequency is detected in a spectrometer. Varying the delay time between the two copies of the pulse and recording the spectra results in a FROG trace or spectrogram. Using a retrieval algorithm, the spectral intensity and phase information of the incident pulse can be reconstructed. Figure 4.14a shows such a measured FROG spectrogram of the broadened near-infrared pulses after passing HC1 and a chirped-mirror compressor. Clearly, the sum frequency strongly depends on the relative delay time and is strongest at a delay time of 0 fs. Utilising an iterative retrieval algorithm, the measured spectrogram is

reconstructed. The reconstructed spectrogram and spectral information are depicted in Fig. 4.14b and Fig. 4.14c, respectively. The spectral phase is relatively flat and shows only a minor curvature. To achieve this flat spectral phase, the pulses exiting HC1 are reflected 14 times off chirped mirrors with a GDD of $-350\,\mathrm{fs}^2$ per reflection prior to the FROG measurement. As the comparison of the calculated intensity envelope with the transform-limited pulse duration reveals, a non-zero GDD remains (see Fig. 4.14d). This is, however, of no issue as the combination of pulse energy and pulse duration leads to sufficiently high peak intensities for spectral broadening in a second HC.

After HC2, the compressed pulses are characterised again to ensure that they are as close to FTL as possible. Implementing a pair of fused silica wedges with variable thickness, facilitating the fine-tuning of the GDD, the pulses can be compressed to 13.5 fs at an FTL of 8 fs. The recorded (panel a) and retrieved FROG traces (panel b) as well as the spectral intensity and phase (panel c) of the compressed pulses are depicted in Fig. 4.15.

Thus, using the above-mentioned range of techniques, we successfully created ultrashort near-infrared laser pulses well-suited for the generation of broadband mid-infrared radiation (see sec. 4.3).

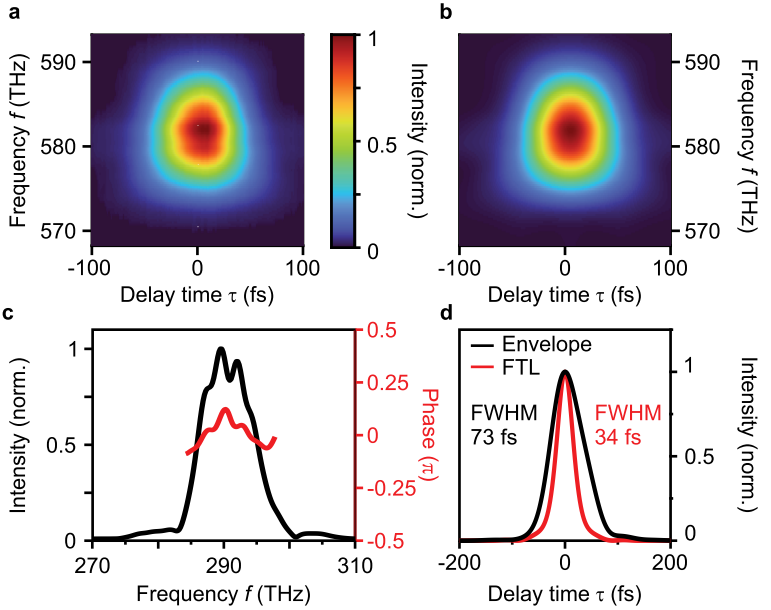


Figure 4.14 | FROG measurements of the HC1 output pulses. a, A FROG spectrogram of the compressed HC1 output pulses is recorded. **b**, The reconstructed FROG spectrogram nicely resembles the measured trace. **c**, The retrieved spectral intensity (black curve) and phase (red curve) of the reconstructed waveform are depicted. **d**, The temporal intensity envelope of the reconstructed pulse (black curve) shows an FWHM width of 73 fs, while the FTL pulse duration lies at 34 fs (red curve).

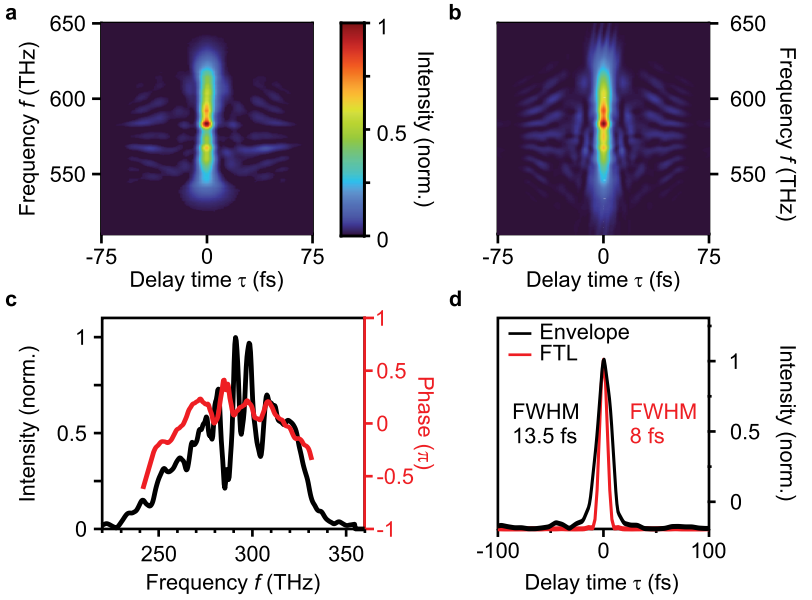


Figure 4.15 | FROG measurements of the dual-stage Herriott cell output pulses. a, A FROG spectrogram is recorded after compression of the HC2 output pulses. **b**, The reconstructed FROG spectrogram closely resembles the measured one. **c**, The retrieved spectral intensity (black) and phase (red) of the reconstructed waveform are depicted. **d**, The temporal intensity envelope of the reconstructed pulse (black) exhibits a duration of 13.5 fs at an FTL (red) of 8 fs.

4.3 Generation of single-cycle mid-infrared pulses

The frequency range between microwave and infrared radiation, typically spanning from a few hundred gigahertz to several terahertz, is commonly labelled the "terahertz gap". This frequency range has long been considered challenging to access and study due to the lack of suitable laser systems and technologies. Over the last decades, many ways to generate laser pulses within the terahertz gap have been developed. For example, vacuum electronic devices usually used for microwave generation, such as magnetrons, gyrotrons, and synchrotrons [Gan06], can undergo targeted modifications to enable their operation at terahertz frequencies. Nevertheless, if pulses suitable for LW-STM within this spectral range are desired, additional requirements must be met by the laser source. Specifically, to support single-cycle pulses in the terahertz and mid-infrared range, a sufficiently broadband frequency spectrum is crucial. The most critical pulse characteristic for the application of laser pulses in the field of lightwave-electronics [Sch14, Hoh15, Rei18, Sch19a] is the stability of the CEP. Thus, the origin of the CEP shall be revisited in the following explanations, closely following previous elaborations by Cundiff and Weiner [Cun10].
A train of pulses emitted by a laser oscillator is characterised by the periodic repetition of intensity spikes with a width of τ and a temporal separation between subsequent laser shots of $\frac{1}{f_{\mathrm{rep}}}$. Figure 4.16a displays such an exemplary pulse train with Gaussian-shaped intensity envelopes for each pulse. The spectrum of a singular pulse is broadband and exhibits a width that is inversely proportional to the pulse duration τ. Notably, subsequent pulses in the pulse train are not identical: The oscillating carrier wave does not necessarily peak at the peak of the envelope. The reason for this variability is that while the laser pulse is travelling back and forth within the oscillator, the different propagation velocities of the carrier wave (travelling at the phase velocity) and the peak of the intensity envelope (travelling at the group velocity) lead to a phase shift. This phase shift is different for each pulse emitted by the oscillator, leading to variations in the exact temporal shape of the electric field underneath the envelope. The temporal evolution of this phase is given by the so-called carrier-envelope phase ϕ_{CE}. As these pulses periodically reoccur, the Fourier transform of the whole pulse train consists of evenly spaced sharp lines underneath the envelope (see Fig. 4.16b). This so-called frequency-comb is however offset from a

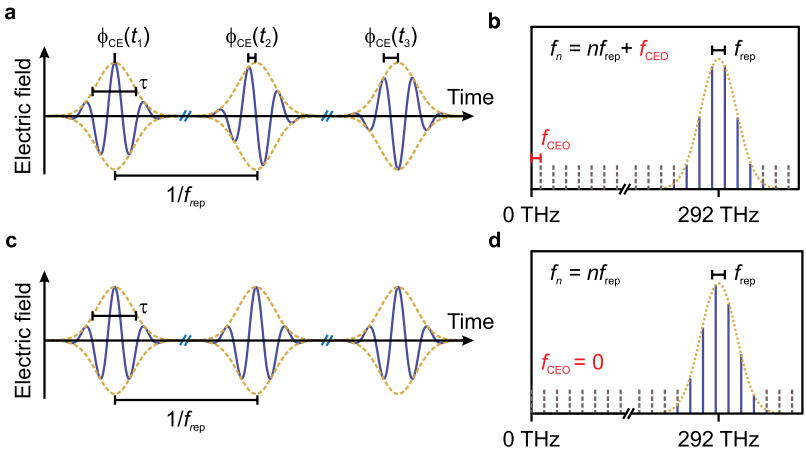


Figure 4.16 | Carrier-envelope phase. A pulse train emitted by a near-infrared laser oscillator ($f_{\text{centre}} = 292\,\text{THz}$), operating at a repetition rate of f_{rep}, features periodically repeating intensity spikes spaced in time by $\frac{1}{f_{\text{rep}}}$. **a**, For a non-CEP-stable pulse train, the temporal evolution of the relative phase between the electric field oscillating within the envelope and the peak of the envelope is described by the carrier-envelope phase ϕ_{CE}. **b**, The resulting Fourier spectrum exhibits periodically spaced spectral lines, the so-called frequency-comb, with a spacing of f_{rep}. The non-constant ϕ_{CE} leads to an offset of the comb from zero frequency by the carrier-envelope offset f_{CEO}. **c**, In the case of a CEP-stable pulse train, all pulses are identical. **d**, The constant carrier-envelope phase ϕ_{CE} in a CEP-stable pulse train leads to a frequency comb in which the individual lines are placed at integer multiples of f_{rep}.

frequency of zero by the carrier-envelope offset frequency f_{CEO} (see Fig. 4.16b). The individual spectral lines within the full pulse train can then be written as:

$$f_{\text{n}} = \text{n} f_{\text{rep}} + f_{\text{CEO}} \tag{4.24}$$

In a CEP-stable pulse train, the carrier envelope phase ϕ_{CE} is constant and therefore the shape of the oscillating electric field remains unchanged over time (see Fig. 4.16c). As a result, the frequency-comb is locked to a carrier-envelope offset frequency $f_{\text{CEO}} = 0$ and the individual spectral lines are given by $f_{\text{n}} = \text{n} f_{\text{rep}}$ (see

Fig. 4.16d). Only such phase-locked ($f_{\mathrm{CEO}} = 0$) pulses are suitable for our application. Creating these CEP-stable pulse trains constitutes the most challenging aspect of the optical setup.

In the endeavour to provide user-friendly and cost-effective laser sources for CEP-stable pulses, numerous research groups and commercial laser suppliers have developed publicly procurable table-top solutions in recent years [Dhi17, Lei23]. The most promising table-top sources providing intense, CEP-stable, and octave-spanning spectra involve the interaction of femtosecond laser pulses in the near-infrared domain with matter in a variety of different fashions. Several approaches that may come into question for our application will be introduced and analysed for their suitability.

For instance, photoconductive switches, which consist of small gold antennas fabricated on top of a semiconductor, can be used to generate low-frequency terahertz radiation [Aus84]. Upon illumination of the semiconductor between the two antenna electrodes with ultrashort laser pulses with a suitable photon energy, electrons are photoexcited over the bandgap and accelerated in a strong DC electric field. Due to the short lifetime of the charge carriers, a short current burst is established that according to Maxwell's equations emits electromagnetic radiation. Applied to suitable materials, this mechanism has been used to successfully generate broadband spectra with frequencies up to 30 THz [Sin20].

A second potentially applicable method is based upon the use of two-coloured laser filaments in a gas [Coo00]. By focusing high-energy near-infrared pulses at a frequency ω_0 and its second harmonic $2\omega_0$ into a gas jet, nonlinear interactions in the generated plasma can produce highly asymmetric terahertz waveforms. Disadvantageously, this technique is relatively unstable and requires extremely high pulse energies, limiting the achievable repetition rates to a few kilohertz when using table-top laser amplifier systems. As we need to work at high repetition rates in the range from hundreds of kilohertz to several megahertz to achieve reasonable data acquisition times, sources based on two-coloured filaments are not fitting for us.

Another promising technique to generate single-cycle terahertz pulses involves spintronic emitters, which harness the inverse spin Hall effect [Sei16]. Specifically, ultrashort near-infrared pulses are absorbed in a bilayer structure consisting of a ferromagnetic and a nonmagnetic metallic film within an external magnetic field. A superdiffusive spin current [Bat12] launched in the ferromagnetic layer is converted

into a charge current due to the inverse spin Hall effect. Owing to the ultrafast optical excitation, superdiffusion, and charge lifetimes of less than 100 fs, the ultrashort current burst leads to the pulsed emission of a phase-stable sub-cycle terahertz pulse [Kam13]. However, the limited optical skin depth of metals hinders the scaling of these emitters in thickness, necessitating large pump-pulse energies to achieve high terahertz field strengths, effectively limiting the repetition rate.

A recent promising technological advancement in the generation of high-field-strength, sub-cycle terahertz pulses utilises ultrafast shift currents in quantum well (QW) emitters [Mei22]. In this approach, ultrashort near-infrared pulses are absorbed in a custom-tailored Type-II band-aligned QW structure. The wave functions of the photoexcited electrons and holes are localised predominantly at opposite ends of the QW due to the inbuilt electric potential. As a result, the photoexcitation creates an ultrafast shift current spike, which is ultimately limited by the near-infrared pulse duration. The shift current leads to the emission of ultrashort terahertz pulses with frequency components up to 21.4 THz (full width at tenth of maximum, FWTM). Although the use of QW emitters offers promising avenues for generating short laser pulses in the multi-terahertz regime, it is still not well established and the achieved maximum multi-terahertz frequency of 21.4 THz (FWTM) is well below what is required for a single-cycle pulse in the mid-infrared spectral range.

Considering the current state-of-the-art single-cycle terahertz sources, the most versatile and tunable emitter for the accomplishment of our desired terahertz pulse parameters remains the highly-efficient and well-established technique of difference-frequency generation (DFG) within a nonlinear optical crystal. This second-order nonlinear optical process is described briefly in the following paragraph.

Difference-frequency generation. The most flexible means of generating ultrashort laser pulses in the terahertz and mid-infrared regime relies on nonlinear optical effects in non-centrosymmetric crystals. As already mentioned in section 4.2 and noted in equation 2.10, for intense light fields the polarisation induced by an external electric field can be expanded in a power series. Taking only the term quadratic in incident field strength into account, two incident electromagnetic waves oscillating at angular frequencies ω_1 and ω_2 create a polarisation oscillating at their beating frequency $\omega_3 = \omega_1 - \omega_2$ that emits an electromagnetic wave at the difference frequency ω_3 (see Fig. 4.17a).

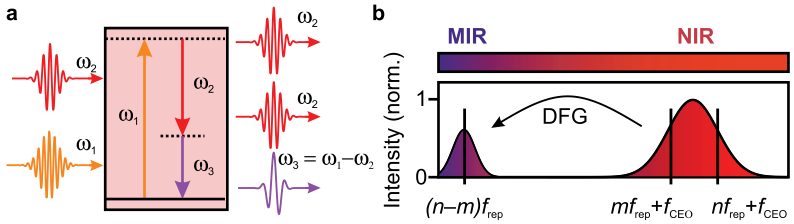


Figure 4.17 | Difference-frequency generation. a, Within a suitable nonlinear medium, the second-order nonlinear optical process of difference-frequency generation (DFG) allows for the creation of radiation at the beating frequency of two incoming electromagnetic waves oscillating at angular frequencies ω_1 and ω_2. Apart from the photon oscillating at $\omega_3 = \omega_1 - \omega_2$, two photons at ω_2 exit the medium. **b**, In the case of optical rectification, two frequencies within the same near-infrared femtosecond pulse $f_n = nf_{rep} + f_{CEO}$ and $f_m = mf_{rep} + f_{CEO}$ interact, leading to the emission of a difference-frequency component with a vanishing carrier-envelope frequency: $f_{diff} = (n - m)f_{rep}$. As any pair of two spectral components within the frequency comb can emit mid-infrared radiation at their difference frequency, a broad spectrum ultimately limited by the near-infrared bandwidth can be generated.

The two high-frequency components for the coherent mixing process do not necessarily have to originate from different pulses. In the special case of intra-pulse DFG, also termed optical rectification (OR), two spectral components within the same pulse interact. As all possible pairs of frequencies within the incident near-infrared spectrum create new frequency components, a broadband low-frequency spectrum will arise, featuring a maximum frequency given by the bandwidth of the near-infrared pulse. Figure 4.17b illustrates such an OR from the near-infrared (hundreds of THz) into the few- to multi-terahertz spectral regime [Ric94, Bon95, Kai00]. Notably, laser pulses generated by means of OR are inherently CEP-stable. As mentioned in the discussion following Fig. 4.16, the frequencies emitted by a pulsed laser are given by:

$$f_n = nf_{rep} + f_{CEO} \tag{4.25}$$

By nonlinear interaction of two frequencies f_n and f_m from the same pulse within a

second-order nonlinear optical medium, the resulting difference frequency:

$$\begin{aligned} f_{\mathrm{diff}} &= f_{\mathrm{n}} - f_{\mathrm{m}} \\ &= (\mathrm{n} f_{\mathrm{rep}} + f_{\mathrm{CEO}}) - (\mathrm{m} f_{\mathrm{rep}} + f_{\mathrm{CEO}}) \\ &= (\mathrm{n} - \mathrm{m}) f_{\mathrm{rep}} \end{aligned} \tag{4.26}$$

is independent of the carrier-envelope offset frequency f_{CEO}. The CEP-stability of mid-infrared pulses generated in this way was first demonstrated by Huber *et al.* [Hub00a]. The authors showed electro-optically sampled mid-infrared waveforms with frequencies surpassing 40 THz .

The different frequencies partaking in the DFG process shall henceforth be called "pump" frequency ($\omega_1 = 2\pi f_{\mathrm{n}}$), "seed" frequency ($\omega_2 = 2\pi f_{\mathrm{m}}$, m<n), and "idler" frequency ($\omega_3 = \omega_1 - \omega_2 = 2\pi(\mathrm{n} - \mathrm{m}) f_{\mathrm{rep}}$). As depicted in Fig. 4.17a, a second frequency component oscillating at ω_2 is generated simultaneously within the DFG process. This component shall be called "signal".

However, in order to attain a nonlinear polarisation within the medium that oscillates at $\omega_3 = \omega_1 - \omega_2$, the two incident spectral components must overlap temporally. Thus, only Fourier-transform-limited pulses can make use of the full near-infrared bandwidth. Therefore, the twofold importance of the duration of the near-infrared pulse for a given spectrum becomes apparent: First, shorter pulses lead to higher peak intensities and thus more efficient OR and second, pulses closer to the FTL increase the effective bandwidth.

Phase matching. In the conversion process from one high-energy photon with the energy E_1 to two low-energy photons (E_2 and E_3), energy and momentum conservation must be fulfilled. The former is satisfied by simply requiring:

$$E_1 = E_2 + E_3 \tag{4.27}$$

$$\rightarrow \hbar\omega_1 = \hbar\omega_2 + \hbar\omega_3 \tag{4.28}$$

The latter can be achieved by meeting the phase-matching (PM) condition, as will be explained in the following. An ultrashort incident near-infrared pulse and thus the phase of the nonlinear polarisation it creates within a crystal propagate at the group velocity of the pulse. At each position within the crystal, the oscillating

nonlinear polarisation leads to the emission of mid-infrared radiation. These mid-infrared frequencies travel at their respective phase velocities $v_{\mathrm{ph}} = \frac{c}{n_{\mathrm{MIR}}(\omega)}$ governed by the refractive index $n_{\mathrm{MIR}}(\omega)$ in the medium. Thus, if the phase velocities do not coincide with the group velocity, the generated mid-infrared radiation and the polarisation may partially interfere destructively. The most efficient mid-infrared generation is therefore achieved in the case of equal propagation velocities of the second-order polarisation and the mid-infrared radiation and can be written in terms of the wavevectors as:

$$\boldsymbol{k}_3 = \boldsymbol{k}_1 - \boldsymbol{k}_2, \tag{4.29}$$

where:

$$k_i = |\boldsymbol{k}_i| = \frac{n_i \omega_i}{c}, \qquad i = 1, 2, 3 \tag{4.30}$$

denote the wavevectors of the three different oscillating electromagnetic fields. This condition of vanishing wavevector mismatch $\Delta k = |\boldsymbol{k}_3 - (\boldsymbol{k}_1 - \boldsymbol{k}_2)|$ is readily fulfilled only in a few media and only for specific combinations of photon energies due to the material's dispersion. This can be circumvented by using birefringent crystals whose refractive index can be tuned. In a uniaxially birefringent crystal, a single direction, the so-called optic axis, governs the optical anisotropy of the medium. Electromagnetic radiation propagating parallel to the optic axis is always subject to the ordinary refractive index n_{o}, regardless of its polarisation. However, light propagating under an angle of θ with respect to the optic axis and with a non-zero electric field component along the optic axis experiences a different, extraordinary refractive index n_{eo}. This can be tuned by the PM angle θ according to the equation:

$$\frac{1}{n_{\mathrm{eo}}(\omega, \theta)^2} = \frac{\cos^2(\theta)}{n_{\mathrm{o}}(\omega)^2} + \frac{\sin^2(\theta)}{n_{\mathrm{e}}(\omega)^2} \tag{4.31}$$

wherein n_{o} and n_{e} are the ordinary refractive index and the principal value of the extraordinary refractive index, respectively. One distinguishes between positively ($n_{\mathrm{e}} > n_{\mathrm{o}}$) and negatively ($n_{\mathrm{e}} < n_{\mathrm{o}}$) birefringent media. In negatively birefringent, normally dispersive ($n_{\omega_3} < n_{\omega_1}$) media, which are used in this work, the photon at the highest frequency ω_1 has to be polarised extraordinarily such that the phase-matching condition in equation 4.29 is fulfilled. As a result, PM can be achieved in two different configurations:

- Type-I phase matching (eoo): The pump photon is polarised extraordinarily. The signal photon and the idler photon are polarised ordinarily.
- Type-II phase matching (eoe): The pump and idler photons are polarised extraordinarily, while the signal photon is polarised ordinarily.

The abbreviations in brackets – eoo and eoe – indicate the polarisation axes of the highest- to lowest-frequency photon from left to right. In summary, based on the available incident spectrum and the angle-dependent nonlinear susceptibility of the crystal, the suitable phase-matching type and angle must be chosen to fulfil the momentum conservation for the desired output spectrum.

Simulation of optical rectification in gallium selenide. All requirements discussed above must be taken into account to identify suitable materials for the generation of mid-infrared pulses using DFG. To briefly recapitulate, relevant factors include a high nonlinear susceptibility $\chi^{(2)}$ and broadband phase matching for the input laser spectra centred around 1030 nm. The layered crystal gallium selenide (GaSe) not only fulfils all these requirements but has also proven very versatile for optical rectification from the near-infrared into the terahertz and mid-infrared spectral regions. Specifically, it is transparent over a wide spectral region (0.7−18 µm) and has a particularly high effective nonlinear susceptibility ($\mathrm{d}_{22} = (86 \pm 17)\,\frac{\mathrm{pm}}{\mathrm{V}}$) [All09].

In order to investigate whether GaSe can act as a suitable emitter for broadband, CEP-stable, mid-infrared, and single-cycle pulses, simulations are performed prior to the experimental realisation. The reader is referred to appendix C.3.1 for details on the theoretical description and the simulation of optical rectification in nonlinear optical crystals. The final result from solving the nonlinear wave equation is given by:

$$I_3 = \frac{8 d_{\mathrm{eff}}^2\, \omega_{\mathrm{MIR}}^2\, I_{\mathrm{NIR},1} I_{\mathrm{NIR},2}}{n_1 n_2 n_3 \varepsilon_0 c^2} L^2 \mathrm{sinc}^2\left(\frac{\Delta k L}{2}\right) \tag{4.32}$$

The intensity I_3 of the generated mid-infrared radiation at a specific frequency ω_3 depends on multiple parameters such as the intensities at the pump and seed frequency (I_1 and I_2), the corresponding refractive indices at all three frequencies (n_1, n_2, and n_3) and the crystal thickness L. Moreover, the effective nonlinearity driven by the incident light d_{eff} and the wavevector mismatch Δk enter the equation. How

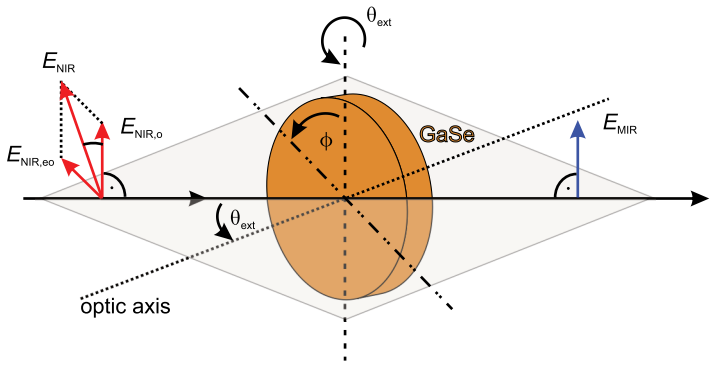


Figure 4.18 | Sketch of relevant angles and polarisations for DFG in gallium selenide. Two incident near-infrared electric fields, $E_{\mathrm{NIR,eo}}$ and $E_{\mathrm{NIR,o}}$, propagating along the horizontal black arrow are polarised along the extraordinary and ordinary axes of a gallium selenide crystal (GaSe). The optic axis is denoted by the dotted black line. In the illustrated case, the generated difference-frequency signal E_{MIR} is polarised parallel to $E_{\mathrm{NIR,\,o}}$, indicating type-I PM. By varying the azimuthal angle ϕ (rotation about the optic axis), the PM type and thus the polarisation of the mid-infrared radiation can be changed. The angle θ_{ext} (rotation around the dashed vertical black line) defines the efficiency of the conversion and the spectrum of the mid-infrared pulses.

the results of the theoretical analyses influence our experimental emitter design and geometry are presented in the following.

In this setup, the near-infrared pulses, which were previously spectrally broadened in the dual-stage multipass cell are to be focused into the nonlinear optical crystal GaSe. For the most efficient mid-infrared generation, the phase mismatch must be minimised and the nonlinear polarisation should be created along the crystal axis with the highest nonlinear susceptibility. The available degrees of freedom to tune OR and satisfy these conditions include the polarisation of the near-infrared pulse and the crystal orientation relative to the incident light. Setting the correct polarisation of the incident near-infrared light with respect to the crystal axes is therefore crucial. As sketched in Fig. 4.18, the nonlinear optical crystal can be rotated in the azimuthal direction by an angle ϕ and in the sagittal direction by the polar angle θ_{ext}. Light

incident on the crystal under the angle θ_{ext} is refracted at the air–crystal interface according to Snell's law. Within the crystal the propagation angle relative to the optic axis is called the (internal) PM angle θ. As explained above, for optimal optical rectification, the PM angle θ must be chosen such that the photon momentum mismatch in the spectral region of interest is minimised (see eq. 4.31). Moreover, the effective nonlinearity of GaSe is influenced by the PM angle θ as well as the azimuthal angle ϕ according to the following equation:

$$d_{\text{eff}}^{\text{eoo}} = d_{22} \cos(\theta) \sin(3\phi) \tag{4.33}$$

$$d_{\text{eff}}^{\text{eoe}} = d_{22} \cos^2(\theta) \cos(3\phi) \tag{4.34}$$

Evidently, the choice of PM type influences the optimal polar angle that should be employed.

For type-I PM, which will be used for the rest of this thesis, the generated mid-infrared radiation E_{MIR} is polarised perpendicular to the pump field ($E_{\text{NIR,eo}}$) and parallel to the seed field ($E_{\text{NIR,o}}$). Importantly, in the case of OR, where pump and seed field originate from the same near-infrared pulse, the polarisation of the incident light must be chosen such that it contains electric field components oscillating parallel as well as perpendicular to the optic axis. In other words, the pulse must exhibit non-zero projections onto the extraordinary and the ordinary axes of the nonlinear optical crystal. If an angle of 45° is chosen between E_{NIR} and the extraordinary axis, this leads to equally large electric fields ($E_{\text{NIR,eo}} = E_{\text{NIR,o}}$) along the two relevant crystal axes.

In order to find the optimal PM angle for the generation of the desired mid-infrared spectrum, the phase mismatch Δk must be determined for different angles θ and idler frequencies. More precisely, the field strength of a certain generated mid-infrared frequency is strongly influenced by the PM factor $|\text{sinc}(\Delta k L/2)|$. Calculating the PM factor for a GaSe crystal with a thickness of 50 µm and an incident near-infrared pump pulse with a centre wavelength of 1028 nm reveals a broadband region from approximately a frequency of 10 THz to 50 THz with a small phase mismatch (see Fig. 4.19a). Thus, in this frequency range, theoretically, large mid-infrared field strengths could be generated. As mentioned above, the possible difference frequencies emitted from a nonlinear crystal are ultimately limited by the bandwidth of

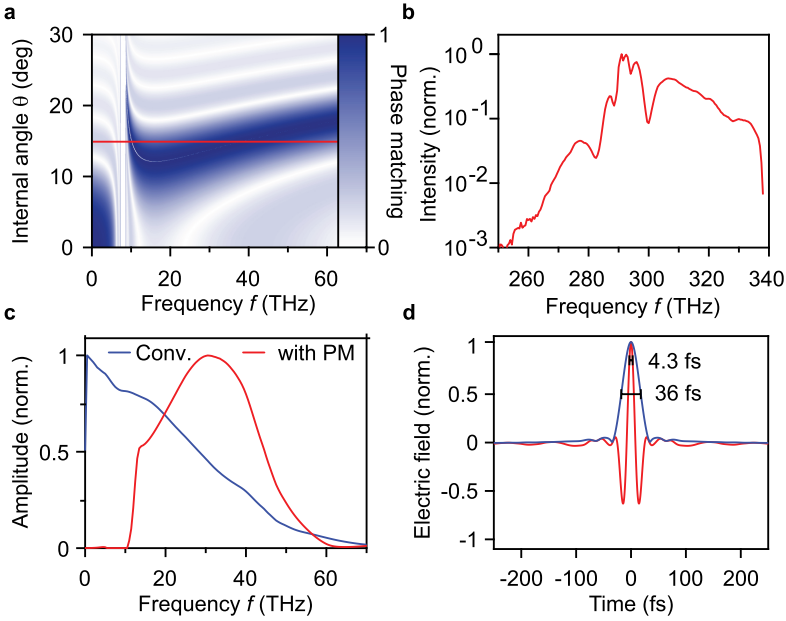


Figure 4.19 | Simulated optical rectification in GaSe. a, The phase-matching factor for a GaSe crystal with a thickness of 50 µm at a pump wavelength of 1028 nm is calculated for various internal PM angles θ. For an angle of 14°, broadband PM for frequencies spanning 10 to 55 THz is achievable. **b**, The HC output spectrum utilised for mid-infrared generation is plotted on a semi-logarithmic scale. **c**, The possible difference frequencies within the near-infrared spectrum are obtained by a convolution of the spectrum with itself (blue). When taking PM into account, the amplitude spectrum drastically changes its shape (red). **d**, The simulated mid-infrared transient (red) exhibits an ultrashort pulse duration with a field crest (top 90%) of only 4.3 fs. The intensity envelope possesses an FWHM of 36 fs.

the incident laser pulse. Taking the spectrum of the HC output into account (see Fig. 4.19b), the mid-infrared spectrum that can theoretically be generated for the known incident near-infrared pulse is simulated for an internal PM angle of $\theta = 14°$ (see Fig. 4.19a, red horizontal line). As Fig. 4.19c shows, the broadband near-infrared spectra would support difference-frequency components up to 60 THz (FWTM), represented by the blue curve in Fig. 4.19c. By multiplying this spectrum with the corresponding phase-matching factors and normalising the result, the red curve in Fig. 4.19c is obtained. This spectrum denotes the relative amplitude of the different mid-infrared frequency components that can be generated within a GaSe crystal with a thickness of 50 µm using the ultrashort near-infrared pulses created in the dual-stage HC. The resulting mid-infrared transient calculated via Fourier transform possesses an asymmetric pulse shape with a peak field duration (top 90%) of only 4.3 fs, rendering it highly suitable for the application as ultrafast bias pulse in LW-STM. To determine the optimal thickness of the employed GaSe crystal, I compared the resulting mid-infrared transients for varying crystal thicknesses. This analysis revealed that in order to allow for broadband PM and generate asymmetric waveforms, GaSe crystals with a thickness of 50 µm or less should be used.

Near-infrared focus in GaSe. Having shown that GaSe with a thickness of 50 µm is indeed a suitable candidate for the generation of ultrashort mid-infrared pulses, the next step would be to create sufficiently high peak intensities within the crystal. As the nonlinear effect of difference-frequency generation scales quadratically with the incident electric field, creating a sharp focus is the most straightforward way of increasing the output power of mid-infrared radiation. However, the smallest possible beam diameter is limited by the damage threshold of GaSe. Inconveniently, the precise damage threshold of GaSe for femtosecond pulses at megahertz repetition rates has not yet been investigated in detail in the literature. A general estimate for the peak intensity that is well tolerated by usual nonlinear optical crystals lies around $200\,\frac{\mathrm{GW}}{\mathrm{cm}^2}$. With the laser specifications of our setup, namely a pulse energy of 17 µJ and a pulse duration of 14 fs, the peak power reaches 1.2 GW. As a result, a Gaussian beam radius of $w_0 = 750\,\mu\mathrm{m}$ (880 µm FWHM) is already sufficiently small to reach a suitable peak intensity of $200\,\frac{\mathrm{GW}}{\mathrm{cm}^2}$. To generate a focus of this size, the output beam of HC2, which exhibits beam radii of $w_\mathrm{x} = 1650\,\mu\mathrm{m}$ and $w_\mathrm{y} = 1850\,\mu\mathrm{m}$, must be only slightly focused. This is achieved by using a deliberately slightly misaligned

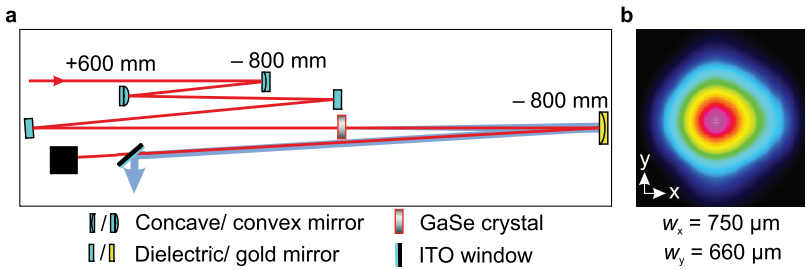


Figure 4.20 | Focusing into the gallium selenide crystal. a, For the generation of a suitable focal size, two curved mirrors (ROC = −800 mm and ROC = +600 mm) are placed at a distance of 137 mm from each other. The generation crystal is placed in the focus that is obtained 1570 mm after the focusing mirror. A gold mirror (ROC = −800 mm) is used to collimate the generated mid-infrared radiation. An ITO window separates the mid-infrared pulses from the remaining pump light. **b**, The experimentally determined intensity distribution in the focal plane reveals a Gaussian beam profile with waist radii of $w_x = 750\,\mu m$ and $w_y = 660$.

mirror telescope (see Fig. 4.20a). Using a calculation of the beam caustic (see Suppl. Fig. C.6) and tests, the optimal focal size was obtained by placing a concave mirror with ROC = −800 mm at a distance of 137 mm from a convex mirror with ROC = +600 mm. The beam waist at the focus amounts to $w_x = 750\,\mu m$ and $w_y = 660\,\mu m$ (Fig. 4.20b). Thus, the peak intensity at the focus exceeds 200 $\frac{GW}{cm^2}$, sufficient for the generation of mid-infrared transients through optical rectification within GaSe and still below the damage threshold.

Generation of ultrashort mid-infrared waveforms. As explained in the previous chapter covering the simulation of optical rectification (see sec. 4.3), in the difference-frequency generation process, pump photons are annihilated, while signal and idler photons are created. As a result, during the propagation through the nonlinear optical medium, the number of pump photons is gradually depleted, whereas the other two increase. In order to guarantee the highest possible mid-infrared output, the number of pump photons should exceed the number of seed photons at the crystal entrance face. In intra-pulse DFG, this is achieved by rotating the incident near-infrared polarisation. When polarised at an angle of 45° with respect to the extraordinary axis, the projections of the electric field onto the extraordinary and

ordinary axes are equally large. Thus, the electric fields of the pump and the seed part are equally strong. In order to optimise the optical rectification process, the near-infrared polarisation is chosen such that a slightly larger fraction of the incident light is polarised along the extraordinary axis, yielding a larger pump intensity than seed intensity of the incident light.

As disclosed above and illustrated in Fig. 4.19a, the spectral components that can be generated with high efficiency within GaSe critically depend on the PM angle θ. In order explore the optimal angle for our application, the experimentally generated waveforms and mid-infrared spectra as a function of θ_{ext} are recorded with an electro-optic sampling (EOS) setup (for details on EOS see sec. 2.3). External phase-matching angle and incidence angle are used interchangeably in the following. For the mid-infrared generation and detection, GaSe crystals with a thickness of 30 µm and 6 µm, respectively, were chosen. As the detection mechanism again is based on a nonlinear optical process, in general PM must also be considered there. The advantage of using such a thin GaSe crystal for detection is the extremely broadband PM, which legitimises neglecting the detector response [Kno18].

Figure 4.21a displays the normalised mid-infrared waveforms generated for the different angles of incidence θ_{ext}. Performing Fourier transforms of all waveforms leads to the spectra in Fig. 4.21b, spanning from a frequency of 10 to 38 THz (FWTM). Based on the data displayed in Fig. 4.21a, one would choose an external PM angle of either 0° or 50°, as this leads to the most asymmetric waveforms with an asymmetry factor (largest positive half-cycle in relation to largest negative excursion) of 1:1.7. Both angles have their respective advantages: In the case of zero-degree phase matching, Fresnel reflection at the air-crystal interface is minimised for s-polarised light. The greater transmission increases the pulse energy available for OR and thus leads to higher intensities of the generated mid-infrared radiation. Yet, the overall spectrum for an external PM angle of $\theta = 0°$ is centred around a lower frequency than the spectra generated for larger angles. Considering that higher frequencies lead to a better temporal resolution in LW-STM operation, an external PM angle of $\theta_{\text{ext}} = 50°$ would seem the best choice. However, the optimal waveform for the application in lightwave-driven STM is determined by its shape in the near field. To obtain the expected near-field waveform, the changes to the spectral amplitude and spectral phase upon propagation into the STM junction must be considered. The

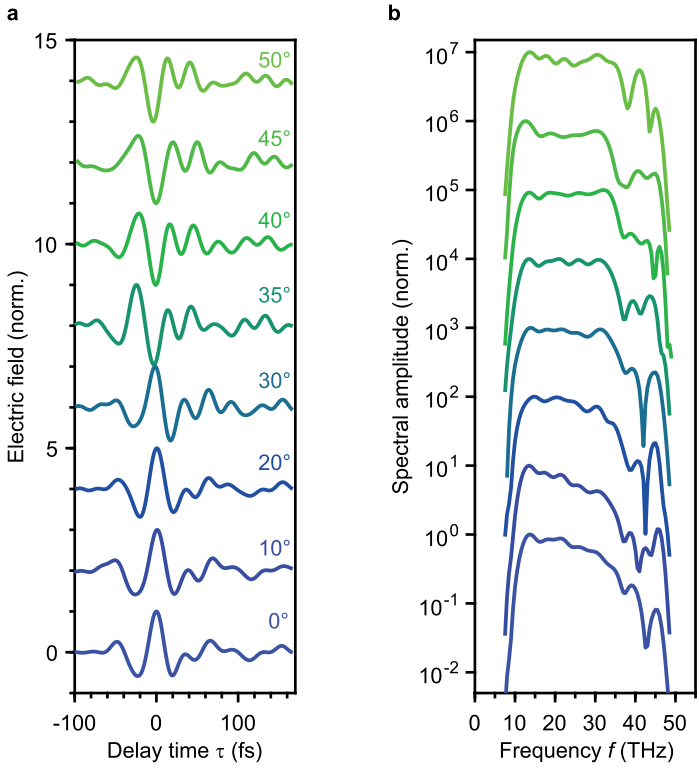


Figure 4.21 | Dependence of the mid-infrared transient on the angle of incidence. a, The mid-infrared transients generated with various angles of incidence θ_{ext} on the GaSe crystal are detected via electro-optic sampling. The most asymmetric waveforms occur for angles of $\theta_{\text{ext}} = 0°$ and $\theta_{\text{ext}} = 50°$. **b**, Performing Fourier transforms of the transients displayed in panel a results in the amplitude spectra shown here. With increasing angle of incidence, a clear shift to higher frequencies is visible.

transfer function from far to near field in the mid-infrared range is given by a field enhancement that is inversely proportional to the frequency and a CEP shift of $-\frac{\pi}{2}$. Applying this transfer function to the complex-valued spectra leads to the various waveforms depicted in Fig. 4.22a. In the near field, the most asymmetrical waveforms arise for angles of incidence of 35° and 40°, each characterised by an asymmetry ratio of 1:1.47. Both measured waveforms exhibit ultrashort pulse durations with a peak width (top 90%) of 6 fs and 6.7 fs for incidence angles θ_{ext} of 35° and 40°, respectively. Due to the shorter pulse duration, an external phase-matching angle of 35° was chosen. As a result of the frequency-dependent transfer function, the centre frequency of 21 THz in the near field is red-shifted compared to the centre frequency of 24 THz in the far field. The near-field waveform for the chosen angle of incidence of 35° exhibits a pulse duration of 32 fs (FWHM of intensity envelope), which warrants that only 1.5 oscillation cycles at the centre frequency occur underneath the intensity envelope. Given the ultrashort peak width of only 6 fs and an asymmetry ratio of almost 1:1.5, these pulses are highly suitable for the application in a MIR-STM.

Prior to guiding the generated mid-infrared pulses into the STM chamber, the remaining near-infrared pump and signal light must be filtered out. Otherwise, this high-intensity laser light would heat the sample system and the STM chamber considerably, making high-precision scanning tunnelling experiments impossible. To this end, a suitable indium-tin-oxide (ITO) covered glass window is used. The plasma frequency of this coating is chosen to lay between the mid-infrared and near-infrared spectral regions. The high-frequency near-infrared radiation is thus transmitted to largest extent and dumped, while the desired mid-infrared spectral region is reflected. After two reflections off ITO windows, all remaining near-infrared radiation is removed and the pure mid-infrared pulse energy can be measured in a pyroelectric detector. Using a suitable lock-in detection technique, the mid-infrared pulse energy is determined to be 179 pJ prior to the incoupling into the vacuum chamber. Guiding the mid-infrared radiation towards the STM, a large loss of pulse energy must be considered. The diamond window, through which the radiation enters the vacuum chamber, exhibits a transmittance of approximately 69% under normal incidence. Thus, only a pulse energy of 124 pJ remains after the window.This pulse energy is then focused into the STM junction using a parabolic mirror mounted on the scan head. The design of the STM scan head (see sec. 4.6) allows for the installation of

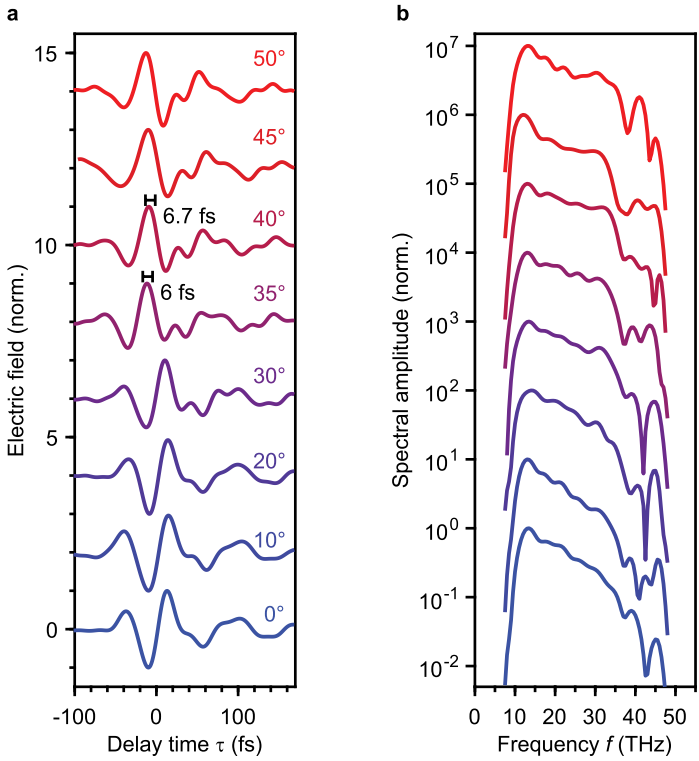


Figure 4.22 | Simulated near-field transients for various angles of incidence. **a**, The expected near-field mid-infrared transients within the tunnelling junction are obtained by applying the far-field-to-near-field transfer function to the transients shown in Fig. 4.21a. The most suitable near-field transients are created for angles of incidence on the GaSe crystal of $\theta_{\text{ext}} = 35°$ and $\theta_{\text{ext}} = 40°$. The corresponding peak widths (top 90%) amount to 6 fs and 6.7 fs, respectively. **b**, The spectral amplitudes are attained via Fourier transform of the waveforms shown in panel a.

a parabolic mirror with a very short focal length ($f = 7.5\,\mathrm{mm}$) close to tunnelling junction, which enables a small focal size and therefore large laser intensities. The collimated Gaussian beam diameter of 5.53 mm (see measurements in appendix C.3.3) leads to a focal size in the range of 30 µm ($\frac{1}{e^2}$-diameter), assuming Gaussian beam propagation with a beam quality of $M^2 = 1.3$. With the pulse duration of 35 fs, this yields an approximate peak intensity of $I_{\mathrm{peak}} \approx 700\,\frac{\mathrm{MW}}{\mathrm{cm}^2}$. The peak field would then amount to:

$$E_{\mathrm{peak}} = \sqrt{\frac{2I_{\mathrm{peak}}}{c\,\epsilon_0}} \approx 730\frac{\mathrm{kV}}{\mathrm{cm}} = 7.3 \cdot 10^{-3}\frac{V}{\mathrm{Å}} \tag{4.35}$$

As mentioned above, the propagation of radiation along the STM tip into the nanoscopic tunnelling junction leads to a considerable near-field enhancement. For mid-infrared radiation, the enhancement factor is expected to be several thousand $(3000 - 5000)$, which would increase the peak field to around $28\,\frac{\mathrm{V}}{\mathrm{Å}}$. Applying such large voltages across the STM junction would give us access energy levels far above the conduction band of solids or empty orbitals in molecules. As most interesting dynamics occur close to the Fermi energy, the mid-infrared transients would need to be attenuated to sufficiently low field strengths. Of course, additional power losses must be considered when including further beam steering in the beampath such as flipping polarity or performing MIR pump – MIR probe experiments. However, the bandgaps of most semiconductors or electronic excitations we would like to investigate lie in the range of a few volts, which provides us with a margin of two orders of magnitude.

4.4 Michelson interferometer and mid-infrared beam steering

In order to enable MIR pump – MIR probe experiments, identical copies of the mid-infrared waveform must be generated and temporally delayed relative to one another. Figure 4.23 shows a schematic of the Michelson interferometer employed. A beamsplitter separates the mid-infrared pulse into two equally strong parts. In each interferometer arm, the polarisation can be rotated and the pulse can be modulated by means of a mechanical chopper. The modulation is ultimately necessary for low-

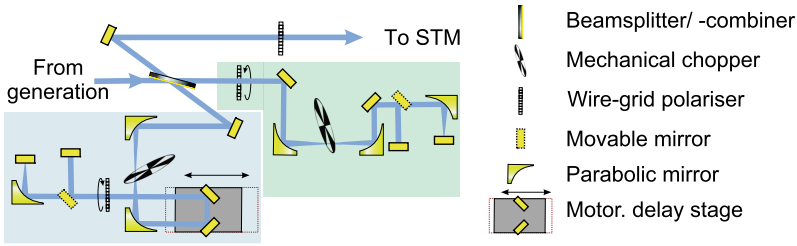


Figure 4.23 | Sketch of the Michelson interferometer. Using a gold-coated Germanium wafer as beamsplitter, identical copies of the incident mid-infrared pulse are generated. The two pulses within each pair are guided into two different interferometer arms (blue and green shaded box) and delayed with respect to each other using a motorised delay stage. By making use of the Gouy phase shift, the polarity of the transients can be separately adjusted in each interferometer arm by removing a movable mirror (indicated by the dashed outline). In each arm, the pulse train is periodically blocked using a mechanical chopper wheel and rotated in polarisation by means of a wire-grid polariser. The combining this with a stationary polariser (analyser) after the beamcombiner allows for the continuous control of the electric field strength of each arm.

noise lock-in detection of the lightwave-driven tunnelling currents. One arm can be delayed with respect to the other by means of a motorised delay stage. As will be explained in detail in the following section, additionally, the setup allows for the individual change of the polarity of the mid-infrared waveforms. Traditionally, to separate terahertz or mid-infrared radiation into two approximately equal parts, the light is polarised perpendicular to the plane of incidence on a bare germanium wafer under an angle of incidence of 45°. At this angle, the reflectivity of germanium for light with a wavelength of 11 µm, for example, lies at 48%, making this type of optical device a simple solution for a 50/50 beamsplitter. However, a main disadvantage of germanium wafers arises from multiple internal reflections of the incident pulse. Problematically, the reflected pulses leave the wafer in the same direction as the main pulse, with a spatial offset of several micrometres and a temporal offset that depends on the number of reflections N and the wafer thickness. Assuming a wafer with a thickness of 500 µm and taking Snell's law into account, the trailing reflections

are separated in time by $\Delta\tau_{\mathrm{N}}$ from the main pulse:

$$\Delta\tau_{\mathrm{N}} = N \cdot \frac{2 \cdot 500\,\mu\mathrm{m}\, n_{\mathrm{Ge}}}{c_0} \cdot \cos\left(\arcsin\left(\frac{\sin 45^\circ}{n_{\mathrm{Ge}}}\right)\right) = N \cdot 13.0\,\mathrm{ps} \tag{4.36}$$

These spurious reflections cannot easily be separated from the main pulse prior to focusing onto the tip, thereby complicating precise experimental control. To circumvent this, the germanium wafer is placed at the Brewster's angle with respect to the incident light. To create a non-zero reflectivity, a gold layer much thinner than the electromagnetic skin depth at 20 THz ($\delta_{\mathrm{Au}} = 25.7\,\mathrm{nm}$) is evaporated on the front side. As the layer is so thin, a non-zero portion of the electric field can still penetrate the gold layer and is transmitted through the wafer. The thickness of the gold layer should be chosen such that the beamsplitter ideally allows for an identical reflectivity and transmissivity of 50%. Due to the limited precision of the evaporation process of the $(3 \pm 2)\,\mathrm{nm}$ thick gold layer onto the wafer, a beamsplitter with a ratio of 55:45 (reflectivity:transmissivity) was fabricated.

After the original beam is split using the described setup, the two parts of the original beam travel along almost identical paths. The identical components are explained below. In order to temporally delay the two interferometer arms with respect to each other, a mechanical delay stage is placed in one arm. This stage allows for a variable travel distance of $\pm 25\,\mathrm{mm}$, corresponding to a delay time of about $\pm 83\,\mathrm{ps}$. Both beams are reflected back towards the gold-coated germanium wafer. Here, again parts of each beam are reflected and transmitted, allowing the wafer to effectively function as a beamcombiner.

The optical components that are alike in the two arms include mechanical choppers, pulse-flipping optics, and wire-grid polarisers. The optical choppers are used to periodically block the mid-infrared pulse trains. As a result, the tunnelling currents driven by the mid-infrared pulses will also be modulated at this frequency. This allows for the application of lock-in amplification of the lightwave-driven tunnelling current to increase the signal-to-noise ratio. When using mechanical choppers, special care must be taken to ensure that the pulse train is periodically either completely blocked or unblocked. It must be avoided that pulses are clipped by the chopper wheel. This can be achieved by choosing the aspect ratio between the chopper opening and the beam diameter as large as possible. To this end, in each arm of the interferometer,

the beam is focused using an off-axis parabolic mirror with a focal length of 3". In each focal plane, a mechanical optical chopper blade is placed to modulate the mid-infrared pulse trains. The diverging mid-infrared beams are recollimated using identical off-axis parabolic mirrors with a focal length of 3".

Flipping the polarity of the mid-infrared waveforms. It is of utmost importance to be able to apply positive and/or negative ultrafast voltage pulses across the tip-sample junction. In our setup, the polarity of the waveforms can be switched by making use of the Gouy phase shift. The so-called Gouy phase denotes an additional phase contribution that a Gaussian beam acquires along its propagation direction. The Gouy phase is related to the distance to the beam waist $z = 0$ and the Rayleigh length z_R (see eq.4.3):

$$\phi_\mathrm{G}(z) = -\arctan\Big(\frac{z}{z_\mathrm{R}}\Big) \tag{4.37}$$

According to this equation, when light passes through a focus from $z = -\infty$ to $z = +\infty$, it acquires an additional CEP shift of:

$$\Delta\phi = -\arctan\Big(\frac{\infty}{z_\mathrm{R}}\Big) - \arctan\Big(\frac{-\infty}{z_\mathrm{R}}\Big) = \pi \tag{4.38}$$

Making use of this phase shift in each of the two interferometer arms, the polarity can be flipped individually. This allows for MIR pump – MIR probe measurements in which either pulse can access either the valence or the conduction band of a semiconducting sample.

Varying the mid-infrared field strength. In addition to flipping the polarity of the bias applied across the tip-sample junction, the bias magnitude is adjustable using electronically rotatable wire-grid polarisers placed in each interferometer arm separately. By rotating the polarisation of the mid-infrared laser pulses and subsequent passing through a fixed analyser, the mid-infrared field strength can be tuned.

In summary, the employed mid-infrared optics allow for the generation of identical copies of the pulse, delaying these with respect to each other. To allow for various different experimental designs, these pulses can be separately flipped in polarity, tuned in field strength, and modulated at different frequencies. Notably, by changing the phase-matching angle of the optical rectification, even the centre frequency can be adapted to the experimental goal. This versatile source is thus able to be used for

the investigation of ultrafast dynamics excited by applying tunable voltage pulses in the positive or negative direction. The bias can either be employed as a force pulse (see sec. 2.6), or to charge the sample by removing (see sec. 2.3) or injecting (see sec. 2.4 and Chapter 3) a single electron within less than half a cycle of the driving field.

4.5 Generation of ultrashort optical pump pulses

The source of ultrashort pump pulses described in this chapter was planned and designed by me. In the build-up, Johannes Hayes, Peter Menden, and Samuel Balonier aided me and played crucial parts for the timely completion.

The previous chapters have explained the generation of mid-infrared transients that open a tunnelling window shorter than 10 fs across the tip-sample junction. By means of a Michelson interferometer and multi-component beam-steering optics, these ultrashort voltage pulses allow for pump-probe experiments with tunable field strength and flippable polarity. Single electrons can thus be injected into the sample or extracted from the sample and the effect of the resulting dynamics on the conductivity can be probed by a second mid-infrared pulse. However, if we are interested in the dynamics that are induced by promoting electrons from filled states into unoccupied states of a semiconductor or molecule, a different excitation pathway is necessary. By the absorption of photons with a suitable energy, a multitude of higher-lying energy levels in the sample can be reached and the dynamics of their relaxation process investigated. For example, pump-probe experiments utilising visible pump and mid-infrared probe pulses have been shown to provide information on the lifetime of excitons and hot electrons [Mer19a] as well as on the propagation of collective excitations such as surface polaritons [Hub16]. For these reasons, we aimed to develop a setup that permits the excitation of the sample with pulses in the visible domain. After the dual-stage HC, the broadened near-infrared spectra already contain spectral components with photon energies in the right ballpark. The spectrum after the first HC spans a wavelength interval between 1013 nm and 1047 nm, while after the second HC already spectral components down to a wavelength of 950 nm are available for pumping the sample. This facilitates the excitation of a multitude of low-energy transitions within solid-state and molecular samples with energies up to 1.3 eV. However, a large number of interesting transitions in condensed-matter systems lie at higher energies and thus in the regime of visible photons. Semiconducting bandgaps are typically situated at energies between 1 eV and 2 eV, which corresponds to wavelengths between 620 nm and 1200 nm. In order to resonantly excite matter at these energies, pump photons with an energy larger than the laser fundamental of 1.2 eV are necessary. This can be achieved by making use of nonlinear frequency

conversion to create a source of ultrashort pulses in the visible regime.
For this purpose, a small portion of the HC1 output energy is split off using a broadband half-wave plate and a polarising beamsplitter (see Fig. 4.9). This portion of the beam containing a pulse energy of 1.74 µJ, henceforth referred to as "pump arm", is used for the generation and amplification of visible pulses. Before doing so, another aspect must be considered: namely the temporal overlap of the different pulses within the setup. Specifically, the majority of HC1 output-pulse energy (23 µJ) is guided through a second HC for further broadening and finally used for the generation of mid-infrared radiation. Ultimately, the pump pulses and the mid-infrared pulses, whose generation scheme has been described in a previous section (sec. 4.3), must be spatially and temporally overlapped within the STM junction. Thus, the optical path length from the laser amplifier output to the STM junction must be precisely controlled and matched with micrometre precision. Particularly, the additional distance that the mid-infrared generating part of the laser beam (referred to as "MIR-arm") traverses within the second broadening stage must also be added to the split-off pump arm. To this end, an empty resonator was set up, as is depicted in Fig. 4.24 (orange background). Subsequently, the near-infrared pulses are used to generate a supercontinuum and the second harmonic of the incident frequencies. Tunable parts of the supercontinuum can then be amplified in a dual-stage non-collinear optical parametric amplifier using the second-harmonic photons. To temporally compress the visible pulses to their respective FTL, a prism compressor will be set up. The three major stations the pulses pass on their way from the HC1 output to the STM – DMPC (see Fig. 4.24, orange background), optical parametric amplifier (Fig. 4.24, green background), and prism compressor (Fig. 4.24, blue background) – will be explained in detail in the remainder of this section.

Delay multipass cell (DMPC). The additional beam path necessary in the pump arm in order to temporally overlap it with the MIR-arm could be simply achieved by reflecting the beam back and forth across the optical table until the necessary length is reached. However, small variations in the incoming beam angle, mirror positions, or even air density will lead to large position fluctuations after the long propagation distance. The issue of air density fluctuations and turbulences can be mitigated by a housing possessing only small holes for the in- and outcoupling of radiation. However, beam-pointing variations and mirror vibrations cannot be overcome in this way.

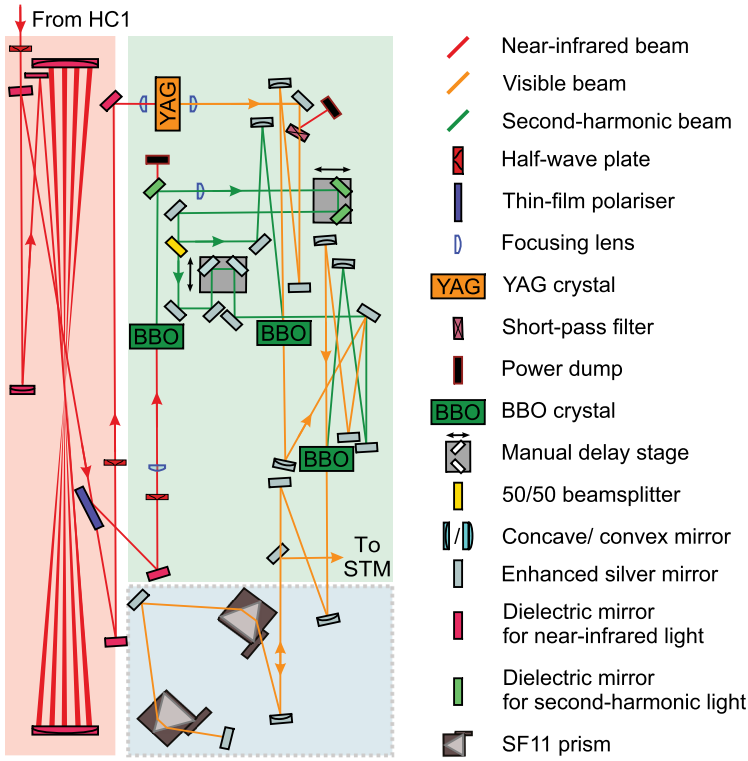


Figure 4.24 | Schematic of delay multipass cell, two-stage non-collinear optical parametric amplifier, and prism compressor. The ultrashort near-infrared laser pulses exiting HC1 are guided through a delay multipass cell to add a beampath of 11.4 m (red box). After the DMPC, using a thin-film polariser, a small part of the incoming near-infrared pulse energy is split off and used to generate a supercontinuum spanning from near-infrared to visible wavelengths inside a YAG crystal. The majority of the pulse energy is focused into a BBO crystal for frequency-doubling. After passing a manual delay stage, this second-harmonic light is split into two equally strong parts using a 50/50 beamsplitter. The two parts are separately focused into two subsequent non-collinear optical parametric amplifier (NOPA) stages using BBO to amplify tunable parts of the supercontinuum spectrum. The whole NOPA setup is enclosed in the green box. In order to compress the visible pulses, a prism compressor on the basis of two highly-dispersive SF11 prisms will be set up (blue box).

Therefore, the necessary elongation of the beampath is achieved using a so-called self-imaging multipass cell.
The designed self-imaging multipass cell, which shall be called delay multipass cell (DMPC) in the following, consists of two focusing mirrors forming a resonator in which the beam is reflected back and forth. In its design, this component is identical to a Herriott-type broadening cell from which the nonlinear medium has been removed. As we do not wish to imprint any GDD on the pulses, the resonator mirrors are not chirped. As explained in section 4.2, the incoming laser beam is focused into the centre of the resonator and then is reflected back and forth between the two focusing end mirrors several times (see Fig. 4.24, top). Importantly, in contrast to an HC, where a high peak intensity at the resonator centre is desired to achieve high SPM, the focal size is not relevant within the DMPC. The exact number of round trips until the beam is coupled out again depends on the exact design choices of the resonator. In case of the DMPC set up, the incoming laser beam is focused into the centre of a resonator using a curved mirror ($f = 750\,\mathrm{mm}$). The two end mirrors ($f = 250\,\mathrm{mm}$) are placed at a distance of 812 mm from each other. The beam is reflected back and forth 14 times allowing for large propagation distances of in my case 11.4 m. During these 14 round trips through the resonator, the laser spot travels circularly around the centre axis of the end mirrors five times. Using the nomenclature introduced in section 4.2 on Herriott cell design, this corresponds to $n = 7$ and $m = 5$. Figure 4.25 displays the calculated order and positions where an incident pulse coupled in through a small scraper mirror placed in front position 0 would impinge on the two resonator mirrors. By deliberate minor misalignment of the cell, the in- and outcoupled beams propagate under slightly different angles and can be separated using a D-cut mirror. By using a DMPC, one can introduce large beam paths while maintaining a setup which has a table-top suitable footprint of around $100\,\mathrm{cm} \times 20\,\mathrm{cm}$. The resonator design chosen exhibits a self-correcting potential against fluctuations of beam pointing and mirror deformations, effectively keeping the total beam length and position very stable.
The focusing mirrors' radii of curvature as well as the mirror distance are set such that the beam diameters on the mirrors stay roughly equally large for all passes through the cell. In order to achieve this, mode-matching considerations must be made. This allows for the minimisation of changes to the beam diameter in spite of

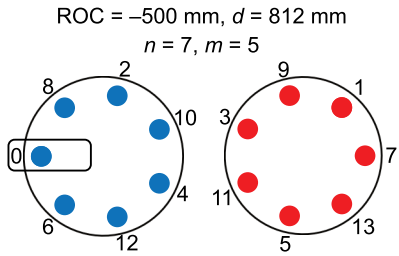


Figure 4.25 | Calculated spots of incidence in the DMPC. In order to obtain a certain optical path travelled within the multipass cell while maintaining a self-imaging behaviour, only certain mirror separations are permitted. For the radii of curvature of the resonator end mirrors (ROC $= -500$ mm), a separation d of 812 mm will result in seven reflections per mirror ($n = 7$) and five revolutions in the azimuthal direction ($m = 5$) prior to the outcoupling of the beam.

a travelling distance of 11.4 m.

The output pulses of the DMPC are then used for the generation and amplification of suitable visible pulses. Figure 4.24 provides a sketch of the relevant steps and optical components used. The pulse energy is split into two parts, which are explained in detail in the following paragraphs. One part is used to drive the highly nonlinear process of supercontinuum generation, while the other part is frequency-doubled by means of the second-order nonlinear process of second-harmonic generation. For the amplification of the supercontinuum, the process of difference-frequency generation is utilised in a dual-stage NOPA.

Supercontinuum generation. For the optical excitation of a multitude of electronic transitions in condensed matter, photon energies in the range of $1-2$ eV would be highly advantageous. The near-infrared spectrum after the first HC spans from a wavelength of 1013 nm to 1047 nm and thus spans photon energies from 1.18 eV to 1.22 eV. In order to create spectral components with higher photon energies, we make use of the extremely nonlinear optical process of supercontinuum or whitelight (WL) generation. Specifically, a pulse energy of 0.26 µJ of the DMPC output is focused tightly ($f = 25$ mm) into an yttrium aluminium garnet (YAG) crystal with a thickness of 5 mm. The tight focus in combination with the short pulse duration leads to large peak intensities so that higher orders of nonlinearities become important (see eq. 2.10). A multitude of processes including self-phase modulation (see eq. 4.18), four-wave mixing, and soliton based dynamics occur simultaneously within a thin plasma filament created by the extreme field strength. The length within the

crystal over which these processes take place is governed by a very delicate balance of two competing effects: self-focusing (see sec. 4.2) and diffraction. Under the correct conditions, this leads to the formation of a filament or plasma channel over the whole crystal thickness. Using an iris diaphragm and a lens with a focal length of 25 mm on a linear stage, the focusing conditions can be adjusted to achieve optimal spectral broadening within the YAG crystal. I deliberately upchirp the supercontinuum by using a short-focal-length ($f = 50\,\mathrm{mm}$) and therefore very thick fused silica lens for recollimation. After filtering out the remaining near-infrared fundamental spectrum using a shortpass filter, the remaining generated visible part of the spectrum spans wavelengths from 550 nm to 1000 nm (see Fig. 4.26). Despite the short-pass filter with a cut-off wavelength of 950 nm, a large contribution from the fundamental beam remains. The broad spectrum now contains photon energies sufficiently large to investigate the multitude of interesting systems mentioned earlier: electronic excitations over the bandgap of semiconductors, the creation of electron–hole pairs, overcoming the HOMO–LUMO energy barrier in molecules, and many more. However, supercontinuum generation is a very inefficient process that reaches an output-pulse energy of 14.3 nJ (within the spectrum spanning wavelengths from 550 nm to 1000 nm) for a near-infrared input energy of 260 nJ. When focusing such weak pulses with a low number of photons onto the sample, the excitation density – i.e., the excited atoms or molecules per unit area – would be very small, leading to small pump-induced changes to the sample properties that may prove challenging to detect. To overcome the issue of low pump pulse energy, the WL spectrum can be amplified using a so-called optical parametric amplifier, which will be explained in the following paragraph.

Second-harmonic generation. The basic idea of a optical parametric amplification (OPA) lies in siphoning energy from a high-intensity photon and splitting it into two lower energy photons. As the spectrum to be amplified, the previously generated WL, lies in the range of photon energies between 1.24 eV and 2.25 eV (wavelengths of 550−1000 nm), photons with an energy higher than that must be generated. By means of second-harmonic generation (SHG) of the near-infrared HC1 output, photons with a frequency of $2 \times \omega_{\mathrm{NIR}} = 2 \times 291\,\mathrm{THz} = 582\,\mathrm{THz}$, which corresponds to energies of 2.13 eV, are created.

The necessary near-infrared pulse energy for this purpose is split off from the DMPC

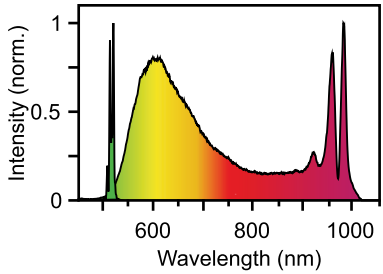


Figure 4.26 | Spectra of supercontinuum and second harmonic. The short-wavelength part of the broadband supercontinuum spectrum generated in a YAG crystal spans from the fundamental wavelength (1028 nm) down to 550 nm with a peak at a wavelength of 600 nm. The second-harmonic spectrum (green) is centred around a wavelength of 515 nm.

output. Near-infrared pulses with an energy of 1.23 µJ are focused (f = 250 mm) into a β-barium borate (BBO) crystal with a thickness of 5 mm. Notably, BBO is a non-centrosymmetric material with a high damage threshold, wide transparency range, and high nonlinear coefficient. Due to its symmetry, it features a non-vanishing second-order susceptibility $\chi^{(2)}$, leading to the emission of the second harmonic of the incident wavelength under the correct phase-matching conditions. Harnessing the nonlinearity of the BBO crystal, the spectrum of the near-infrared pulses is partially frequency-doubled into the visible regime spanning from 490 nm to 550 nm. The process of SHG is very efficient with a conversion rate of 60%, leading to a pulse energy of 732 nJ in the green spectral range. The remaining laser fundamental is filtered out using dielectric mirrors whose reflectivity is maximised for a central wavelength of 515 nm while maintaining a high transmissivity for 1030 nm.

Subsequently, the power of the green second-harmonic light (see Fig. 4.26, green) is split equally into two parts, which are used to amplify the visible supercontinuum in two optical parametric amplification stages placed in series.

Non-collinear optical parametric amplifier. In general, OPA relies on difference-frequency generation (see explanation following eq. 2.10 and sec. 4.3) between a pump photon and a seed photon, creating a so-called idler photon. Simultaneously, in order to satisfy energy conservation:

$$\hbar\omega_{\text{pump}} = \hbar\omega_{\text{signal}} + \hbar\omega_{\text{idler}} \tag{4.39}$$

a second photon with the same energy as the seed, the signal, is emitted. In this OPA setup, the green second harmonic centred at a wavelength of 515 nm provides the

pump photon, while the WL pulse donates the seed photon. Both pulses are spatially and temporally overlapped within a second-order nonlinear medium. Specifically, BBO was selected for the same reasons as outlined in the preceding subsection, along with the potential to perform broadband phase matching for DFG. For parametric amplification to set in, energy and momentum conservation must be fulfilled, as described in section 4.3 and appendix C.3.1. To achieve this, the phase-matching angle of the BBO crystal can be adjusted. In contrast to collinear OPA, where pump and seed beams travel collinearly, we opted for the so-called non-collinear OPA (NOPA). In this implementation, the pump and seed beams impinge on the nonlinear crystal under a small acute angle. The freedom to choose the exact angle creates an additional tuning knob on the minimisation of the group-velocity mismatch in a large spectral range [Rie00, Cer03].

To choose the part of the WL spectrum to be amplified, the strong chirp introduced by the collimating lens becomes beneficial. As a result of the dispersion of the thick lens used for collimation, longer wavelengths travel at the front of the beam, while short wavelengths are located further in the back. By changing the length of the SHG arm using a delay stage, temporal overlap with different parts of the strongly chirped WL pulse within the BBO crystal can be reached. To accomplish the most efficient amplification, the PM angle of the BBO crystal can additionally be modified.

To further increase the pulse energy of the visible spectrum, the pre-amplified pulses are focused into a second NOPA stage with the second half of the second-harmonic light acting as pump pulses. Besides the advantage of higher-energy visible pulses, a double-stage NOPA is much more stable than a single-stage NOPA when used in the so-called pump-saturation regime. The issue of instability poses a grave limitation to the beam quality, especially when a supercontinuum is used as seed pulse for DFG: In general, the delicate process of supercontinuum generation is highly prone to irregularities like small beam-pointing or power fluctuations and the experimental conditions including air pressure and composition. Such factors may thus lead to large variations in the supercontinuum output power of the YAG crystal. As a consequence, if such an unstable pulse is amplified within a NOPA, the instabilities are amplified as well. This can be prevented by operating the NOPA in the pump-saturation regime, which is achieved if the number of pump photons and not signal photons limits the conversion process. Advantageously, the pump pulse exhibits less

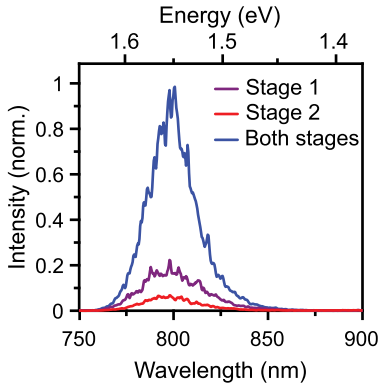


Figure 4.27 | Amplified spectrum for one or two NOPA stages. The amplification of the spectrum centred at a wavelength of 800 nm in stage 1 (purple), stage 2 (red), and both simultaneously (blue) is depicted. The spectra are normalised to the peak of the spectral intensity after the dual amplification. A clear nonlinear increase in intensity when amplified twice becomes apparent.

sources of instabilities like variations in the number of photons, since the second-order nonlinear process of SHG is relatively stable. Hence, we split the pump beam into two parts and use each half as a pump pulse for one NOPA setup, respectively, which are positioned consecutively.

Figure 4.27 impressively demonstrates the nonlinear increase from single- to dual-stage amplification for a central wavelength of 800 nm. The NOPA allows for the amplification of the seed from a pulse energy of few nanojoules to 70 nJ for a second-harmonic pump energy of around 333 nJ in each stage.

The resulting pulse energies for other seed spectra are summarised in the following table.

Central wavelength (nm)	640	700	750	800	850	900
Power (mW)	125	79	75	68	66	77
Pulse energy (nJ)	123	72	77	70	68	79

Figure 4.28 displays the amplified spectra for different centre wavelengths after the two stages. The intensities are normalised to the integral over the respective spectral intensity and multiplied by the detected power to remove artefacts created through different incoupling into the spectrometer fibre.

In summary, the designed and aligned dual-stage NOPA is able to create high-energy, tunable pulses in the visible spectral region exhibiting central wavelengths between 640 nm and 900 nm. This corresponds to photon energies between 1.37 eV and 1.94 eV.

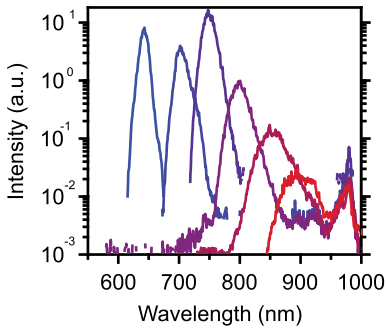


Figure 4.28 | Wavelength-dependent amplification. Amplified spectra after two NOPA stages for various centre wavelengths (640 nm, 700 nm, 800 nm, 850 nm, and 900 nm) are plotted on a semi-logarithmic axis. Tuning the phase-matching angle and the delay time between the pump and supercontinuum pulses, the spectrum to be amplified can be chosen. The integrated spectral intensity is normalised and then scaled by the measured power.

Pulse compression. In order to ensure a temporal resolution below 35 fs within the STM, the optical pulse should possess a duration in the same range. The duration of the NOPA output pulses for a centre wavelength of 800 nm were determined by means of SHG-FROG. The measured and retrieved FROG spectrograms are illustrated in Fig. 4.29. The measured spectrogram is relatively uniform and exhibits its centre at a frequency of 750 THz. The retrieved spectrogram resembles the measured one best for the intensity and phase spectrum depicted in Fig. 4.29c. A clear residual chirp is visible in the strong curvature of the spectral phase. The pulse envelope with an FWHM of 130 fs can theoretically be compressed close to the FTL of 25 fs by imprinting a GDD of $-390\,\mathrm{fs}^2$. The amplified pulses are relatively narrowband due to the limited temporal overlap of the long WL pulses with the short SHG pulses within the DFG crystal as well as due to the narrowband phase matching within the thick BBO crystal. Using a prism compressor, we plan to remove the residual chirp and achieve close-to-FTL pulses.

Shorter NOPA output pulses are technically achievable by reducing the chirp of the WL pulses prior to amplification or the thickness of the BBO crystals within the amplification stages. However, a delicate balance between power, temporal resolution in the STM, and resonant excitation must be attained. Moreover, for shorter pulses and therefore more broadband spectra, higher-order dispersion becomes ever more crucial. Since third-order dispersion cannot be compensated for in a prism compressor, a hybrid prism-grating compressor would be required. The trade-off between pulse

duration and monochromatic/resonant excitation established in this setup allows for pulses with bandwidths between 9 THz and 13 THz while maintaining pulse durations below 40 fs.

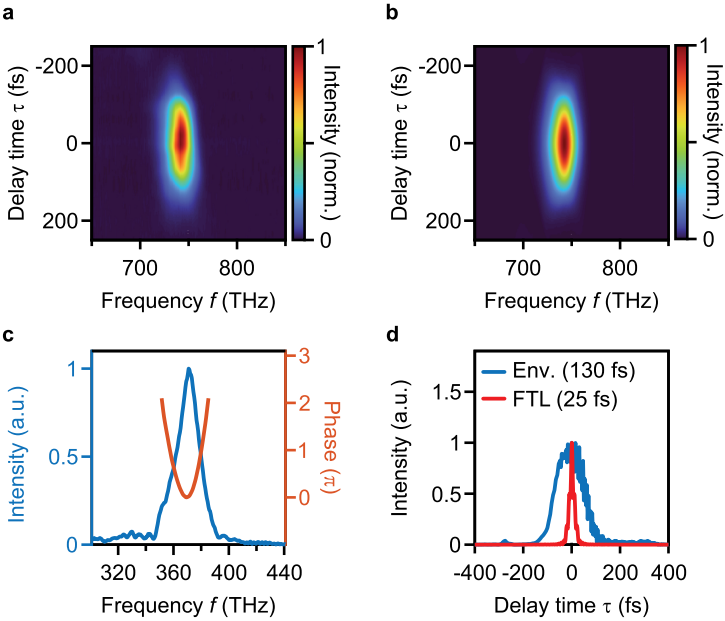


Figure 4.29 | FROG measurements for the NOPA output at a wavelength of 800 nm. a, The measured FROG spectrogram features a relatively uniformly shaped peak centred around 750 THz. **b**, The reconstructed FROG spectrogram nicely resembles the measured trace. **c**, The retrieved intensity spectrum peaks at a frequency of around 372 THz. The phase spectrum reveals a residual chirp of 390 fs^2. **d**, The retrieved intensity envelope (Env.) of the pulse features a duration (FWHM) of 130 fs, while the Fourier-transform-limited pulse duration lies at 25 fs.

In conclusion, by making use of spectral broadening in a Herriott cell, SHG, and a dual-stage NOPA, ultrashort visible pulses with adjustable centre wavelengths can be generated. In various experimental designs, these pulses can be used to optically excite a multitude of different systems in the LW-STM and visualise femtosecond dynamics on the atomic scale. For instance, the tunable wavelength (640−900 nm) and thus photon energies between 1.4 eV and 2.0 eV of the NOPA output pulses make resonant pumping of a wide range of excitonic or electronic transitions possible. Moreover, if smaller photon energies (<1.2 eV) are desired, the broad near-infrared pulses emitted by the dual-stage HC can be directly focused onto the sample. For large-bandgap semiconductors or the excitation of electrons far over the band edge, one can employ the second-harmonic pulses (with a photon energy of 2.4 eV) generated in a BBO crystal.
The high pulse energy of the visible pulses (200−350 nJ) together with the short focal length within the STM leads to high excitation densities that considerably alter the sample's properties and thus easily detectable signals. This opens the avenue for the investigation of the intriguing nanoscale physics occurring within tens of femtoseconds to picoseconds after optical excitation.

4.6 Optimised STM scan head for lightwave-driven scanning tunnelling microscopy

The new LW-STM setup utilising mid-infrared carrier waves was designed and built in close collaboration with the research group of Prof. Dr. Jascha Repp. All components for smooth STM operation were and are assembled by Dr. Thomas Buchner, Raffael Spachtholz, and Christoph Rohrer. For the sake of completeness, the most important considerations in the design and construction of the STM scan head are summarised here.
As the scan head constitutes the centrepiece of any STM, its design should be optimised for the exact desired experiment and measurement scheme. The standard considerations required for DC- and electronic AC-STM operated at low temperatures include sufficient thermal and vibrational decoupling from the environment, an

ample scan range for relative motion between tip and sample, biasing of tip and/or sample, as well as low-resistance wiring to the measurement device. The first two requirements are achieved by suspending the scan unit from springs and attaching additional eddy current dampers (see Fig. 4.30) to the bottom of the copper baseplate. By surrounding the scan head with two layers of cryogenically cooled shields (liquid nitrogen and liquid helium) as well as an ultrahigh vacuum of 10^{-10} mbar, the thermal exchange is limited to blackbody radiation of the 4.2 K cold surrounding.
Besides these basic prerequisites, the scan head employed for LW-driven STM necessitates additional adaptations. For example, an additional optical access to the STM tip must be accommodated. The radiation coupled in from outside the vacuum chambers through these windows must be guided towards and focused onto the tip-sample junction. In our first implementation of a lightwave-driven STM (see Fig. 2.5, sketch in the bottom part), the collimated terahertz beam is focused using an off-axis parabolic mirror with a focal length of 75 mm. A following planar mirror, which is tiltable via piezo driving motors, allows for the precise alignment of the beam to the tip-sample junction. Both these optical components are mounted directly on the scan head, resulting in a highly stable focus position even for vertical and horizontal motion of the scan head relative to the incident beam.
In the first-generation setup [Coc13], the STM tip is scanned across the sample with Ångström-precision using piezo stacks. Given the relatively large focus of the terahertz radiation, measuring few millimetres in diameter, the tip can be moved in a spot size of hundreds of micrometres without a noticeable reduction of the field strength. However, during standard STM operation, it is regularly necessary to change the tip position on the millimetre scale for various reasons, including the search for an optimal measurement spot or a tip cleaning procedure. In these cases, the laser alignment must be repeated, which substantially increases the time spent on the alignment procedure at the expense of measurement time. In the new MIR-STM, where the frequency of the incoupled radiation is higher by a factor of more than 20 relative to the first-generation THz-STM, the focal size will shrink accordingly, which theoretically amplifies the need for frequent realignment of the focal spot. For these considerations, the sample should ideally be scanned with a fixed tip position. This eliminates the need for realignment of the beam when changing the measurement spot, reducing undesired interruptions of the data acquisition.

Furthermore, for the characterisation and investigation of samples, it is also of interest to study radiation scattered or even emitted from the tip-sample junction. In many techniques related to STM, the scattered light is in fact the main information extracted in the measurement. For example, using tip-enhanced Raman spectroscopy (TERS) [vSL12, Zha16] or scanning near-field optical microscopy (SNOM) [Eis14, Hub16, Pla21], fascinating physical properties and dynamics can be investigated.

Aiming to harness this additional channel of information in our new MIR-STM, we designed the optics to collect the radiation scattered from and emitted by the tip-sample junction. To this end, a large curved mirror is placed in close proximity to the tip (see Fig. 4.30 and Fig. 4.31). One half of the parabolic mirror is used to focus the radiation onto the tip and the other to collect light scattered or emitted out of the junction, collimate it, and guide it out of the vacuum chamber for detection. The large mirror covering $\frac{1}{4}$ of the solid angle and of which one half is used to gather scattered light is expected to be able to collect sufficiently large intensities of scattered radiation for analysis. Having this channel of information on the spectral composition of the scattered and emitted radiation aside of the tunnelling current will greatly improve the capabilities of the new MIR-STM.

Another advantage of placing the mirrors in close proximity to the tip is the high numerical aperture and thus the smaller focal size. This increases the mid-infrared field strength and therefore the ultrafast bias applied across the junction, enabling the access to energy levels further away from the Fermi energy. Furthermore, restricting the light to a smaller area leads to a higher excitation density. For example, when pumping the sample with photons in the visible range whose energy exceeds the bandgap of a semiconductor, the excited electrons and holes or even their paired quasiparticles, excitons, can strongly influence the local conductivity of the sample. A higher density of excitons would lead to a clearer signal, reducing acquisition times substantially.

As a result of the changes to the geometry of tip and mirror, the new scan head needed to accommodate additional modifications of the ramp ring. The ramp ring holds the sample and allows for the relative motion between the tip and the sample surface (see Fig. 4.30). As first introduced by Besocke *et al.* [Bes87] in 1987, this movement can be achieved by placing a ring with three ramps on top of three piezo-electric

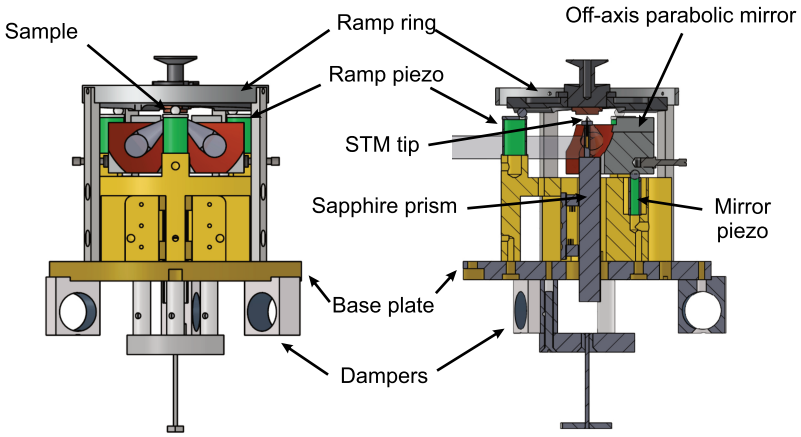


Figure 4.30 | Schematic front- and side-view of the new STM scan-head design. A copper base plate with three eddy current dampers forms the foundation for the scan head. The sample is screwed to the centre of the ramp ring that is placed on top of three ramp piezos (green cylinders). A large mirror covering $\frac{1}{4}$ of the solid angle allows for focusing the radiation onto the tip and recollimating light originating from the tip-sample junction. The mirror can be moved by means of three additional piezos (here only visible in the side-view on the right). The STM tip is mounted on top of a sapphire prism, which facilitates the vertical positioning of the tip.

driving motors. By applying suitable voltage pulses to different parts of the crystals, vertical elongations as well as lateral slants are possible, providing a full 3D freedom of motion of the ramp ring (see Fig. 4.31, ramp piezos). Additionally, the tip can be moved in the vertical direction using a sapphire prism in a so-called Pan design [Pan99]. In order to apply a bias voltage to the sample, which is necessary for STM operation, the ramp ring is conductively connected to the top of the piezo stacks using phosphor bronze half-spheres (Fig. 4.31, small spheres on top of the ramp piezos). Each contact can be separately addressed, which allows for the electrical gating of the sample in addition to biasing the sample relative to the tip. Modifying the doping of the sample by shifting the Fermi energy serves as an additional tuning knob in the measurements, which will facilitate the investigation of numerous fascinating and

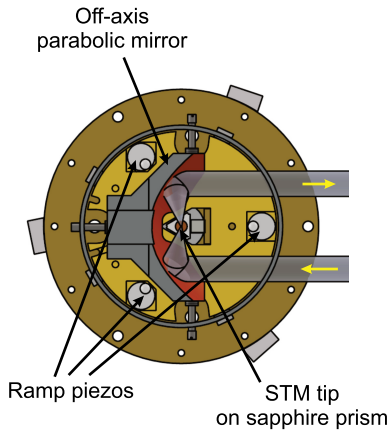


Figure 4.31 | Top-view of the new scan-head design without the ramp ring. The off-axis parabolic mirror covers a large solid angle and is used for the in- as well as outcoupling of radiation (grey areas with yellow arrows) into/from the tip-sample junction. The STM tip can be moved up and down (into and out of the image plane) using a sapphire prism.

highly relevant samples, including complex two-dimensional semiconductor devices. In summary, the novel modified STM scan head overcomes the drawbacks of the first-generation scan head by inverting the ramp-ring geometry such that the sample is scanned over a stationary tip. This necessitates only a single alignment procedure of the mid-infrared pulses to the tip prior to the measurements. Moreover, the large parabolic mirror mounted in close proximity to the tip allows for the generation of high field strengths in the focus as well as the recollimation of radiation scattered out of the tunnelling junction for further analysis. These improvements will allow for the study of ultrafast dynamics on the femtosecond scale within a plethora of appealing sample structures by investigating the pump-induced changes to the lightwave-driven tunnelling currents and the emitted light. For example, the experimental design alluded to in the end of section 2.7 could become a reality. In that context, the idea was to resolve the quantum state of a molecular harmonic oscillator by investigating the statistics of the absorbed and/or emitted photons as a function of the oscillation phase. The vibration can alter the absorption and emission properties by for example modulating the transition dipole. Thus, studying the light scattered out of/emitted from the STM junction spectrally, will yield a new information channel in MIR-STM. All in all, over the last four years, we set up a new laboratory in which we combine the incredible atomic spatial resolution of a custom-designed STM with the fem-

tosecond temporal resolution lightwave electronics have to offer. Starting with an empty laboratory (see Fig. 4.32a), we planned, designed, purchased, and aligned the complex optical and electronic components necessary to achieve such a novel ultrafast nanoscopy setup. Figure 4.32b displays the status of the laboratory in July 2023. The STM with its preparation chamber and cryostat can be seen in the foreground of the right panel. The left panel partially shows the Yb:fibre amplifier as well as the dual-stage Herriott cell and illustrates the density of optical components necessary to achieve the goal of ultrashort CEP-stable pulses in the mid-infrared range. The final optical setup and a sketch of the position of the STM relative to the optical table is provided in Fig. 4.33. The near-infrared output pulses of the Yb:fibre amplifier undergo spectral broadening and temporal compression in a Herriott-cell setup. The output pulses are split into multiple parts, each used for a different purpose in the final nanoscopy setup. With one arm, CEP-stable mid-infrared transients are generated, which will be used to drive tunnelling events in the MIR-STM within a predicted time window as short as 6 fs. The other arm is used to drive a dual-stage NOPA for the generation of tunable ultrashort visible pulses. These can be used to optically excite the sample under investigation. For details on the individual components, the reader is referred to the respective sections within this chapter.
In the near future, this setup will enable us to attain deeper insights into the dynamics of nanoscopic entities. For example, intramolecular vibrations, high-frequency phonons in crystals, and potentially even electronic motion can be investigated with the improved temporal resolution of this novel setup. Dynamics excited either by means of electron injection/extraction or absorption of a tunable visible pump pulse can be studied in real-time on the atomic scale. With these capabilities, we may obtain new insights in the excitation/relaxation pathways of charge carriers and vibrations as well as their non-trivial interaction. This may elucidate nanoscale variations in phase transitions and help us understand the origins of high-temperature superconductivity. Altogether, a wide range of interesting open questions in the field of condensed-matter physics may be addressed with the new MIR-STM.

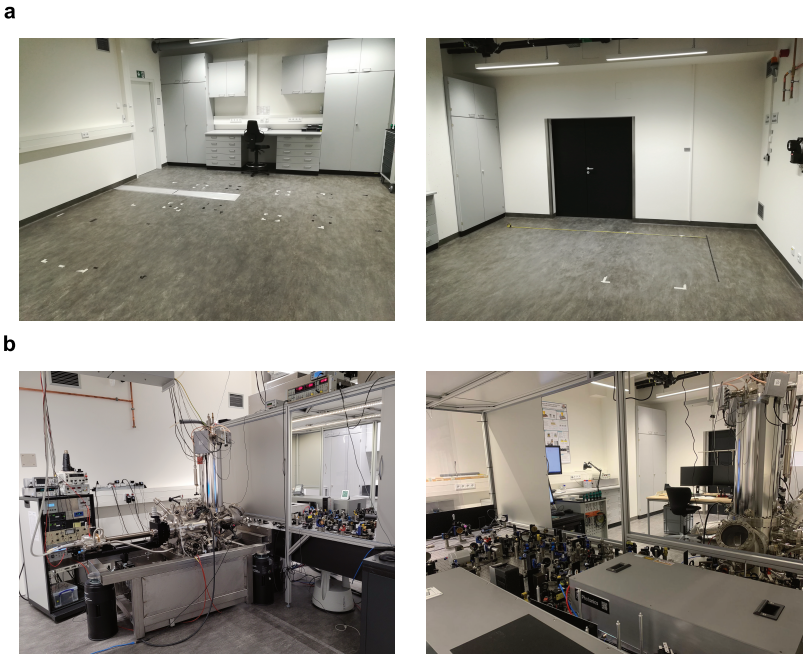


Figure 4.32 | Laboratory photographs. a, Photographs of the empty laboratory in June 2021 reveal the floor markers for the planned optical table and STM position. **b**, A picture taken in July 2023 shows the installed STM and optical setup. In the left panel, the STM is visible in the foreground (left half of the picture). In the right panel, the camera angle was chosen to capture more of the optical table (Yb:fibre amplifier and Herriott cells in the foreground).

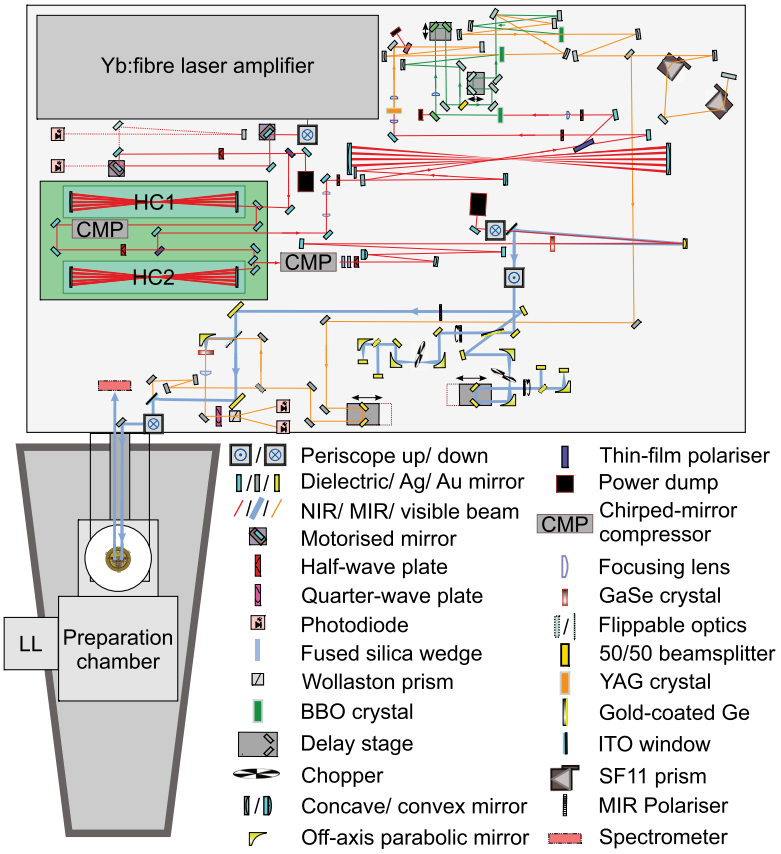


Figure 4.33 | Complete MIR-STM setup. The complete optical setup on the optical table (grey rectangle) and a sketch of the STM (grey trapezium, not to scale) are depicted. The three different chambers of the STM setup – the load lock (LL), the preparation chamber, and the actual STM measurement chamber – are outlined. The output pulses of the Yb:fibre amplifier are spectrally broadened and temporally compressed in Herriott cells (HC1, HC2) and employed to drive a multitude of nonlinear optical processes. The generated CEP-stable mid-infrared transients as well as tunable visible pump pulses from a NOPA are coupled into the vacuum chamber of the STM and focused into the tunnelling junction. Details on the individual steps are provided in the respective sections within this chapter.

Chapter 5

Summary and outlook

In this thesis, I presented our recent achievements in improving the capabilities of lightwave-driven scanning tunnelling microscopy (LW-STM). Building on the successes of the first-generation LW-STM system, providing a simultaneous Ångström and 100 fs spatio-temporal resolution [Coc16], I introduced two series of experiments investigating a single-molecule switch, magnesium phthalocyanine (MgPc), and a tungsten diselenide (WSe_2) monolayer.

The extension of LW-STM to single-molecule ultrafast action spectroscopy allows for the acquisition of data visualising the result of every attempt of switching MgPc [Pel20b]. We have detected intriguing position- and probe-strength-dependent variations in the switching probability. In the near future, chemical videography shedding light on elaborate reactions such as between multiple components [Rep06] or complex folding networks [Sti11] directly on surface and in real space could become a reality.

Utilising ultrafast action spectroscopy and an atom-scale voltage gauge, we are able to map out the full near-field waveform in an atom-scale gap [Pel21]. State-of-the-art nanophotonic simulations can now be assessed experimentally, providing the basis for an improved design of optical nanotechnology.

Applying local femtosecond forces to the MgPc molecule prior to the readout, we have successfully demonstrated coherent control of the molecule's motion on the atomic scale. Specifically, the electric field of a terahertz pulse was used to excite an in-plane librational motion, leading to coherent oscillations of the switching probability. Moreover, the interaction between key atoms of the molecule and the substrate

leads to strong local differences in the observed libration characteristics [Pel20b]. This motivates the investigation of chemical reactions and their response to precisely pinpointed excitation to enhance our understanding of catalysts in numerous processes.

By controlled excitation of the molecular oscillator MgPc, engineering of non-classical vibrational states may be possible. Investigating the switching statistics of the librating MgPc, we detected intriguing temporal variations of the switching variance. These modulations could indicate that a squeezed oscillator state is created. Extended statistical analysis and additional experimental evidence are necessary to verify these assumptions. Thus, we have introduced quantum-limited counting statistics, taking the first step towards imaging an individual molecule's quantum state at its intrinsic spatial and temporal scales.

Intricate motion patterns of atomic nuclei are not only present in molecular systems but are ubiquitous in all forms of matter at temperatures above absolute zero. Notably, the electronic and vibronic attributes of a system are highly interconnected. Investigating a monolayer of the semiconducting transition-metal dichalcogenide tungsten diselenide (WSe_2), we were able to show the interaction of a phonon mode with the electronic degrees of freedom. Specifically, the localised energy levels, arising from a single-atomic selenium vacancy in the WSe_2 monolayer, were studied. An acoustic vertical phonon mode excited by a terahertz pump pulse modulated the tunnelling current driven by a time-delayed probe pulse. Extending our capabilities within the LW-STM to allow for the first time-resolved scanning tunnelling spectra on sub-picosecond timescales, we revealed unipolar, non-monotonic shifts of the tunnelling onset of the lower defect energy level by up to 45 meV.

The local conductance of WSe_2 can not only be modified by terahertz field excitation, but I have also shown that in optical pump – THz probe measurements, changes to the tunnelling current can be detected. By comparing pump-probe traces on mono- and bilayer WSe_2 samples, we potentially detected signatures of intralayer exciton decay. Further experiments are expected to clarify the origins of the time-dependent tunnelling current and resolve variations in the charge-carrier dynamics with Ångström spatial resolution.

To overcome the limited resolution of visible pump – THz probe experiments and investigate even faster dynamics, we constructed an improved LW-STM facilitating

tunnelling currents with unprecedented temporal precision. To this end, the pulses generated in a near-infrared high-power laser amplifier were modified by harnessing numerous nonlinear optical processes: Particularly, we implemented a dual-stage Herriott-type multipass cell to temporally compress the pulse duration from 270 fs to 13.5 fs (FWHM) and used intrapulse difference-frequency generation within a gallium selenide crystal to produce inherently phase-stable ultrashort mid-infrared pulses with a broadband intensity spectrum spanning from a frequency of 10 THz to 38 THz (FWTM). These single-cycle pulses can ultimately be employed as ultrafast transient bias for the novel MIR-STM.

Moreover, to optically excite charge carriers within the sample, I designed a setup to create intense pulses in the visible range. Exploiting filamentation within a YAG crystal, a broadband supercontinuum is created. Variable spectral regions of its short-wavelength part (550−1000 nm) can be amplified by means of a dual-stage non-collinear optical parametric amplifier. This allows for tunable ultrashort pump pulses with pulse energies of up to 120 nJ that can be used to resonantly drive specific energy transitions up to photon energies of 2 eV within the sample.

With the combination of a tunable optical excitation and a temporal resolution of tunnelling below 10 fs, our capabilities to investigate ultrafast dynamics are greatly enhanced. In the future, this novel setup will enable the visualisation of intramolecular vibrations, high-frequency phonons, as well as their interaction with the electronic degrees of freedom [Mer21]. For example, the bound electrons in semiconductors can be optically excited, leading to a hot electron–hole plasma out of which a plethora of quasi-particles such as intra- and interlayer excitons [Ste17, Mer19a, Mer20] or trions [Mak13] can emerge. With the enhanced temporal resolution of the new MIR-STM, the decay and conversion between these different excited states and their nanoscale spatial variations can be tracked.

Many properties of quantum materials originate from the complex interplay of charge, orbital, spin, and lattice degrees of freedom. With LW-STM, local variations in topographic and electronic properties and their effect on electron–phonon interaction can be studied. Furthermore, we believe that the temporal resolution within this next-generation LW-STM will suffice to investigate phase correlations and possibly their statistical behaviour. For example, the ability to track phase coherence [Cor99] could yield insights into how phase fluctuations of preformed Cooper pairs contribute

to shaping the phase diagram of high-temperature superconductors [Kei15].
In addition, intramolecular vibrations greatly influence the light-harvesting capabilities as they play an important role in the efficiency of excitation energy transfer [Kun22], singlet fission [Dua20], and even photosynthesis [Ful14]. Furthermore, molecular vibrations are ubiquitous and govern the efficiency of chemical reactions. For a deeper understanding of chemical and biological processes, even the vibrations within the "simple" yet highly relevant molecules carbon dioxide [Joh95] or water [Per16] must be closely understood. Femtosecond dynamics of vibrations induced by light [Dan90] or by electron transfer [Raf21] that have so far only been investigated in ensembles can now be studied at the nanoscale.
I believe that the novel mid-infrared LW-STM with its time, energy, and spatial resolution is able to address these exciting questions and many more in order to improve our understanding of the ultrafast nanoscopic phenomena all around us.
To elevate the capabilities of LW-STM even further, one can consider the following potential enhancements. Combining LW-STM with a strong magnetic field, the investigation of completely new sample systems and the physical processes therein become possible. Previous studies have made use of the tunnelling magnetoresistance with spin-polarised STM tips and have shown the control over individual skyrmions [Rom13] and single-atom magnets [Nat17] as well as the exploration of spin relaxation dynamics on nanosecond timescales [Lot10].
To further explore the propagation of ultrafast excitations on the nanoscale, a possible avenue lies in the use of two STM tips. Launching phonons, plasmons, or other collective excitations from one position and monitoring the effect at a different tip position as a function of time would allow for the investigation of propagation velocities, interference, or dispersion. Hence, this could open another interesting perspective on the transfer of energy and momentum on nanoscopic scales.
To conclude, building on the groundbreaking insights facilitated by the exceptional spatio-temporal resolution of LW-STM, the next-generation MIR-STM setup will expand our understanding and control of the nanoworld by providing even better temporal resolution and more versatile excitation pathways.
Having commenced this thesis with the famous words of Goethe's Faust who pondered on the ultimate truth regarding the inner workings of this world, I would like to conclude my dissertation with an equally well-known quote by the same persona:

"Da steh' ich nun, ich armer Tor,
Und bin so klug als wie zuvor!"

"For here I stand, a fool once more,
No wiser than I was before!"
(loose translation)
Faust, Part I (1808), Johann Wolfgang von Goethe

Though at first glance these words convey frustration about scientific stagnation, they also contain the understanding that every question answered raises novel queries worth investigating. However, in contrast to Faust's purely pessimistic view, I wholeheartedly perceive that in my time as PhD student I have continuously learned and added to my experiences, satisfying my personal curiosity on the inner workings of the minuscule building blocks of matter around us. Additionally, I hope to have committed a tiny sliver of knowledge about our world that has not existed prior to my research. Perhaps one day, our everyday lives will even benefit from the developments made on the basis of the insights gained through my efforts.

Appendix A

Single-molecule oscillator

A.1 Spatial map of the switching probability

In order to investigate the spatial variations in switching behaviour, we performed experiments for various tip positions. The top row of Suppl. Fig. A.1 depicts the resulting maps of local switching probability for the two initial configurations: left (panel a) and right (panel b). In each panel, eight bright lobes of increased switching probability, of which every other is slightly more distinct, become apparent. Comparing these switching maps to the calculated maps of charging probability (panels c and d for molecule in the left and right configuration initially), reveals clear similarities. From this we can conclude that in first approximation the local charging probability dominates the switching probability.

A.2 Extended data on quantitative sampling of the near field

A.2.1 Far-to-near-field transfer function

From the amplitude and phase spectra for the far and near field displayed in Figs. 2.14c and 2.14d, respectively, a complex-valued far-to-near-field transfer function including all effects such as coupling, propagation, and quantum phenomena, can be obtained (Suppl. Fig. A.2). The experimental results for both amplitude (dark blue) and phase

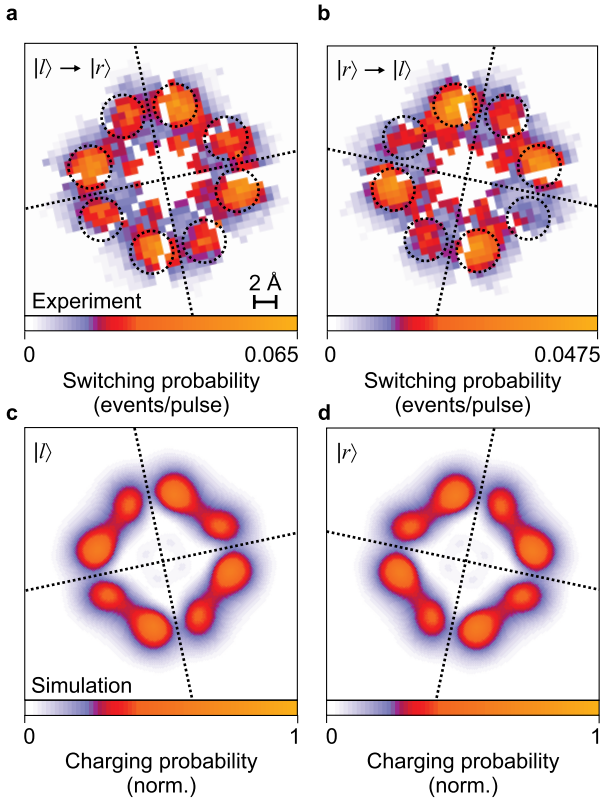


Figure A.1 | Spatial map of the local switching probability. a+b, The measured spatial profiles of switching probability (events/pulse) for toggling the molecule from left to right (panel a) and right to left (panel b) exhibit eight bright lobes each, where every other is slightly stronger in intensity. **c+d**, Comparing the measured switching probability to the simulated charging probability shows strong similarities. This indicates that the switching probability for a certain tip position dramatically depends on the local charging probability and thus wave-function overlap. The slight asymmetry in the lobe structure stems from an internal torsion the molecular frame.

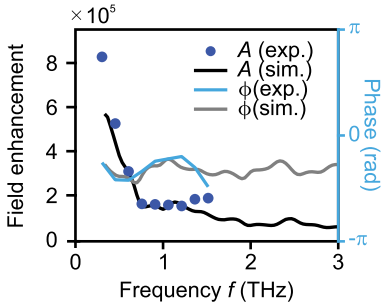


Figure A.2 | Complex-valued transfer function. By comparing the far- and near-field spectra, a complex-valued transfer function can be extracted. The propagation into the near field leads to a field enhancement of the amplitude A on the order of 1.8×10^5 and a phase shift of around $-\frac{\pi}{3}$ for a frequency of 1 THz.

(light blue) agree very well with the simulated transfer function (black and grey, respectively). Among the most notable frequency-dependent changes are the $\frac{1}{f}$-like field enhancement and additional finer structure in the phase as well as amplitude transfer functions. On the curves for both the amplitude and phase of the simulated transfer function, a small oscillatory structure with a periodicity of approximately 1.3 THz is observed. Even small features such as the flat part of the field enhancement around 1 THz and the oscillation in the phases concur between simulation and experiment.

The origin of the oscillatory feature becomes apparent upon close examination of the shape of the tip. Due to the electrochemical etching process, a sharp edge is created, at which surface plasmons can be reflected, forming a standing-wave pattern between this lip and the tip apex. For the tip employed in our experiments, this etched regions span approximately a length of $\mathrm{L} \approx 200\,\mu\mathrm{m}$, acting as a resonator for surface plasmons whose wavelength λ fulfils $\mathrm{L} = \mathrm{n} \cdot \lambda/2$, with n being an integer. This leads to a slight frequency-periodic modulation of the coupling efficiency of lightwaves into the junction.

A.2.2 Simulation: Retardation effects in the near field

Adding the molecule into the junction strongly affects the temporal evolution of the electric field at the centre of the simulated tunnelling junction (see Suppl. Fig. A.3). The first part of the transient is not modified and only drives electrons from the tip

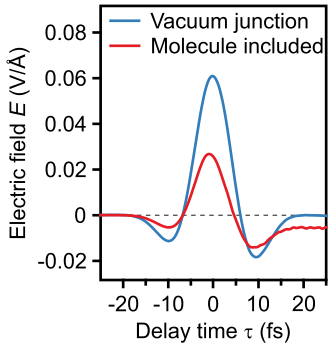


Figure A.3 | Retardation effects. The presence of the molecule alters the temporal evolution of the near-field waveform. Compared to the vacuum junction (blue curve), the main maximum of the waveform is suppressed when the molecule is present in the junction (red curve). Additionally, retardation effects occur, leading to a long trailing field longer than 25 fs, which is not observed in the vacuum junction.

into the molecule. Due to accumulation of negative charges in the molecule and the hole left in the tip, an electric field with opposite polarity to the driving field builds and reduces the peak field. The lifetime of the charges features a trailing field, which persists for a long time after the electric field of the pulse has shrunk back to zero.

A.3 Field-induced in-plane rotation

Pump-probe experiments at different positions on the MgPc molecule (marked A-E in Suppl. Fig. A.4b) reveal interesting spatial variations (see Suppl. Fig. A.4a). For positions A to C, a large mean switching probability but almost no temporal variations are detected. In contrast, at the spots marked D-E, the mean switching yield is lower, while the oscillation amplitude is large. Comparison with density-functional theory calculations exposes a possible explanation. The colour map in Suppl. Fig. A.4b encodes the local in-plane rotation as a response to an electric field applied locally. The region where datasets D-F were recorded corresponds to an area where a large angle rotation of 1° is caused by the electric field. At positions A-C, almost no in-plane rotation is caused. The underlying mechanism for these disparities can be understood when examining the adsorption energetics of the molecule on top of the sodium-chloride lattice. The negatively charged chlorine ions of the salt interact strongly with select atoms of the molecular frame. Particularly, the outer hydrogen atoms of the isoindole units (marked in pink) undergo a strong polar interaction,

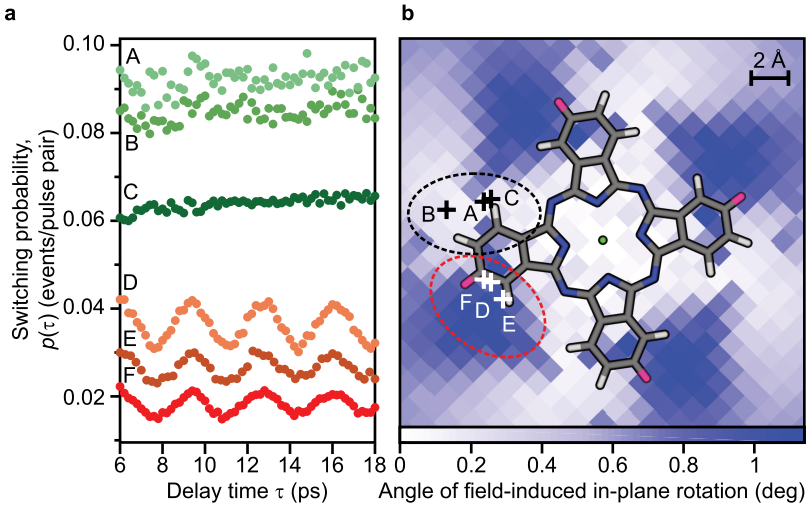


Figure A.4 | Local field-induced angular excursion. a, Measuring the position- and time-dependent switching probability for 6 tip positions (A-F) shows stark differences. The data recorded at A-C possess a high mean switching probability over the measured time window ($6 - 18$ ps), while no clear oscillations are detected. The inverse behaviour is seen at locations D-F, where the mean probability is low but strong time-dependent variations persist. **b**, Comparing the relative positions of A-F on the molecular frame with DFT calculations of field-induced rotations reveals the positional coincidence of large switching-probability oscillations and a large field-induced in-plane rotation of the molecule.

locally binding the molecule tightly to the surface. Disruption of this interaction with an external electric field thus strongly modifies the adsorption energetic of MgPc, leading to an in-plane rotation. After the ultrafast electric field pulse in our experiment has passed, a librational motion about the potential minimum is triggered. This strongly alters the delay-time-dependent switching probability. The origin of the local discrepancies becomes clear upon revisiting the analysis of the unpumped molecule. The four isoindole units exhibit a small internal torsion due to the same strong polar interaction between specific hydrogen atoms with chlorine ions of the substrate (see Suppl. Fig. A.5).

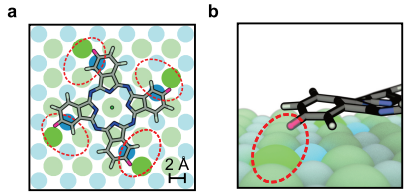


Figure A.5 | Internal torsion of MgPc. **a**, Top view of the molecule in its left configuration on a NaCl lattice. The strong interaction of select hydrogen atoms (marked in pink) of the isoindole units with the negatively charged chlorine ions dominates the adsorption energetics of the molecule. The areas of strong interaction are marked with red dashed ellipses. **b**, As a result of the strong interaction, the molecule is not completely flat. The side view reveals an internal torsion of the isoindole arms.

A.4 Extended data on non-classical oscillator states

A.4.1 Bootstrapping

Boostrapping is a statistical method used to infer information on a population based on a set of samples taken [Efr82]. The basic idea of bootstrapping is explained on a straightforward example. Assume a large population of entities with distinct properties: balls of different sizes. Each random sampling with replacement leads to a detected distribution of sizes. From a finite number of draws, the true mean and variability cannot be exactly determined. However, resampling the recorded data allows for inferences to be made on the variability of the mean. Importantly, this statistical resampling approach is purely data-driven and can help us deduce estimations on fluctuations of the data without making any assumptions on the underlying statistics of the population.

In the case of the molecule's switching probability (see Chapter 2.7), to estimate $\sigma_{\langle M \rangle}$, we employ this bootstrap resampling method. In this process, we draw L = 4000 values randomly from the observed dataset of M, allowing for the possibility of selecting the same value more than once. This resampling procedure is repeated 87 500 times. For each resampled dataset (indexed as i), we compute the mean value, m_i. The standard deviation of this resulting distribution of m_i provides the bootstrap-

derived estimate for $\sigma_{\langle M \rangle}$.

In our work on single-defect spectroscopy (see Chapter 3), the bootstrap method was utilised to improve statistics and estimate error bars on the detected lightwave-driven current I_{LW} and its derivative $\frac{\mathrm{d}I_{\mathrm{LW}}}{\mathrm{d}E_{\mathrm{peak}}}$. For each set-specific external parameters (delay time, pump and probe fields), the measurement and analysis procedure was the same. The lightwave-driven current was recorded over a certain time period and subsequently the average value was calculated. The mean values of N measurement traces were subject to random resampling with replacement $M \times N$ times. This generated M sequences, each containing N elements. Following this, the mean value $\langle I_{\mathrm{LW}} \rangle_m$ ($m \in M$), for each sequence, was computed. Subsequently, the error bar of I_{LW} was determined as the standard deviation of $\langle I_{\mathrm{LW}} \rangle_m$. Likewise, the error bar of the derivative was derived as the standard deviation of $\langle \frac{\mathrm{d}I_{\mathrm{LW}}}{\mathrm{d}E_{\mathrm{peak}}} \rangle_m$.

Appendix B

Single-atom vacancy in monolayer WSe_2

B.1 Ultrafast tunnelling model

The tunnelling process from the metallic tip into a position on the semiconducting WSe_2 monolayer with a single-atomic vacancy was simulated. For our investigations we were interested in the tunnelling currents driven up to voltages of 500 mV (between the first and second defect levels D_1 and D_2). To do so, two contributions needed to be incorporated: the tunnelling through the lower-lying defect level D_1 and the "background" non-resonant co-tunnelling into the pristine WSe_2 monolayer on gold. As the electrons tunnelling through D_1 possess a non-zero dwell time within D_1, which effectively blocks this tunnelling channel for subsequent electrons, this had to be considered in the model. A rate equation is able to capture such dynamics in the following way:

$$\frac{dN}{dt} = -\frac{N}{\tau_{LT}} + \alpha \cdot (1 - N) \tag{B.1}$$

As per equation B.1, the temporal evolution of the occupation number N of the first defect level depends on time t, and the lifetime of the charged D_1-state τ_{LT}. The parameter α encodes the instantaneous tunnelling rate into the defect for a dwell-time of zero.

The uncertainty principle relates the lifetime of an excited state (here D_1) with its energy uncertainty. Therefore, by closely examining the width of D_1 in the DC-STS, the lifetime can be inferred. Fitting the D_1-level recorded in conventional DC-STS

using a Lorentzian line shape $L_{D_1}(V)$ yields a lifetime broadening of $\tau_{LT} = 5\,\text{fs}$.
To be able to compare the simulation to measurement data, we needed to simulate the autocorrelation trace we recorded. In the experiment, two weak terahertz transients with peak voltages in the nearfield below the first defect level were superimposed and temporally delay with respect to another. The total current was maintained slightly below one electron per pulse, which corresponds to around 100 fA for a laser repetition rate of 610 kHz. In combination with the short lifetime of 5 fs, the low currents ensured that artefacts arising from Coulomb-blockade effects were negligible. In the simulation, this was included by setting the number of electrons tunnelling through D_1 to one electron per 300 fs. This tunnelling rate of one electron per 300 fs corresponds to tunnelling through the whole defect level, which means:

$$\int_{-\infty}^{\infty} \mathrm{d}V L_{D_1}(V) = \frac{\mathrm{e}}{300\,\text{fs}} \tag{B.2}$$

Lower tip-confined voltages will induce lower tunnelling rates. Depending on the instantaneous tip-confined voltage, the tunnelling probability into the defect level is calculated as a fraction in the following equation:

$$I_{D_1,\,\text{inst}}(V_{\text{inst}}) = \frac{\int_0^{V_{\text{inst}}} \mathrm{d}V L_{D_1}(V)}{\int_{-\infty}^{\infty} \mathrm{d}V L_{D_1}(V)} \cdot \frac{\mathrm{e}}{300\,\text{fs}} \tag{B.3}$$

This allowed us to determine the absolute scaling of α with the instantaneous voltage V_{inst}. Notably, when as the tunnelling rate α varies with instantaneous voltage it becomes a function of t and autocorrelation time τ. By solving the rate equation numerically for each autocorrelation delay time τ with the known near-field voltage transient, the instantaneous tunnelling current $I_{D_1,\,\text{inst}}(t, \tau)$ was extracted.
As mentioned above, apart from state-selective tunnelling into the defect level, electrons with energies lower than the conduction band minimum can tunnel into the non-zero density of states of the gold-hybridised monolayer WSe_2 – the background. The STS spectrum of the pure background was modelled by subtracting the two Lorentzian curves, describing the two defect levels from the measured DC $\mathrm{d}I/\mathrm{d}V$ and fitting the remainder with a 5^{th} order polynomial $\text{Poly}_5(V)$. The contribution $I_{\text{background}}(t, \tau)$ of the background to the total current was calculated by integrating

the polynomial fit from 0 to $V_{\mathrm{inst}}(t)$ and the same scaling.

$$I_{\mathrm{background}}(V_{\mathrm{inst}}) = \frac{\int_0^{V_{\mathrm{inst}}} \mathrm{d}V \mathrm{Poly}_5(V)}{\int_{-\infty}^{\infty} \mathrm{d}V L_{\mathrm{D}_1}(V)} \cdot \frac{\mathrm{e}}{300\,\mathrm{fs}} \tag{B.4}$$

Importantly, for modelling the tunnelling into the background, no rate-equation model is necessary, as charges tunnelling into the monolayer or the gold substrate immediately propagate away due to their high mobility. To obtain the total LW-driven current for each interpulse delay time τ, the sum $I_{\mathrm{defect}} + I_{\mathrm{background}}$ was integrated over time.

B.2 Analysis of femtosecond snapshots

Within the main text, lightwave-driven snapshots of the single-atomic vacancy within a WSe_2 monolayer were recorded (see Fig. 3.12). Further analysis of the images allows for more elaborate statements on the inherent symmetries. Supplementary Figure B.1 shows the mean signal within a circular region of radius $r = 6.7\,\text{Å}$ centred on the vacancy indicated as a white circle in Fig. 3.12d. The radial averaging is represented by the red curve, while the blue curve corresponds to the azimuthal averaging. With increasing peak field strength, several different trends can be identified. The radially averaged signal first shows a clear peak at around $2 - 3\,\text{Å}$ for low field strengths, while above $0.4\,E_0$, an increase of the background signal (distance $> 5\,\text{Å}$) becomes evident, hinting towards the conduction band onset and thus tunnelling channels opening into the pristine semiconductor. In azimuthal direction below the defect level onset, no peaks can be discerned. The angle of 0° corresponds to the white line drawn in Fig. 3.12d, moving counter-clockwise with increasing angle. Once the peak field exceeds the defect levels ($0.24\,V$), three peaks with 120° separation arise, reproducing the threefold symmetry of the image. Above the conduction band, the C_3-symmetry disappears, showing only a depression at around 180°. By carefully comparing the steady-state and ultrafast STM images, we can thus confidently assign the first (second) peak in the $\mathrm{d}I_{\mathrm{LW}}/\mathrm{d}E_{\mathrm{LW}}^{\mathrm{peak}}$ spectrum to the D1 (D2) level associated with the selenium vacancy.

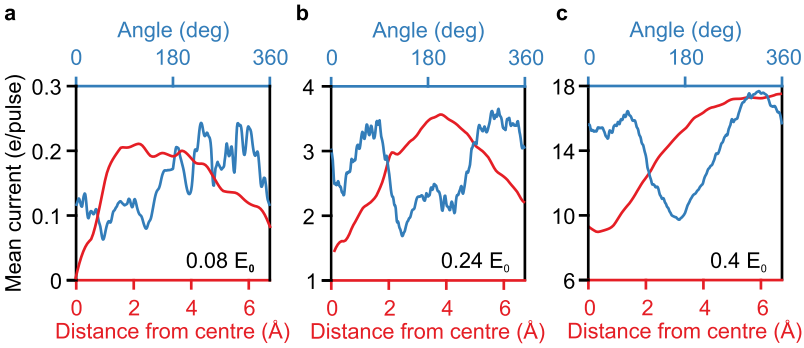


Figure B.1 | Analysis of annularity and background signal. Averaging the LW-STM images displayed in the bottom rows of panels a-c in Fig. 3.12 radially and azimuthally reveals their dominant appearance. Integrating radially from the centre of the selenium vacancy to a distance of 6.7 Å (exemplarily shown as a white circle in the bottom row of Fig. 3.12a) and azimuthally over the full 360° (0° marked by white line, moving counter-clockwise) accentuates the structure of the orbitals for the three energies used. **a**, For a peak field of $E_{\mathrm{LW}}^{\mathrm{peak}} = 0.08\,E_0$, no azimuthal dependence (blue curve) and a wide peak in the radial dependence (red curve) at $2-3$ Å is detected, indicating an annular shape. **b**, Above the defect energy levels ($E_{\mathrm{LW}}^{\mathrm{peak}} = 0.24\,E_0$), the plateau-like peak in radial direction becomes sharper. In azimuthal direction, three peaks with a spacing of 120° appear. The peak at 180° is attenuated in strength most likely by another defect nearby. **c**, Within the conduction band ($E_{\mathrm{LW}}^{\mathrm{peak}} = 0.4\,E_0$), the threefold symmetry blurs out and the radial signal increases towards larger radii from a minimum at the vacancy location. The remaining azimuthal asymmetry can be again attributed to another disruption of the crystal periodicity nearby.

B.3 Tunnelling decay constant

The tunnelling current possesses a characteristic dependence on the tip-sample distance. Supplementary Figure B.2 shows the retraction curves ($I_{\mathrm{LW}}^{\mathrm{pr}}$ vs. z) for two different delay times, $\tau = 8.8$ ps (blue) and $\tau = 10.2$ ps (orange), between the pump ($E_{\mathrm{LW}}^{\mathrm{peak}} = 0.32\,\mathrm{E}_0$) and the probe pulse ($E_{\mathrm{LW}}^{\mathrm{peak}} = 0.3\,\mathrm{E}_0$) plotted in a semi-logarithmic scale. As expected, the current scales exponentially with tip-sample distance $I_{\mathrm{LW}}^{\mathrm{pr}} \propto \exp(-2\kappa\Delta\mathrm{z})$. Linear fits to the curves yield the decay constants of $\kappa = 1.51\,\mathrm{Å}^{-1}$ and $1.59\,\mathrm{Å}^{-1}$ for the two delay times, respectively. The tip-sample

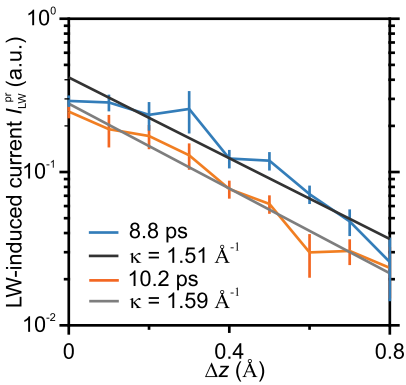


Figure B.2 | Retraction curves for two different delay times. When recording the LW-driven current as a function of tip-sample separation from a predefined height setpoint ($V = 30\,\mathrm{mV}$, $I = 49\,\mathrm{pA}$, 1.8 Å approach) for two different delay times, 8.8 ps (blue) and 10.2 ps (orange), the expected exponential dependence on tunnelling-barrier thickness, $I \propto \exp(-2\kappa\Delta z)$, is seen. Fitting linear curves to the semi-logarithmically plotted data allows for the retrieval of the tunnelling decay constants, $\kappa = 1.51\,\mathrm{Å}^{-1}$ and $\kappa = 1.59\,\mathrm{Å}^{-1}$, respectively.

distance of $\Delta z = 0$ corresponds to a height setpoint of $V = 30\,\mathrm{mV}$, $I = 49\,\mathrm{pA}$ and an additional approach by 1.8 Å.

B.4 Long-lived defect oscillations

In order to investigate the dynamics a pump pulse with a peak voltage above the conduction band onset can excite within the monolayer WSe_2, pump-probe experiments were conducted. These showed a distinct oscillation frequency of about 0.33 THz (see Fig. 3.6 on and off the lobes of a single-atom defect in a monolayer WSe_2. Extending the relative delay between excitation and read-out pulse to more than 95 ps for a measurement position on the lobe (position 2 in Fig. 3.6c) reveals the long-lived oscillation excited by the pump pulse. Over the measured region, no clear decay was conceivable, pointing towards a weakly-damped phonon mode – in our case the drum-like vertical oscillation mode of the whole monolayer.

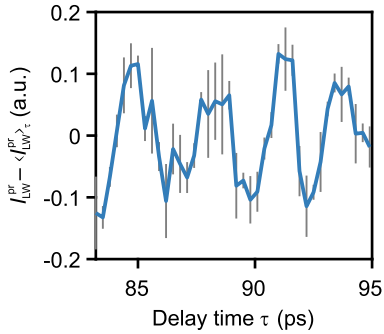


Figure B.3 | Current oscillations at long delay times. Monitoring the tunnelling current driven by the probe pulse for pump-probe delay times between 83 ps and 95 ps shows the long lifetime and thus weak damping of the current oscillations. The mean current $\langle I^{pr}_{LW} \rangle_\tau$ was subtracted for better visualisation. The error bars (grey) indicate the standard deviation of the bootstrapped current signal.

B.5 DFT calculations of ultrafast energy shifts of localised defect levels

It is crucial to thoroughly investigate the impact of atomic displacements resulting from DM on the energy spectrum at the defect site. The centre-of-mass oscillation, which is the most prominent motion, causes periodic changes in the monolayer's position within the gap by Δz_{Se}. As a result, the Coulomb interaction of the defect with its image charge in the tip and substrate is also affected.
Supplementary Figure B.4 displays the nomenclature for the calculation of the energy shifts. Upon negative displacement ($z_{Se} < 0$), the distance between the defect and its image charge in the tip is increased from $2z_1$ to $2(z_1 - z_{Se})$. Correspondingly, the distance to the image charge in the substrate is reduced from $2z_2$ to $2(z_2 + z_{Se})$. As as result, the energy shift $\Delta E_{IC}(\Delta z_{Se})$ due to this effect can be simply written as a sum of two Coulomb terms:

$$\Delta E_{IC}(\Delta z_{Se}) = \Delta E_{IC,tip} + \Delta E_{IC,\ substrate} \tag{B.5}$$

$$= -\frac{e^2}{2}\Big(\frac{1}{2(z_1 - \Delta z_{Se})} - \frac{1}{2z_1}\Big) - \frac{e^2}{2}\Big(\frac{1}{2(z_2 + \Delta z_{Se})} - \frac{1}{2z_2}\Big) \tag{B.6}$$

where $z_1 = 4\,Å$ and $z_2 = 5.5\,Å$ are the equilibrium distances from the defect to image plane of tip and substrate, respectively. With the peak-to-peak displacement

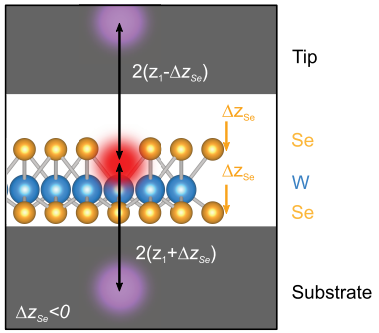


Figure B.4 | Image-charge effect and intracell distortion. To investigate the shift of energy levels due to a displacement of the different layers, two contributions were taken into account. Moving the top selenium layer out of its equilibrium position (z_1) by $\Delta z_{\mathrm{Se}} < 0$ leads to image charges in both the metallic tip and substrate at distances $2(z_1 - \Delta z_{\mathrm{Se}})$ and $2(z_1 + \Delta z_{\mathrm{Se}})$, respectively. Concurrently, an intracell distortion occurs, leading to the tungsten layer to settle closer to the bottom than the top selenium layer.

of 10 pm extracted from measurements (see Fig. 3.8b), this leads to:

$$\Delta E_{\mathrm{IC}}(\Delta z_{\mathrm{Se}} = +5\,\mathrm{pm}) - \Delta E_{\mathrm{IC}}(\Delta z_{\mathrm{Se}} = -5\,\mathrm{pm}) = 10\,\mathrm{meV}. \tag{B.7}$$

Figure B.5 visualises both the energy shifts due to the image-charge effect (red data points, bottom x-axis) and due to intracell distortions (green data points, top x-axis). The energies are given relative to the valence band maximum E_{VBM} of WSe_2 for tip-sample distances $-10\,\mathrm{pm} < \Delta z < 10\,\mathrm{pm}$. The illustration shows the corresponding intracell distances $-5.4\,\mathrm{pm} < \Delta z_{\mathrm{W\text{-}Se}} < 5.3\,\mathrm{pm}$ according to the drum-mode as well as the combined effect (blue data points). In total, the energy shifts lie in the same order of magnitude as the energy shifts we have seen in ultrafast STS. This indicates that indeed image-charge effects and intracell distortions could be the origin of the effects we observe in the experiment. Nonetheless, further contributions are possible.

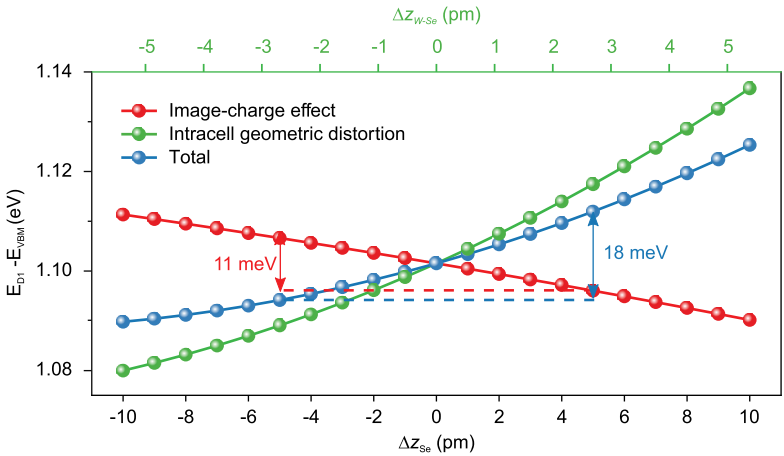


Figure B.5 | Effect of defect energy level on atom displacement. The D_1 energetic position relative to the valence band maximum shifts down monotonously with increasing Δz_{Se} as a result of the image-charge effect (red), while the intracell distortion (green) leads to a competing drift. The cumulative effect (blue) amounts to an energy shift of +18 meV for the peak-to-peak displacement of 10 pm extracted from the experiment.

B.6 Characterisation of the ultrafast voltage transient

Injecting terahertz waveforms into the STM junction leads to ultrashort bias pulses, which under the correct circumstances allow for tunnelling. Importantly, the measured far-field waveform does not reflect the actual transient voltage in the atomic junction, as different effects modify the intensity and phase spectrum. Using a molecular switch as atom-sized voltage gauge, the atomically-confined NFWF could be quantitatively extracted (see sec. 2.5). This very precise measurement scheme gives information on the exact shape of the bias voltage with extreme precision. However, this measurement technique is relatively complex, as it relies on the growth of an insulating layer and deposition of the molecular switch. Hence, it is valuable to utilise a more broadly accessible and versatile protocol to measure the NFWF within the tunnelling gap.

Notably, in this approach the atomically-confined waveform cannot be recorded but a voltage pulse within a larger tip-sample junction is extracted.
The basic idea is to utilise a visible pump – terahertz probe scheme where the visible gate pulse leads to a temporally strongly confined change of the junction properties, which is probed by the terahertz pulse as described by Yoshida *et al.* [Yos19] and Müller *et al.* [Mül20]. Supplementary Figure B.6a shows the dependence of the photocurrent on the DC bias for gate pulses with an average power of 960 µW at 610 kHz. As the calculated focal spot is in the range of tens of microns and thus considerably larger than tip-sample distance set in these experiments, both tip and sample are illuminated and photoelectrons are excited. For positive biasing, the photocurrent is dominated by photoelectrons flowing from the tip to the sample and for negative biasing vice versa. As the employed visible pulses have frequencies in the hundreds-of-terahertz-range, the Keldysh parameter (see eq. 2.8) is expected to be above 10, indicating dominating multiphoton emission. For low tip-sample distances and large bias voltages, an additional current contribution through photo-assisted tunnelling must be considered [Yan11]. Illumination of the junction with a terahertz pulse in conjunction with the DC bias will lead to a changed photocurrent signal [Coc13, Wim17] due to the additional transient biasing. As the terahertz pulse duration is much longer than the photoemission process, which is at the timescale of the near-infrared pulse envelope (<50 fs), photoemission can be estimated to be instantaneous. In addition, it can be assumed that the terahertz-induced change in photocurrent is be determined by the slope of the static I_{ph}–V characteristics (see Suppl. Fig. B.6a). If the I_{ph}–V curve is linear, the instantaneous terahertz voltage $V_{\mathrm{LW}}^{\mathrm{inst}}$ can be directly obtained by dividing the measured photocurrent by the slope. Varying the relative arrival time of the near-infrared and the terahertz pulse will then directly lead to a calibrated near-field voltage transient (see Suppl. Fig. B.6b). To test the effect of the applied DC bias on the delay-time-dependent photocurrent, three curves were recorded at $V_{\mathrm{DC}} = 0\,V$, $6\,V$, $10\,V$ (indicated by dashed lines in Suppl. Fig. B.6a). Clearly, in the linear regime ($6\,V$ and $10\,V$) the two curves agree very well (Suppl. Fig. B.6b, blue and purple transients). Setting a DC bias of $0\,V$ leads to the enhancement of negative bias features of the waveform, as a direct result of the increased slope in the I_{ph}–V_{DC} characteristic. Thus, by guaranteeing that the terahertz voltages remain within the linear region of the I_{ph}–V_{DC} curve

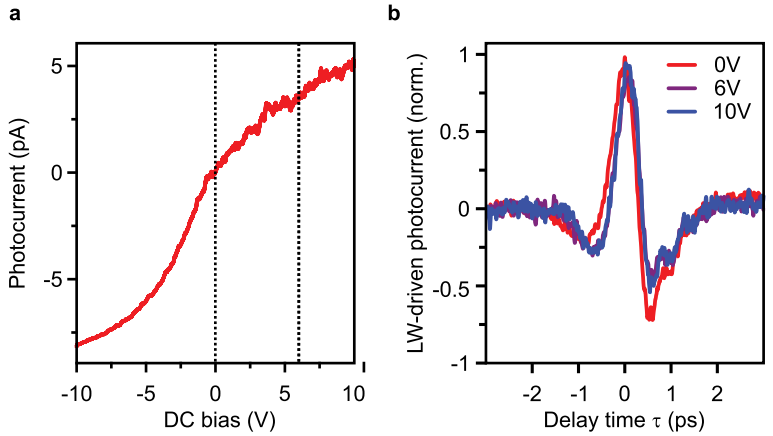


Figure B.6 | Voltage transients in STM junction. a, The relation between the laser-induced photocurrent and the applied DC voltage is strongly nonlinear for negative bias voltages, while it is linear for positive bias voltages. **b**, The photocurrent as a function of pump-probe delay for different DC bias voltages ($0\,V$, $6\,V$, and $10\,V$) was recorded. The nonlinear I_{ph}–V characteristic (panel a) is reflected in a change of the shape of the transient current.

and by making use of the quasi-instantaneous photoemission process, the near-field waveform within the tip-sample gap of 100 nm can be determined in a phase-resolved manner without the need of a specific sample system.

B.7 Bias-modulation for quicker spectroscopy measurements

As explained in section 3.4, in order to obtain ultrafast snapshots of the local density of states, we monitor the probe-pulse-driven tunnelling current for different peak fields $E_{\mathrm{LW}}^{\mathrm{peak}}$. By differentiating the resulting current I_{LW} with respect to the terahertz field strength ($\mathrm{d}I_{\mathrm{LW}}/\mathrm{d}E_{\mathrm{LW}}^{\mathrm{peak}}$), we obtain access to the local density of states. When conducting DC-STS measurements, this derivative $\mathrm{d}I_{\mathrm{DC}}/\mathrm{d}V_{\mathrm{DC}}$ is immediately

recorded while slowly varying V_{DC} by modulating the bias with a sine signal of several hundred hertz and peak-to-peak amplitudes of around 20 mV. By demodulating the measured current curve $I(V)$ in a lock-in amplifier, the slope can be immediately extracted. In LW-STS, a similar approach would significantly improve the signal-to-noise ratio and reduce the measurement time. Therefore, a modulation of the terahertz field strength in the range of 10-100 Hz would be desirable. I will introduce the possibility of such a rapid field modulation without the need for additional optical components that could alter the shape of the waveform.

To achieve this, we modulate the near-infrared polarisation prior to the terahertz generation in the lithium niobate crystal. Since the nonlinear process of optical rectification is strongly polarisation-dependent, the field strength of the generated terahertz pulses will change accordingly. To extract the exact scaling of the terahertz field strength with the polarisation of the incident near-infrared pulse and rule out any changes to the waveform, we electro-optically sample the terahertz transients for different waveplate rotation angles. The waveform remains practically unchanged within the range of $\pm 25°$ (Suppl. Fig. B.8a). For larger rotations of the near-infrared polarisation, changes to the waveform become apparent. The peak signal in EOS, which directly encodes the peak field of the terahertz transient, shows a smooth dependence on the near-infrared polarisation angle (Suppl. Fig. B.8a). Supplementary Figure B.7 reveals that the asymmetry (panel a) follows a smooth cosine^2-dependence on the polarisation angle, whereas the peak position and peak width possess some jumps as a function of angle. In summary, small angle modulations ($\pm 5°$) allow for the modulation of the terahertz power while keeping the other parameters of the waveform intact.

To achieve field variations of a few percent, only small angle changes of $\pm 5°$ are necessary. In combination with a motorised mount allowing for rotation frequencies of 60 Hz, this facilitates the fast modulation of the terahertz peak voltage in the junction and the immediate extraction of the instantaneous local density of states with LW-STS.

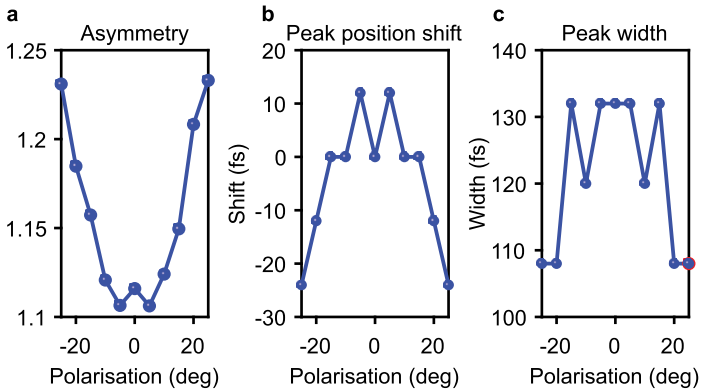


Figure B.7 | Dependence of terahertz waveform on near-infrared polarisation. Variation of the near-infrared polarisation incident on the nonlinear generation crystal $LiNbO_3$ leads to changes in various parameters of the terahertz transients. **a**, The asymmetry ranges from 1.1 for optimal generation efficiency to 1.23 for angles of ±25°. **b**, The maximum field of the waveforms shifts in time from +12 fs to -24 fs with respect to the 0° case. **c**, The peak width (top 90%) remains relatively stable with a mean of 121 fs and a standard deviation of 11.3 fs. For small angle variations of ±5° the peak width is constant at 132 fs.

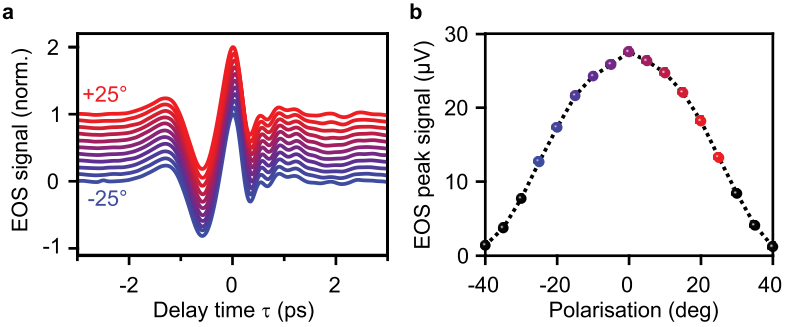


Figure B.8 | Terahertz electric field and waveform vs. near-infrared polarisation. a, The normalised EOS traces of the terahertz waveform for various near-infrared polarisations between −25° and 25° look practically identical, showing no angle-dependence. Subsequent traces are offset by 0.1 for clarity. **b**, The field strength of the generated terahertz transients, however, critically depends on the polarisation of the near-infrared pulses at the nonlinear crystal. The peak signal of terahertz transients acquired using electro-optic sampling are plotted as a function of the near-infrared polarisation. The spheres are coloured in the same scheme as the waveforms in panel a. Black spheres indicate the peak fields of transients not shown in panel a.

Appendix C

Ultrashort laser pulses for MIR-STM

C.1 Extended data on the amplifier characterisation

Extensive characterisation was performed on the commercially acquired high-power amplifier to ensure that the beam quality in the spatial as well as temporal domain was suitable for our application.

C.1.1 Spatial homogeneity

The laser output spectrum can vary across the beam profile, exhibiting what is known as spatial inhomogeneity or spatial chirp. This spatial variation can lead to less effective nonlinear optical processes, which necessitates its minimisation.
An analysis of the intensity spectrum highlights significant differences in the spectral composition across the diameter of the beam. Supplementary Figures C.1a and b clearly display the marked distinctions in spectra obtained from various positions within the beam area. To illustrate this, the spectra have been normalised to their respective maxima, revealing a gradual transition from a greater spectral emphasis on the long-wavelength end at the bottom of the beam profile to stronger short-wavelength components at the top (see Suppl. Fig. C.1a). Similarly, though to a lesser extent, the short-wavelength spectral content increases from left to right in the horizontal direction (see Suppl. Fig. C.1b).

Upon investigating the issue, it was determined that the problem originated in the grating compressor within the laser housing after the CPA. Rectifying the alignment of the grating and the end mirror significantly reduced the inhomogeneity. Panels c and d of Suppl. Fig. C.1 present spectra obtained from positions comparable to those in panels a and b. The spatial variation is now hardly discernible in the vertical direction, and only subtle alterations in the horizontal direction can be detected.

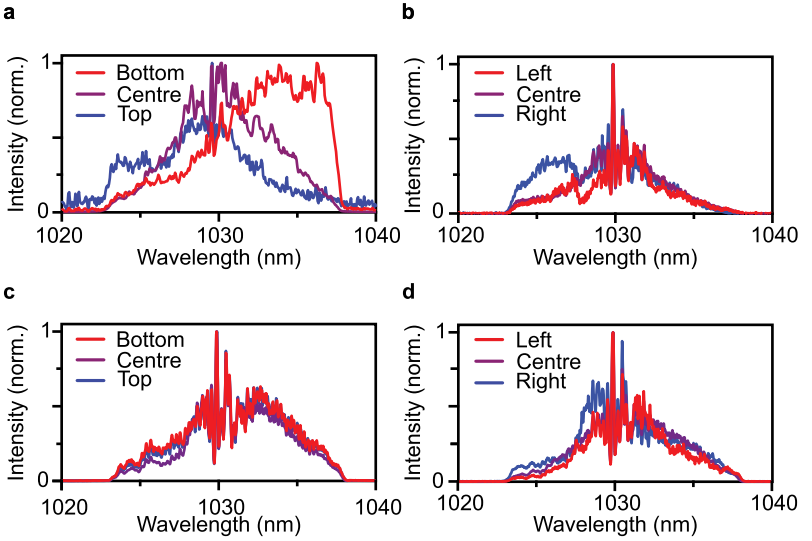


Figure C.1 | Spatial inhomogeneity within the output pulses of the Yb:fibre amplifier. The spatial variations of the spectra across the mode profile of the laser output pulses are determined. **a**+**b**, The normalised spectra differ dramatically between the top and bottom as well as the left and right positions, indicating a considerable spatial chirp. **c**+**d**, After the realignment of the compression gratings and the end mirror within the pulse compressor of the laser amplifier, the spatial discrepancies of the spectra between the upper and lower as well as the left and right ends of the mode could be effectively eliminated.

C.1.2 Pulse contrast

Intensity autocorrelations allow us to precisely identify the temporal characteristics of the pulse, separating the main peak from its broad pedestal. To this end, the output pulses of the laser amplifier are guided through a 50/50 beamsplitter, generating two identical copies of the incident pulse that are delayed with respect to one another using a motorised stage and subsequently spatially overlapped under a small acute angle within a β-barium borate (BBO) crystal as depicted in Suppl. Fig. C.2. A long motorised stage allows for the relative delay of the two autocorrelator arms by up to $\pm$ 400 ps. Apart from the second harmonic of each near-infrared beam separately, at the temporal overlap a sum-frequency signal arises. By recording the intensity of this signal for varying delay times, information on the temporal structure of the pulses can be obtained. The sum-frequency signal resembles a convolution of the pulse with the time-delayed copy of itself:

$$I_{\mathrm{SHG}}(\tau) \propto \left| \int_{-\infty}^{+\infty} E(t) \cdot E(t-\tau)\,\mathrm{d}t \right|^2 \tag{C.1}$$

Due to the spatial separation of the second-harmonic and the sum-frequency signal achieved by placing an aperture in front of the detector photodiode, only the signal arising from the interaction of the two arms is filtered out for detection. The signal-to-noise ratio can be further improved by periodically chopping the pulse train in one arm and demodulating the photodiode signal in a lock-in amplifier.

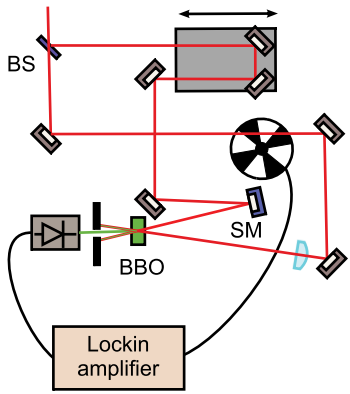


Figure C.2 | Autocorrelator for pulse characterisation. An incident transient is split into two identical copies using a 50/50 beamsplitter (BS). The beam path in one part can be varied using a motorised delay stage (DS). Subsequently, both beams are focused into the same β-barium borate (BBO) crystal under a small acute angle using a spherical mirror (SM) and a fused silica lens (L), respectively. When both beampaths temporally and spatially overlap apart from the second harmonic of each arm separately, a substantial sum-frequency signal will arise. The autocorrelation signal of the two pulses is encoded in the sum-frequency signal and therefore only this part is detected.

C.2 Extended data on the Herriott cells

The spectral broadening scheme in Herriott-type multipass cells is critically dependent on multiple laser characteristics. In the pursuit to achieve optimal broadening, temporal compression, while still operating the cells in a stable regime, different parameters were optimised. These parameters include the type of noble gas, its pressure, the near-infrared input-pulse duration and energy, as well as the pulse compression between the two individual cells.

C.2.1 Saturation effect in Herriott cells.

Of note, the second Herriott cell (HC2) is driven in the energy saturation regime. This implies that the pulse energy focused into HC2 is larger than the minimal pulse energy necessary to achieve the desired spectral broadening. The higher pulse energy leads to a higher peak intensity and thus larger self-phase modulation in the first few passes. As, therefore, the imprinted dispersion is larger than what the chirped mirrors can compensate for, the pulses are dispersed after the first few resonator

passes. Thus, in contrast to the case with an optimal input pulse energy, the efficient spectral broadening is pushed to the first few passes instead of evenly spacing it over all 32 round trips. Additionally, the amount of higher orders of dispersion picked up by the pulse in all subsequent passes is larger. This necessitates a chirped mirror compressor with more negative GDD than in the case of optimal input pulse energy. Yet, this poses no real issue and the additional group delay dispersion (GDD) is easily compensated for by more reflections on the chirped mirrors. The mentioned limitations are acceptable in favour of the crucial advantage of a reduced sensitivity to small variations of the input pulse energy. This allows for an increased stability of the output spectrum, which is indispensable for our applications.

C.2.2 Optimised spectral broadening in Herriott cells

Operating the laser amplifier at low repetition rates (100−500 kHz) and upon optimising the dispersion of the near-infrared input pulse and the gas pressure, the spectral broadening efficiency of the dual-stage Herriott cell could be greatly improved. The gradual broadening of the spectra after passing the first and both multipass cells is visualised in Suppl. Fig. C.3a and Suppl. Fig. C.4, respectively. The Fourier-transform-limit (FTL) drops to around 75 fs after the first stage (see Suppl. Fig. C.3b) and to well below 20 fs when both stages are operated in series (Suppl. Fig. C.4b).

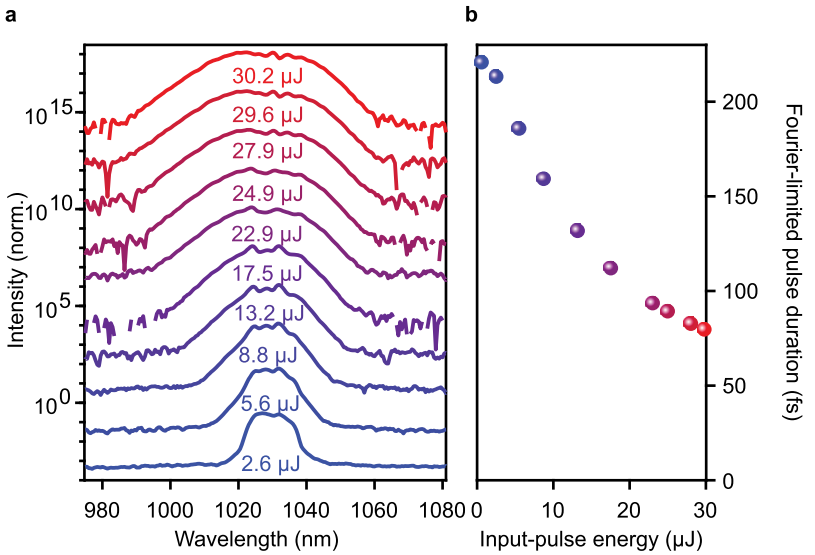


Figure C.3 | Energy-dependent spectral broadening of HC1 at a repetition rate of 500 kHz. a, The output spectra of HC1 for increasing input pulse energies are plotted in a semi-logarithmic scale. Subsequent curves are offset by one order of magnitude and measurement values below 10^{-3} of the maximum are omitted for clarity. **b**, As a result of the spectral broadening, the FTL pulse duration decreases with rising input pulse energy, dropping below 80 fs for an input pulse energy of 29 µJ.

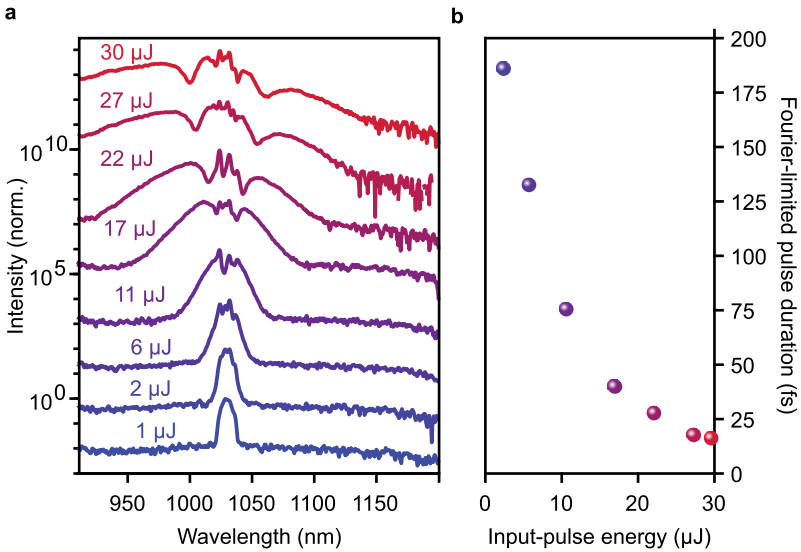


Figure C.4 | Spectral broadening of the dual-stage Herriott cell at a repetition rate of 500 kHz. a, The output spectra of the dual-stage HC for increasing input pulse energies are plotted in a semi-logarithmic scale. Subsequent curves are offset by one order of magnitude for clarity. **b**, The FTL pulse duration drops continuously for increasing input pulse energies, reaching a duration below 20 fs for an input pulse energy of 30 µJ.

C.2.3 Compression of the HC2 output pulses

The pulses emitted by HC2 after spectral broadening are compressed by means of a chirped mirror compressor within the housing of the device. In order to quantify the remaining chirp of the pulses they are characterised by means of frequency-resolved optical gating (FROG). For the mode of operation see Fig. 4.13. The measured and reconstructed spectrograms are depicted in Suppl. Fig. C.5a+b. The retrieved intensity envelope exhibits a width of 22 fs. The intensity and phase spectra corresponding to the retrieved spectrogram reveal the broadband spectrum and the remaining chirp. In order to further compress the extremely broadband pulses to close to their FTL duration, a series of additional chirped mirrors are necessary. Implementing four additional reflections, each imprinting $-50\,\mathrm{fs}^2$, and a pair of fused silica wedges with variable thickness, the pulses can be compressed to 13.5 fs at an FTL of 8 fs (see Fig. 4.15). Supplementary Figure C.5 illustrates the recorded (panel d) and reconstructed (panel e) frequency-resolved optical gating (FROG) traces as well as the spectral intensity and phase (panel f) after the additional chirped mirrors.

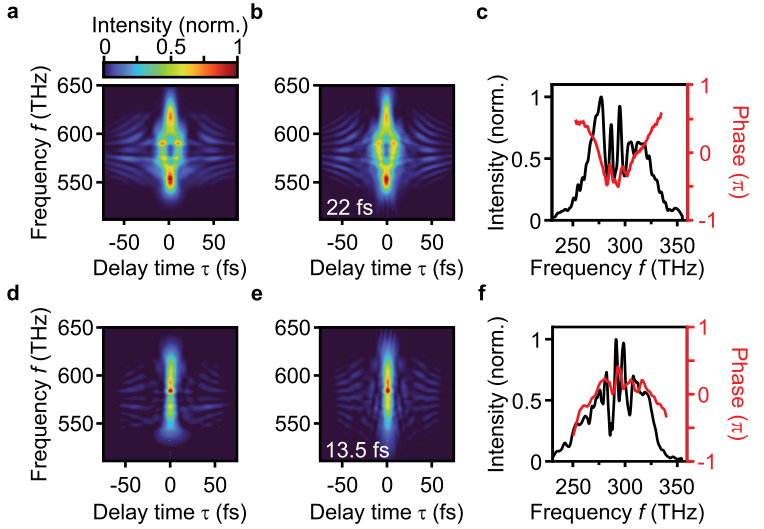


Figure C.5 | FROG measurements of uncompressed and compressed HC2 output pulses. a, A FROG spectrogram is recorded directly after the second broadening stage. **b**, The reconstructed FROG trace shows a pulse duration (intensity envelope) of 22 fs. **c**, The broadband spectral intensity (black) and the spectral phase (red) of the retrieved uncompressed pulse are plotted. **d**, The measured trace after compressing the output of HC2 with additional chirped mirrors is more concentrated around zero delay-time than the spectrogram depicted in panel a. **e**, The reconstructed FROG trace after compression exhibits a pulse duration (intensity envelope) of 13.5 fs. **d**, The spectral intensity (black) and the spectral phase (red) of the compressed retrieved pulse are depicted.

C.3 Extended data on mid-infrared generation

C.3.1 Simulation of optical rectification

The generation of sub-single-cycle mid-infrared transients from optical rectification of broadband near-infrared pulses in nonlinear crystals can be simulated in the following manner. The argumentation and the equations closely follow the book by Boyd [Boy08] and the diploma thesis of Huber [Hub00b]. As light is an electromagnetic wave, its propagation through the nonlinear crystal can be described by the nonlinear wave equation for each frequency component of the field. The electric field and the nonlinear polarisation (introduced in eq. 2.10) are written as sums of their n positive frequency components ω_n:

$$\tilde{\mathbf{E}}(\mathbf{r},t) = \sum_n \tilde{\mathbf{E}}_n(\mathbf{r},t); \quad \tilde{\mathbf{P}}^{\mathrm{NL}}(\mathbf{r},t) = \sum_n \tilde{\mathbf{P}}^{\mathrm{NL}}_n(\mathbf{r},t) \tag{C.2}$$

where each frequency component is a complex quantity defined as:

$$\tilde{\mathbf{E}}_n(\mathbf{r},t) = \mathbf{E}_n(\mathbf{r})e^{-i\omega_n t} + c.c. \tag{C.3}$$

$$\tilde{\mathbf{P}}_n(\mathbf{r},t) = \mathbf{P}^{\mathrm{NL}}_n(\mathbf{r})e^{-i\omega_n t} + c.c. \tag{C.4}$$

This notation leads to the nonlinear wave equation given by the following formula:

$$\nabla^2\tilde{\mathbf{E}}_n - \frac{\varepsilon(\omega_n)}{c^2}\frac{\partial^2\tilde{\mathbf{E}}_n}{\partial t^2} = \frac{1}{\varepsilon_0 c^2}\frac{\partial^2\tilde{\mathbf{P}}^{\mathrm{NL}}_n}{\partial t^2} \tag{C.5}$$

The variable $\varepsilon(\omega)$ denotes the linear dielectric function and ε_0 is the vacuum dielectric constant. Assumptions and simplifications allow for the analytical solution of this differential equation. First, a plane-wave ansatz can be made, as the focal spot is orders of magnitude larger than the wavelengths involved, warranting minimal angular divergence. Additionally, due to low nonlinear susceptibilities χ_{NL} in all materials, the source term on the ride hand side of equation C.5 is small ($\tilde{\mathbf{P}}^{\mathrm{NL}}_n \propto \chi_{\mathrm{NL}}$). If a plane wave propagating in the z-direction with the field amplitude A is employed, the electric field can be written as:

$$\mathbf{E}_{\mathrm{MIR}}(z,\omega) = \mathbf{A}_{\mathrm{MIR}}(z,\omega_{\mathrm{MIR}})e^{i(k(\omega_{\mathrm{MIR}})z-\omega_{\mathrm{MIR}}t)} \tag{C.6}$$

Inserting this ansatz into equation C.5 results in a differential equation for A_{MIR}. The incident electric fields shall be abbreviated in the same fashion:

$$\mathbf{E}_{\mathrm{NIR},1}(z,\omega_1) = \mathbf{A}_{\mathrm{NIR},1}(z,\omega_1)e^{i(k_1 z-\omega_1 t)} \tag{C.7}$$

$$\mathbf{E}_{\mathrm{NIR},2}(z,\omega_2=\omega_1-\omega_{\mathrm{MIR}}) = \mathbf{A}_{\mathrm{NIR},2}(z,\omega_2)e^{i(k_2 z-\omega_2 t)} \tag{C.8}$$

The generation of mid-infrared radiation along the propagation direction is thus given by the following differential equation for the amplitude:

$$\left[\frac{\partial^2 A_{\mathrm{MIR}}(z,\omega_{\mathrm{MIR}})}{\partial z^2}+2ik_{\mathrm{MIR}}\frac{\partial A_{\mathrm{MIR}}(z,\omega_{\mathrm{MIR}})}{\partial z}\right]\cdot e^{ik_{\mathrm{MIR}}z} = - \frac{4d_{\mathrm{eff}}\omega_{\mathrm{MIR}}^3}{c^2}A_{\mathrm{NIR},1}A_{\mathrm{NIR},2}\cdot e^{i(k_1+k_2)z} \tag{C.9}$$

with the effective nonlinearity d_{eff} for a given generation geometry:

$$d_{\mathrm{eff}}^{\mathrm{eoo}} = d_{22}\cos(\theta)\sin(3\phi) \quad d_{\mathrm{eff}}^{\mathrm{eoe}} = d_{22}\cos^2(\theta)\cos(3\phi) \tag{C.10}$$

wherein θ and ϕ denote the phase-matching and azimuthal angle, respectively, as depicted in Fig. 4.18.
Inserting the phase mismatch Δk as introduced previously:

$$\Delta k = |\mathbf{k}_1-\mathbf{k}_2-\mathbf{k}_3| = |\mathbf{k}_1(\theta)-\mathbf{k}_2-\mathbf{k}_{\mathrm{MIR}}(\theta)| \tag{C.11}$$

into eq. C.9, we arrive at

$$\frac{\partial^2 A_{\mathrm{MIR}}(z,\omega_{\mathrm{MIR}})}{\partial z^2}+2ik_{\mathrm{MIR}}\frac{\partial A_{\mathrm{MIR}}(z,\omega_{\mathrm{MIR}})}{\partial z} = -\frac{4d_{\mathrm{eff}}\omega_{\mathrm{MIR}}^3}{c^2}A_{\mathrm{NIR},1}A_{\mathrm{NIR},2}\cdot e^{i\Delta kz} \tag{C.12}$$

Since the nonlinearities are usually moderate, the mid-infrared amplitude grows very slowly compared to the timescales of the oscillation frequency:

$$\left|\frac{\partial^2 A_{\mathrm{MIR}}(z,\omega_{\mathrm{MIR}})}{\partial z^2}\right| \ll \left|\frac{\partial A_{\mathrm{MIR}}(z,\omega_{\mathrm{MIR}})}{\partial z}\right| \tag{C.13}$$

This simplifies equation C.9 to:

$$\frac{\partial A_{\mathrm{MIR}}}{\partial z} = \frac{2id_{\mathrm{eff}}\omega_{\mathrm{MIR}}^3}{k_{\mathrm{MIR}}c} A_{\mathrm{NIR},1} A_{\mathrm{NIR},2} e^{i\Delta kz} \tag{C.14}$$

As no mid-infrared electric field exists before the nonlinear interaction within the crystal, the boundary condition $A_{\mathrm{MIR}}(z = 0, \omega_{\mathrm{MIR}}) = 0$ is set. This first-order differential equation is then solved, resulting in an electric field oscillating at the mid-infrared frequency after passing a crystal of length L:

$$\begin{aligned} A_{\mathrm{MIR}}(L) &= \frac{2d_{\mathrm{eff}}\omega_{\mathrm{MIR}}}{k_{\mathrm{MIR}}} c^2 A_{\mathrm{NIR},1} A_{\mathrm{NIR},2} \cdot \int_0^L dz\, e^{i\Delta kz} \\ &= \frac{2d_{\mathrm{eff}}\omega_{\mathrm{MIR}}}{k_{\mathrm{MIR}}} c^2 A_{\mathrm{NIR},1} A_{\mathrm{NIR},2} \cdot \left(\frac{e^{i\Delta kz} - 1}{i\,\Delta k} \right) \end{aligned} \tag{C.15}$$

Therefore, given the material dispersion at all involved frequencies and the amplitude spectrum of the incident field A_{NIR}, the mid-infrared components can be calculated for a crystal with known thickness and nonlinear susceptibility $\chi^{(2)}$.
The intensity of the wave at ω_{MIR} can be calculated as:

$$I_{\mathrm{MIR}} = 2n_3\varepsilon_0 c |A_{\mathrm{MIR}}|^2 \tag{C.16}$$

$$= \frac{8n_3\varepsilon_0 d_{\mathrm{eff}}^2 \omega_{\mathrm{MIR}}^3}{k_{\mathrm{MIR}}}^2 c^3 |A_{\mathrm{NIR},1}|^2 |A_{\mathrm{NIR},2}|^2 \cdot \left| \frac{e^{i\Delta kz} - 1}{\Delta k} \right| \tag{C.17}$$

The last term appearing in this equation, the phase-matching factor, can be expressed as:

$$\left| \frac{e^{i\Delta kz} - 1}{\Delta k} \right| = L^2 \mathrm{sinc}^2\left(\frac{\Delta kL}{2}\right) \tag{C.18}$$

yielding the final result for the intensity of the generated mid-infrared radiation:

$$I_3 = \frac{8d_{\mathrm{eff}}^2 \omega_{\mathrm{MIR}}^2 I_{\mathrm{NIR},1} I_{\mathrm{NIR},2}}{n_1 n_2 n_3 \varepsilon_0 c^2} L^2 \mathrm{sinc}^2\left(\frac{\Delta kL}{2}\right) \tag{C.19}$$

Hence, the efficiency of MIR generation critically depends on the intensities of the incident waves, the phase-matching factor, and the properties of the crystal (length and effective nonlinearity). Figure 4.19 shows the result of a phase-matching simulation using the known near-infrared input pulse.

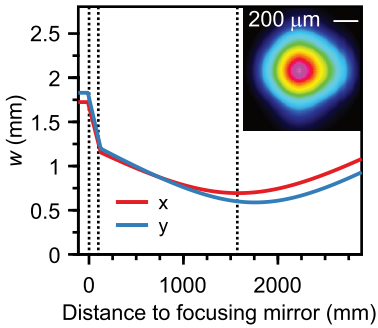


Figure C.6 | Beam caustic and measured focus. The change of the horizontal and vertical beam radii is calculated for the focusing optics used. Curved mirrors with radii of curvatures of ROC = −800 mm and ROC = +600 mm are placed in a distance of 137 mm (first two dashed vertical lines) from each other. The generation crystal is placed at a distance of around 1570 mm from the first mirror (third dashed line). Inset: The experimentally determined intensity distribution in the focal plane reveals the Gaussian beam profile.

C.3.2 Mid-infrared generation focus

In order to generate mid-infrared radiation by means of optical rectification within a gallium selenide (GaSe) crystal, the beam must be focused to create sufficiently large peak intensities. To focus the near-infrared pulses into the nonlinear crystal a concave mirror with a radius of curvature (ROC) of −800 mm and a convex mirror with an ROC of +600 mm were placed at a distance of 138 mm from each other(see Suppl. Fig. C.6a, vertical dashed lines). Since the calculated locations of the focal planes in the x- and y-direction do not perfectly coincide, the GaSe crystal is placed in between the two focal positions at a distance of 1570 mm from the focusing mirror. The calculated beam waists at the crystal position of $w_{\mathrm{x,\,sim}} = 641\,\mu\mathrm{m}$ and $w_{\mathrm{y,\,sim}} = 630\,\mu\mathrm{m}$ differ from the experimental values of $w_{\mathrm{x}} = 750\,\mu\mathrm{m}$ and $w_{\mathrm{y}} = 660\,\mu\mathrm{m}$ (Suppl. Fig. C.6b) The reasons for this could include a deviation from the estimated beam quality of $M^2 = 1.2$ and a remaining small beam convergence of the pulses exiting the Herriott cells.

C.3.3 Optimisation of mid-infrared generation

The ultra-broadband near-infrared spectrum created by the dual-stage multipass cell can support difference frequencies from 0 to 80 THz. Therefore, these close to Fourier-transform-limited near-infrared input pulses are a perfect starting point for

the generation of ultrashort mid-infrared transients with a multi-octave spanning spectrum. However, as mentioned in previous chapters, the phase-matching condition must be fulfilled in the process of optical rectification (see eq. 4.29). This limiting factor leads to a mid-infrared spectrum much narrower than the possible difference frequencies given by a convolution of the spectrum with itself. The blue curve in Fig. 4.19c shows the simulated spectra for infinitely broadband phase matching and for a GaSe crystal with a thickness of 50 µm and a phase-matching angle of $\theta = 14°$ when using the measured HC2 output spectra. It becomes apparent that in spite of the broad input spectrum, only a fraction of spectral components contribute to the output spectrum.

Cutting the spectrum. As mentioned before, the generation of single-cycle mid-infrared transients necessitates a broad spectrum. The highest possible frequency component that can be generated is determined by the bandwidth of the incident near-infrared spectrum. In general, there is no limit for the lowest frequency component. However, terahertz spectral components below 11 THz do not contribute much to the formation of an ultrashort mid-infrared pulse at a centre frequency of 25 THz and could even increase the pulse duration. As the mid-infrared spectrum is created by means of difference-frequency generation (DFG), near-infrared components in the incident pulse that are separated by less than 11 THz create this radiation. Reducing the spectral components in the original near-infrared spectrum with a difference of less than 11 THz can thus suppress the generation of these few-terahertz components. A possibility to do so lays in cutting out the central part of the broadband near-infrared spectrum. Using solely the spectral wings, which are more than 11 THz apart from each other, for mid-infrared generation would suffice to create the desired pulse.

In fact, the central part of the near-infrared spectrum even diminishes the capabilities of our mid-infrared generation. Even after the spectral broadening cells, a large portion of the power is still located close to the central frequency of 295 THz. These unwanted spectral components increase the thermal stress on the nonlinear material and limit the focus size that the crystal can sustain over long periods of time. Additionally, these low-frequency components are critical for gallium selenide specifically, as an optical phonon lies at around a frequency of 11 THz, which reduces the crystal's damage threshold if exited resonantly. Thus, these frequencies are undesirable and

should be removed for optimisation of the mid-infrared generation.
To this end, a prism compressor based on highly-dispersive prisms fabricated out of SF11 and a specially designed end mirror was set up. The near-infrared pulses are focused through two SF11 prisms under the Brewster's angle onto the end mirror using a broadband dielectric mirror with a focal length of 1 m. The first prism leads to angular dispersion of the broadband spectral components that are then refracted to travel collinearly again by the second prism. The end mirror consists of two D-cut dielectric mirrors, which can be individually inserted into the line-shaped beam profile from the two sides. The central part of the spectrum thus passes between the two mirrors and is dumped. Subsequently, the two wings are reflected back the same beam path and superimposed again after the first prism. The name prism compressor may be misleading in this context as perfectly compressed near-infrared pulses are incident. The two functions of this setup are to cut out the central spectral region and recompress the two separate spectral wings. This can be achieved by varying the separation and the material insertion of the two prisms. By placing one D-cut mirror on a stage with a micrometer screw, one of the wings can be temporally delayed with femtosecond precision with respect to the other.
Supplementary Figure C.7 shows the near-infrared spectrum after the prism compressor setup with the fundamental around 295 THz removed. A percentage of 72 % of the incident power remains in the wings and can be used for mid-infrared generation. The simulated difference-frequency spectrum spans from a frequency of 6.2 THz to 35.5 THz ($\frac{1}{e}$-width) and would therefore support a sub-cycle mid-infrared pulse.
A large drawback of the cutting and compression setup used became apparent when attempting to compress the two wings. As we wish to very precisely choose the spectral region to cut out using the end mirror, the spectral components must be spatially separated considerably. In the case of a distance of 0.7 m between the two prisms, this meant a splitting of around 20 mm between the components with wavelengths 950 nm and 1110 nm. Due to this broad splitting, prism no. 2 had to be fully inserted in order to accommodate all spectral components. Two chirped broadband mirrors with -50 fs^2 had to be inserted between the prisms no. 1 and 2 to even reach an overall GDD of 0 for for the central wavelength 1030 nm. At such long propagation distances within the SF11 medium, third-order dispersion of the material becomes significant, reaching values larger than a TOD of -10 000 fs^3. In combination with the

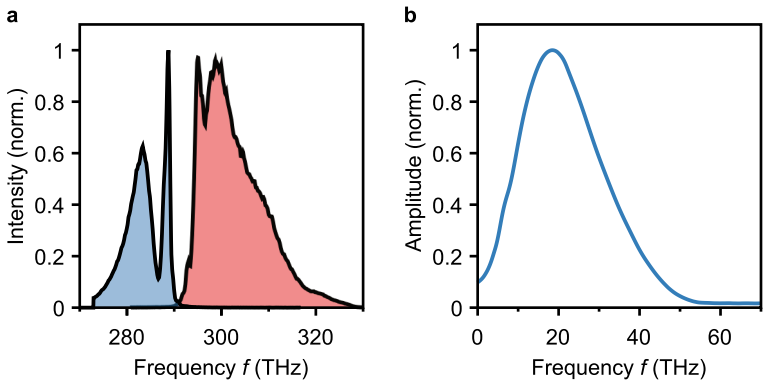


Figure C.7 | Cut spectrum and possible difference frequencies. a, Spectra of the two wings – high energy (red) and low energy (blue) – after the prism compressor setup. **b**, Calculated possible difference-frequency spectrum in the absence of phase-matching constraints.

broadband spectrum, this rendered it impossible to simultaneously compress both spectral wings of the cut spectrum.

Due to this issue and the reflection losses occurring at the eight SF11/air interfaces as well as slightly too narrowband mirrors, it was decided to put this prism compressor on hold. Instead, the broadband, transform-limited pulses out the multipass cell are directly focused into the nonlinear crystal for mid-infrared generation.

In a future setup, these issues could be addressed in part by employing finer tunable D-mirror insertions that would render it unnecessary to spatially split the spectrum this immensely. Additionally, using chirped D-cut end mirrors, the compression of each arm could be done separately. Moreover, one could even think about using two prisms such that two spectral wings propagate through different prisms, allowing for separate optimisation.

Pulse synthesis. With the gallium selenide crystal used, phase matching not only limits the generated spectrum on the low-frequency side due to an optical phonon but also leads to a cut-off at high frequencies of around 40 THz (see spectra in Fig. 4.21). This means that the possible difference frequencies in the HC2 output pulses from the most outer wings of the near-infrared spectrum with a difference of 40-70 THz

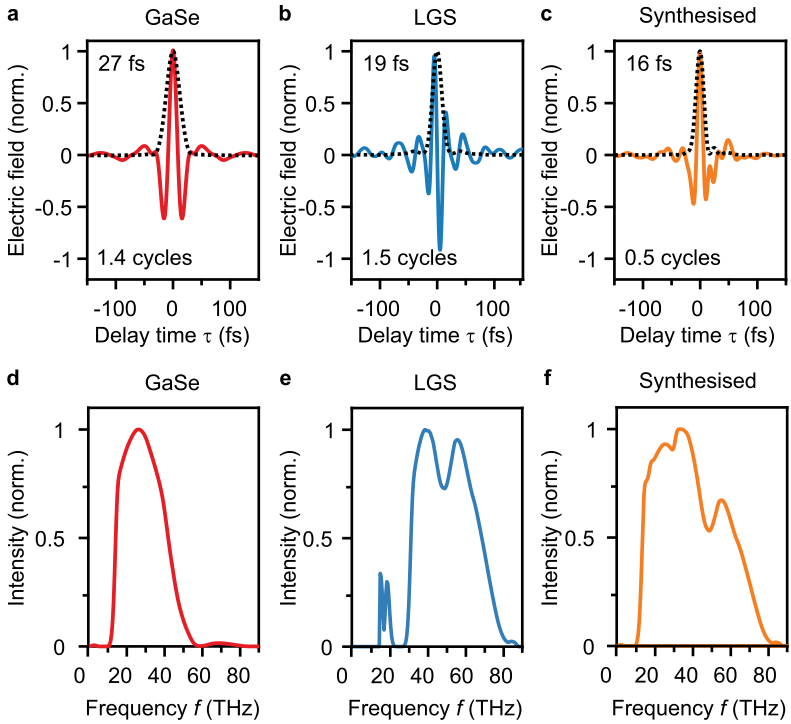


Figure C.8 | Pulse synthesis. a-c, Simulated transients for a GaSe crystal (red) with a thickness of 50 µm, a LGS crystal (blue) with a thickness of 250 µm and coherent superposition of the two waveforms (orange). The FWHM duration of the intensity envelope and the number of optical cycles at the centre frequency are written in the top and bottom left corners, respectively. **d-f**, The corresponding normalised ultrabroadband intensity spectra are obtained by performing Fourier transforms of the waveforms displayed in panels a-c.

are barely generated. However, MIR radiation at this frequency could be generated by employing a different nonlinear crystal with a smaller phase mismatch for DFG in this spectral region. The nonlinear crystal lithium-gallium-disulphide ($LiGaS_2$ or short LGS) is widely used for this MIR regime. It exhibits a broad transparency for wavelengths between 0.33µm and 11.6 µm (frequencies between 25 THz and 90 THz), a high nonlinear susceptibility of 5.6 pm/V [Nam21], and a high damage threshold. A drawback of LGS, however, is the optical phonon at around 25 THz with a bandwidth of 6 THz. As a result, its damage threshold is drastically reduced if the phonon resonantly driven by the generated MIR radiation at this frequency. These properties make this crystal suitable for the generation of MIR frequencies from above 28 THz up to 100 THz. It has been shown by Merkl *et al.*[Mer19b] that by coherent superposition of the radiation generated in two separate nonlinear crystals, extremely short, sub-cycle mid-infrared pulses can be generated. By using the 20 fs short output pulses of a Ti:sapphire laser amplifier operating at a repetition rate of 0.4 MHz, they showed that making use of type-II DFG in a GaSe crystal with a thickness of 70 µm and type-I DFG in a LGS crystal with a thickness of 300 µm, a spectrum spanning 1.5 octaves from a frequency of 11 THz to 30 THz can be generated. This approach should in general be scalable to higher repetition rates and to different pump laser fundamental wavelengths (from 800 nm to 1030 nm). Supplementary Figure C.8 displays the results of a simulation utilising the Herriott cell output spectrum and assuming transform-limited near-infrared pulses. The mid-infrared transients for a GaSe crystal with a thickness of 50 µm in type-II phase matching and 14° internal phase-matching angle as well as a LGS crystal with a thickness of 250 µm (type I, 47°) are plotted in Suppl. Fig. C.8a and b. The intensity envelope, represented by the black dashed line, exhibits an FWHM of 27 fs and 19 fs, respectively. The octave-spanning spectra (d+e) exhibit maxima at frequencies of 26 THz (GaSe) and 39 THz (LGS), respectively.

By coherent superposition of the two waveforms using a beamcombiner and precise control over the relative femtosecond delay time between the two waveforms generated in the two separate crystals, a half-cycle mid-infrared transient with an FWHM of only 16 fs can be created. The extreme asymmetry ratio of 2.1:1 and peak width (top 90%) of less than 4 fs makes this type of transient optimal for the application in LW-STM. The mid-infrared spectrum obtained by Fourier transform spans from

14 THz to 63 THz and thus two optical octaves.
In the spectra and waveforms generated in LGS, the strong phonon absorption around 26 THz becomes clearly visible. Notably, due to the much lower intensity within the wings of the near-infrared spectrum, the generated high-frequency spectral components are equally meagre. To achieve the synthesised transient in Suppl. Fig. C.8c, the field strengths of the two waveforms generated in GaSe and LGS must be equally strong. This can be facilitated by splitting the near-infrared power unequally and thus taking into account the differences in nonlinearities and generation efficiencies.
Mid-infrared beam diameter. In order to estimate the focal size of the mid-infrared radiation in the tip-sample junction, the beam caustic in the far field must be determined. To this end, knife-edge experiments are performed. A razor blade is inserted gradually from one side while the transmitted power is recorded in a pyrometer (see Suppl. Fig. C.9). The curves are normalised and the data for the horizontal and vertical direction are offset horizontally for clarity. Assuming a Gaussian beam profile, the beam diameters in the vertical and horizontal direction can be extracted by fitting error functions to the resulting knife insertion vs. power curves. Thereby, the mid-infrared beam diameter is determined to be 5.53 mm ($\frac{1}{e^2}$) in both directions, which corresponds to an FWHM of 4.72 mm. The beam diameter remains roughly the same after a metre of propagation hinting towards a well collimated beam.

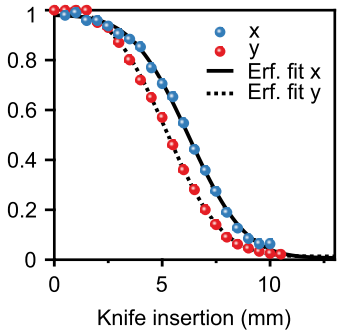


Figure C.9 | Determining the mid-infrared beam diameter. Gradually inserting a knife in the horizontal (x, blue) or vertical direction (y, red) reduces the detected power (here normalised to the respective maximum). Assuming a Gaussian beam profile, fitting the curve with error functions allows for the determination of the $\frac{1}{e^2}$ beam diameters of 5.53 mm (FWHM = 4.72 mm) in both directions. The curves are offset horizontally for clarity.

Appendix D

Publications

Publications in peer-reviewed journals

- Philipp Merkl, Matthias Knorr, Christian Meineke, Lukas Kastner, Dominik Peller & Rupert Huber
  Multibranch pulse synthesis and electro-optic detection of subcycle multi-terahertz electric fields
  Optics letters, **44**, 5521–5524 (2019).

- Dominik Peller, Lukas Z. Kastner, Thomas Buchner, Carmen Roelcke, Florian Albrecht, Nikolaj Moll, Rupert Huber & Jascha Repp
  Sub-cycle atomic-scale forces coherently control a single-molecule switch
  Nature, **585**, 58–62 (2020).

- Dominik Peller, Carmen Roelcke, Lukas Z. Kastner, Thomas Buchner, Alexander Neef, Johannes Hayes, Franco Bonafé, Dominik Sidler, Michael Ruggenthaler, Ángel Rubio, Jascha Repp & Rupert Huber
  Quantitative sampling of atomic-scale electromagnetic waveforms
  Nature Photonics, **15**, 143–147 (2021).

- Christian Meineke, Michael Prager, Johannes Hayes, Qiannan Wen, Lukas Z. Kastner, Dieter Schuh, Kilian Fritsch, Oleg Pronin, Markus Stein, Felix Schäfer, Sangam Chatterjee, Mackillo Kira, Rupert Huber & Dominique Bougeard

Scalable high-repetition-rate sub-half-cycle terahertz pulses from spatially indirect interband transitions
Light: Science & Applications, **11**, 151 (2022).

- Carmen Roelcke*, Lukas Z. Kastner*, Maximilian Graml, Andreas Biereder, Jan Wilhelm, Jascha Repp, Rupert Huber & Yaroslav A. Gerasimenko
  * equal contributions
  Ultrafast videographic tunnelling spectroscopy of a single vacancy in an atomically thin crystal
  Nature Photonics, accepted in principle (2023).

- Yaroslav A. Gerasimenko, Vincent Eggers, Lukas Z. Kastner, Carmen Roelcke, Dominik Peller, Jascha Repp & Rupert Huber
  Quantum squeezing of a single molecule by direct femtosecond atomic forces
  in preparation.

Publications in international conference proceedings

- L.Z. Kastner, D. Peller, T. Buchner, C. Roelcke, F. Albrecht, N. Moll, R. Huber & J. Repp
  Coherent Control of Single-Molecule Switching Reactions with Femtosecond Atomic Forces
  Conference on Lasers and Electro-Optics, OSA Technical Digest (Optical Society of America, 2021) | paper FW3M.5
  doi:https://doi.org/10.1364/CLEO_QELS.2021.FW3M.5

Contributions at conferences

- **International Conference on Ultrafast Phenomena (UP) 2020, conference**
  November 16 – 19, 2020| virtual conference
  Highlighted talk
  Lukas Z. Kastner, Dominik Peller, Thomas Buchner, Carmen Roelcke, Florian Albrecht, Nikolaj Moll, Jascha Repp & R. Huber
  Coherent Control of a Single-Molecule Switch with Sub-Cycle Atomic Forces

- **Laser Science to Photonic Applications (CLEO 2021)**
  May 9 – 14, 2021| virtual conference
  Highlighted talk
  L.Z. Kastner, D. Peller, T. Buchner, C. Roelcke, F. Albrecht, N. Moll, R. Huber, & J. Repp
  Coherent Control of Single-Molecule Switching Reactions with Femtosecond Atomic Forces

- **DPG-Tagung der Sektion Atome, Moleküle, Quantenoptik und Plasmen (SAMOP 2021)**
  September 20 – 24, 2021| virtual conference
  Poster
  L. Kastner, D. Peller, C. Roelcke, T. Buchner, A. Neef, J. Hayes, F. Bonafé, D. Sidler, M. Ruggenthaler, A. Rubio, J. Repp & R. Huber,
  Quantitative Waveform Sampling on Atomic Scales

- **84. Jahrestagung der DPG und DPG-Tagung der Sektion Kondensierte Materie (SKM 2021)**
  September 27 – October 1, 2021| virtual conference
  Poster
  L. Kastner, D. Peller, C. Roelcke, T. Buchner, A. Neef, J. Hayes, F. Bonafé, D. Sidler, M. Ruggenthaler, A. Rubio, J. Repp & R. Huber,
  Quantitatively sampling Atomic Scale Waveforms

- **Internal SFB 1277 Workshop**
  October 4 – 6, 2021| Waldsassen, Germany
  Contributed talk
  L. Kastner, D. Peller, C. Roelcke, T. Buchner, A. Neef, J. Hayes, F. Bonafé, D. Sidler, M. Ruggenthaler, A. Rubio, J. Repp & R. Huber
  Sampling Near-field Waveforms on the Atomic Scale

- **International Conference on Ultrafast Phenomena (UP) 2022, conference**
  July 18 – 22, 2022| Montréal, Canada
  Highlighted talk
  L.Z. Kastner , D. Peller, C. Roelcke, T. Buchner, A. Neef, J. Hayes, F. Bonafé, D. Sidler, M. Ruggenthaler, A. Rubio, J. Repp & R. Huber
  Quantitative sampling of atomic-scale half-cycle terahertz waveforms

- **47th International Conference on Infrared, Millimiter and Terahertz Waves (IRMMW 2022), conference**
  August 28 – September 2, 2022| Delft, Netherlands
  Keynote talk
  L.Z. Kastner, D. Peller, T. Buchner, C. Roelcke, F. Albrecht, N.Moll, J. Repp & R. Huber
  Coherent Control of Single-Molecule Switching Reactions with Sub-Cycle Atomic Forces

- **9th International Conference on Antennas and Electromagnetic Systems (AES 2023), conference**
  June 5–9, 2022| Torremolinos, Spain
  Invited talk
  L.Z. Kastner , D. Peller, C. Roelcke, T. Buchner, A. Neef, J. Hayes, F. Bonafé, D. Sidler, M. Ruggenthaler, A. Rubio, J. Repp & R. Huber
  Terahertz Waveforms in the Atom-Scale Gap of a Scanning Tunnelling Microscope

Bibliography

Abb73. E. Abbe, *Beiträge zur Theorie des Mikroskops und der mikroskopischen Wahrnehmung*, Archiv für mikroskopische Anatomie **9**, 413–418 (1873).

Abd21. M. Abdo, S. Sheng, S. Rolf-Pissarczyk, L. Arnhold, J. A. J. Burgess, M. Isobe, L. Malavolti, and S. Loth, *Variable Repetition Rate THz Source for Ultrafast Scanning Tunneling Microscopy*, ACS Photonics **8**, 702–708 (2021).

All09. K. R. Allakhverdiev, M. O. Yetis, S. Özbek, T. K. Baykara, and E. Y. Salaev, *Effective nonlinear GaSe crystal. Optical properties and applications*, Laser Physics **19**, 1092–1104 (2009).

Amm21. S. E. Ammerman, V. Jelic, Y. Wei, V. N. Breslin, M. Hassan, N. Everett, S. Lee, Q. Sun, C. A. Pignedoli, P. Ruffieux, R. Fasel, and T. L. Cocker, *Lightwave-driven scanning tunnelling spectroscopy of atomically precise graphene nanoribbons*, Nature Communications **12**, 1–9 (2021).

Amm22. S. E. Ammerman, Y. Wei, N. Everett, V. Jelic, and T. L. Cocker, *Algorithm for subcycle terahertz scanning tunneling spectroscopy*, Physical Review B **105**, 1–14 (2022).

Ara22a. Y. Arashida, H. Mogi, M. Ishikawa, I. Igarashi, A. Hatanaka, N. Umeda, J. Peng, S. Yoshida, O. Takeuchi, and H. Shigekawa, *Subcycle Mid-Infrared Electric-Field-Driven Scanning Tunneling Microscopy with a Time Resolution Higher Than 30 fs*, ACS Photonics **9**, 3156–3164 (2022).

Ara22b. Y. Arashida, N. Umeda, H. Mogi, M. Ishikawa, A. Hatanaka, O. Takeuchi, S. Yoshida, and H. Shigekawa, *Phase-controllable multi-laser system with coaxially combined near-infrared and subcycle mid-infrared and THz pulsed beams*, Applied Physics Express **15** (2022).

Aro15. A. Arora, M. Koperski, K. Nogajewski, J. Marcus, C. Faugeras, and M. Potemski, *Excitonic resonances in thin films of WSe_2: From monolayer to bulk material*, Nanoscale **7**, 10421–10429 (2015).

Asp14. M. Aspelmeyer, T. J. Kippenberg, and F. Marquardt, *Cavity optomechanics*, Reviews of Modern Physics **86**, 1391–1452 (2014).

Aus84. D. H. Auston, K. P. Cheung, and P. R. Smith, *Picosecond photoconducting Hertzian dipoles*, Applied Physics Letters **45**, 284–286 (1984).

Avo11. A. Avogadro, *Masses of the Elementary Molecules of Bodies*, Journal de Physique **73**, 58–76 (1811).

Bar61. J. Bardeen, *Tunneling from a many-particle point of view*, Physical Review **6**, 57–59 (1961).

Bar03. W. L. Barnes, A. Dereux, and T. W. Ebbesen, *Surface plasmon subwavelength optics*, Nature **424**, 824–830 (2003).

Bar15. M. Barbry, P. Koval, F. Marchesin, R. Esteban, A. G. Borisov, J. Aizpurua, and D. Sánchez-Portal, *Atomistic near-field nanoplasmonics: Reaching atomic-scale resolution in nanooptics*, Nano Letters **15**, 3410–3419 (2015).

Bar19. S. Barja, S. Refaely-Abramson, B. Schuler, D. Y. Qiu, A. Pulkin, S. Wickenburg, H. Ryu, M. M. Ugeda, C. Kastl, C. Chen, C. Hwang, A. Schwartzberg, S. Aloni, S.-K. Mo, D. Frank Ogletree, M. F. Crommie, O. V. Yazyev, S. G. Louie, J. B. Neaton, and A. Weber-Bargioni, *Identifying substitutional oxygen as a prolific point defect in monolayer transition metal dichalcogenides*, Nature Communications **10**, 1–8 (2019).

Bas16. D. N. Basov, M. M. Fogler, and F. J. García De Abajo, *Polaritons in van der Waals materials*, Science **354** (2016).

Bat12. M. Battiato, K. Carva, and P. M. Oppeneer, *Theory of laser-induced ultrafast superdiffusive spin transport in layered heterostructures*, Physical Review B **86**, 1–16 (2012).

Bes87. K. Besocke, *An easily operable scanning tunneling microscope*, Surface Science **181**, 145–153 (1987).

Bin82. G. Binnig, H. Rohrer, C. Gerber, and E. Weibel, *Surface Studies by Scanning Tunneling Microscopy*, Physical Review Letters **49**, 57–61 (1982).

Böc19. H. Böckmann, M. Müller, A. Hammud, M. G. Willinger, M. Pszona, J. Waluk, M. Wolf, and T. Kumagai, *Near-Field Spectral Response of Optically Excited Scanning Tunneling Microscope Junctions Probed by Single-Molecule Action Spectroscopy*, Journal of Physical Chemistry Letters **10**, 2068–2074 (2019).

Bon95. A. Bonvalet, M. Joffre, J. L. Martin, and A. Migus, *Generation of ultrabroadband femtosecond pulses in the mid-infrared by optical rectification of 15 fs light pulses at 100 MHz repetition rate*, Applied Physics Letters **67**, 2907–2909 (1995).

Boy08. R. Boyd, *Nonlinear Optics*, 3 ed., Academic Press, USA, 2008.

Bur19. S. C. Burd, R. Srinivas, J. J. Bollinger, A. C. Wilson, D. J. Wineland, D. Leibfried, D. H. Slichter, and D. T. C. Allcock, *Quantum amplification of mechanical oscillator motion*, Science **364**, 1163–1165 (2019).

Cav81. C. M. Caves, *Quantum-mechanical noise in an interferometer*, Physical Review D **23**, 1693–1708 (1981).

Ceb16. F. Ceballos, Q. Cui, M. Z. Bellus, and H. Zhao, *Exciton formation in monolayer transition metal dichalcogenides*, Nanoscale **8**, 11681–11688 (2016).

Cer03. G. Cerullo and S. De Silvestri, *Ultrafast optical parametric amplifiers*, Review of Scientific Instruments **74**, 1–18 (2003).

Cha20. A. Chaves, J. G. Azadani, H. Alsalman, D. R. da Costa, R. Frisenda, A. J. Chaves, S. H. Song, Y. D. Kim, D. He, J. Zhou, A. Castellanos-Gomez, F. M. Peeters, Z. Liu, C. L. Hinkle, S.-H. Oh, P. D. Ye, S. J. Koester, Y. H. Lee, P. Avouris, X. Wang, and T. Low, *Bandgap engineering of two-dimensional semiconductor materials*, npj 2D Materials and Applications **4**, 29 (2020).

Che15. A. Chernikov, C. Ruppert, H. M. Hill, A. F. Rigosi, and T. F. Heinz, *Population inversion and giant bandgap renormalization in atomically thin WS_2 layers*, Nature Photonics **9**, 466–470 (2015).

Cia22. A. Ciarrocchi, F. Tagarelli, A. Avsar, and A. Kis, *Excitonic devices with van der Waals heterostructures: valleytronics meets twistronics*, Nature Reviews Materials **7**, 449–464 (2022).

Coc13. T. L. Cocker, V. Jelic, M. Gupta, S. J. Molesky, J. A. J. Burgess, G. D. L. Reyes, L. V. Titova, Y. Y. Tsui, M. R. Freeman, and F. A. Hegmann, *An ultrafast terahertz scanning tunnelling microscope*, Nature Photonics **7**, 620–625 (2013).

Coc16. T. L. Cocker, D. Peller, P. Yu, J. Repp, and R. Huber, *Tracking the ultrafast motion of a single molecule by femtosecond orbital imaging*, Nature **539**, 263–267 (2016).

Coo00. D. Cook and R. Hochstrasser, *Intense terahertz pulses by four-wave rectification in air*, Optics letters **25**, 1210–1212 (2000).

Cor99. J. Corson, R. Mallozzi, J. Orenstein, J. N. Eckstein, and I. Bozovic, *Vanishing of phase coherence in underdoped $Bi_2Sr_2CaCu_2O_{8+\delta}$*, Nature **398**, 221–223 (1999).

Cro93. M. F. Crommie, C. P. Lutz, and D. M. Eigler, *Confinement of electrons to quantum corrals on a metal surface*, Science **262**, 218–220 (1993).

Cun10. S. T. Cundiff and A. M. Weiner, *Optical arbitrary waveform generation*, Nature Photonics **4**, 760–766 (2010).

Dal08. J. Dalton, *A new system of chemical philosophy: part 1*, 2 ed., Printed by S. Russell for R. Bickerstaff, 1808.

Dan90. M. Dantus, R. M. Bowman, and A. H. Zewail, *Femtosecond laser observations of molecular vibration and rotation*, Nature **343**, 737–739 (1990).

DeM67. F. DeMartini, C. H. Townes, T. K. Gustafson, and P. L. Kelley, *Self-Steepening of Light Pulses*, Physical Review **164**, 312–323 (1967).

Dhi17. S. S. Dhillon, M. S. Vitiello, E. H. Linfield, A. G. Davies, M. C. Hoffmann, J. Booske, C. Paoloni, M. Gensch, P. Weightman, G. P. Williams, E. Castro-Camus, D. R. S. Cumming, F. Simoens, I. Escorcia-Carranza, J. Grant, S. Lucyszyn, M. Kuwata-Gonokami, K. Konishi, M. Koch, C. A. Schmuttenmaer, T. L. Cocker, R. Huber, A. G. Markelz, Z. D. Taylor, V. P. Wallace, J. A. Zeitler, J. Sibik, T. M. Korter, B. Ellison, S. Rea, P. Goldsmith, K. B. Cooper, R. Appleby, D. Pardo, P. G. Huggard, V. Krozer, H. Shams, M. Fice, C. Renaud, A. Seeds, A. Stöhr, M. Naftaly, N. Ridler, R. Clarke, J. E. Cunningham, and M. B. Johnston, *The 2017 terahertz science and technology roadmap*, Journal of Physics D: Applied Physics **50**, 043001 (2017).

Dir27. P. Dirac, *The quantum theory of the emission and absorption of radiation*, Proceedings of the Royal Society A **114**, 243–265 (1927).

Dua20. H.-G. Duan, A. Jha, X. Li, V. Tiwari, H. Ye, P. K. Nayak, X.-L. Zhu, Z. Li, T. J. Martinez, M. Thorwart, and R. J. D. Miller, *Intermolecular vibrations mediate ultrafast singlet fission*, Science Advances **6**, eabb0052 (2020).

Dun93. T. J. Dunn, J. N. Sweetser, I. A. Walmsley, and C. Radzewicz, *Experimental determination of the dynamics of a molecular nuclear wave packet via the spectra of spontaneous emission*, Physical Review Letters **70**, 3388–3391 (1993).

Dun95. T. J. Dunn, I. A. Walmsley, and S. Mukamel, *Experimental Determination of the Quantum-Mechanical State of a Molecular Vibrational Mode Using Fluorescence Tomography*, Physical Review Letters **74**, 884–887 (1995).

Efr82. B. Efron, *The Jackknife, the Bootstrap and Other Resampling Plans*, Society for Industrial and Applied Mathematics, 1982.

Eic10. M. Eichberger, H. Schäfer, M. Krumova, M. Beyer, J. Demsar, H. Berger, G. Moriena, G. Sciaini, and R. J. Miller, *Snapshots of cooperative atomic motions in the optical suppression of charge density waves*, Nature **468**, 799–802 (2010).

Eis14. M. Eisele, T. L. Cocker, M. A. Huber, M. Plankl, L. Viti, D. Ercolani, L. Sorba, M. S. Vitiello, and R. Huber, *Ultrafast multi-terahertz nano-spectroscopy with sub-cycle temporal resolution*, Nature Photonics **8**, 841–845 (2014).

Esa18. T. Esat, N. Friedrich, F. S. Tautz, and R. Temirov, *A standing molecule as a single-electron field emitter*, Nature **558**, 573–576 (2018).

Esm19. M. Esmann, S. F. Becker, J. Witt, J. Zhan, A. Chimeh, A. Korte, J. Zhong, R. Vogelgesang, G. Wittstock, and C. Lienau, *Vectorial near-field coupling*, Nature Nanotechnology **14**, 698–704 (2019).

Far16. M. Farmanbar and G. Brocks, *First-principles study of van der Waals interactions and lattice mismatch at MoS_2/metal interfaces*, Physical Review B **93**, 085304 (2016).

Fei17. A. Feist, N. Bach, N. Rubiano da Silva, T. Danz, M. Möller, K. E. Priebe, T. Domröse, J. G. Gatzmann, S. Rost, J. Schauss, S. Strauch, R. Bormann, M. Sivis, S. Schäfer, and C. Ropers, *Ultrafast transmission electron microscopy using a laser-driven field emitter: Femtosecond resolution with a high coherence electron beam*, Ultramicroscopy **176**, 63–73 (2017).

Fib00. G. Fibich and A. L. Gaeta, *Critical power for self-focusing in bulk media and in hollow waveguides*, Optics Letters **25**, 335–337 (2000).

För16. B. Förg, J. Schötz, F. Süßmann, M. Förster, M. Krüger, B. Ahn, W. A. Okell, K. Wintersperger, S. Zherebtsov, A. Guggenmos, V. Pervak, A. Kessel, S. A. Trushin, A. M. Azzeer, M. I. Stockman, D. Kim, F. Krausz, P. Hommelhoff, and M. F. Kling, *Attosecond nanoscale near-field sampling*, Nature Communications **7**, 1–7 (2016).

Ful14. F. D. Fuller, J. Pan, A. Gelzinis, V. Butkus, S. S. Senlik, D. E. Wilcox, C. F. Yocum, L. Valkunas, D. Abramavicius, and J. P. Ogilvie, *Vibronic coherence in oxygenic photosynthesis*, Nature Chemistry **6**, 706–711 (2014).

Gan06. S. Ganichev and W. Prettl, *Intense Terahertz Excitation of Semiconductors*, Oxford University Press, 2006.

Gar97. G. A. Garrett, A. G. Rojo, A. K. Sood, J. F. Whitaker, and R. Merlin, *Vacuum squeezing of solids: Macroscopic quantum states driven by light pulses*, Science **275**, 1638–1640 (1997).

Ger17. S. Gerber, S.-L. Yang, D. Zhu, H. Soifer, J. A. Sobota, S. Rebec, J. J. Lee, T. Jia, B. Moritz, C. Jia, A. Gauthier, Y. Li, D. Leuenberger, Y. Zhang, L. Chaix, W. Li, H. Jang, J.-S. Lee, M. Yi, G. L. Dakovski, S. Song, J. M. Glownia, S. Nelson, K. W. Kim, Y.-D. Chuang, Z. Hussain, R. G. Moore, T. P. Devereaux, W.-S. Lee, P. S. Kirchmann, and Z.-X. Shen, *Femtosecond electron-phonon lock-in by photoemission and x-ray free-electron laser*, Science **357**, 71–75 (2017).

Gla63. R. J. Glauber, *Coherent and Incoherent States of the Radiation Field*, Physical Review **131**, 2766–2788 (1963).

Gri10. S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, *A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu*, The Journal of Chemical Physics **132**, 154104 (2010).

Hal99. J. Hald, J. L. Sørensen, C. Schori, and E. S. Polzik, *Spin Squeezed Atoms: A Macroscopic Entangled Ensemble Created by Light*, Physical Review Letters **83**, 1319–1322 (1999).

Ham90. R. J. Hamers and D. G. Cahill, *Ultrafast time resolution in scanned probe microscopies*, Applied Physics Letters **57**, 2031–2033 (1990).

He14. K. He, N. Kumar, L. Zhao, Z. Wang, K. F. Mak, H. Zhao, and J. Shan, *Tightly bound excitons in monolayer WSe_2*, Physical Review Letters **113**, 1–5 (2014).

Heb02. J. Hebling, G. Almasi, I. Kozma, and J. Kuhl, *Velocity matching by pulse front tilting for large-area THz-pulse generation*, Optics Express **10**, 1161–1166 (2002).

Hei02. A. J. Heinrich, C. P. Lutz, J. A. Gupta, and D. M. Eigler, *Molecule cascades*, Science **298**, 1381–1387 (2002).

Hel21. S. Helmrich, K. Sampson, D. Huang, M. Selig, K. Hao, K. Tran, A. Achstein, C. Young, A. Knorr, E. Malic, U. Woggon, N. Owschimikow, and X. Li, *Phonon-Assisted Intervalley Scattering Determines Ultrafast Exciton Dynamics in* $MoSe_2$ *Bilayers*, Physical Review Letters **127**, 157403 (2021).

Her64. D. Herriott, H. Kogelnik, and R. Kompfner, *Off-Axis Paths in Spherical Mirror Interferometers*, Applied Optics **3**, 523–526 (1964).

Her65. D. R. Herriott and H. J. Schulte, *Folded Optical Delay Lines*, Applied Optics **4**, 883-889 (1965).

Her12. G. Herink, D. R. Solli, M. Gulde, and C. Ropers, *Field-driven photoemission from nanostructures quenches the quiver motion*, Nature **483**, 190–193 (2012).

Hir11. H. Hirori, A. Doi, F. Blanchard, and K. Tanaka, *Single-cycle terahertz pulses with amplitudes exceeding 1 MV/cm generated by optical rectification in* $LiNbO_3$, Applied Physics Letters **98**, 091106 (2011).

Hla00. S. W. Hla, L. Bartels, G. Meyer, and K. H. Rieder, *Inducing all steps of a chemical reaction with the scanning tunneling microscope tip: Towards single molecule engineering*, Physical Review Letters **85**, 2777–2780 (2000).

Hoh15. M. Hohenleutner, F. Langer, O. Schubert, M. Knorr, U. Huttner, S. W. Koch, M. Kira, and R. Huber, *Real-time observation of interfering crystal electrons in high-harmonic generation*, Nature **523**, 572–575 (2015).

Hub00a. R. Huber, A. Brodschelm, F. Tauser, and A. Leitenstorfer, *Generation and field-resolved detection of femtosecond electromagnetic pulses tunable up to 41 THz*, Applied Physics Letters **76**, 3191–3193 (2000).

Hub00b. R. Huber, *Diploma Thesis: Femtosekunden-lmpulse im Mittelinfrarot: Abstimmbare Erzeugung, feldaufgelöste Detektion und erste Anwendungen*, Technische Universität München, Fakultät für Physik, Physik-Department E11, 2000.

Hub16. M. A. Huber, F. Mooshammer, M. Plankl, L. Viti, F. Sandner, L. Z. Kastner, T. Frank, J. Fabian, M. S. Vitiello, T. L. Cocker, and R. Huber, *Femtosecond photo-switching of interface polaritons in black phosphorus heterostructures*, Nature Nanotechnology **12**, 207–211 (2016).

Jan94. J. Janszky, A. V. Vinogradov, T. Kobayashi, and Z. Kis, *Vibrational Schrödinger-cat states*, Physical Review A **50**, 1777–1784 (1994).

Jel17. V. Jelic, K. Iwaszczuk, P. H. Nguyen, C. Rathje, G. J. Hornig, H. M. Sharum, J. R. Hoffman, M. R. Freeman, and F. A. Hegmann, *Ultrafast terahertz control of extreme tunnel currents through single atoms on a silicon surface*, Nature Physics **13**, 591–597 (2017).

Jes19. R. Jestädt, M. Ruggenthaler, M. J. Oliveira, A. Rubio, and H. Appel, *Light-matter interactions within the Ehrenfest-Maxwell-Pauli-Kohn-Sham framework: fundamentals, implementation, and nano-optical applications*, Advances in Physics **68**, 225–333 (2019).

Joh95. W. M. Johnstone, P. Akther, and W. R. Newell, *Resonant vibrational excitation of carbon dioxide*, Journal of Physics B: Atomic, Molecular and Optical Physics **28**, 743–753 (1995).

Jun10. F. Junginger, A. Sell, O. Schubert, B. Mayer, D. Brida, M. Marangoni, G. Cerullo, A. Leitenstorfer, and R. Huber, *Single-cycle multiterahertz transients with peak fields above 10 MV/cm*, Optics Letters **35**, 2645–2647 (2010).

Kai00. R. A. Kaindl, M. Woerner, T. Elsaesser, D. C. Smith, J. F. Ryan, G. A. Farnan, M. P. McCurry, and D. Walmsley, *Ultrafast Mid-Infrared Response of* $YBa_2Cu_3O_{7-\delta}$, Science **287**, 470–474 (2000).

Kam13. T. Kampfrath, M. Battiato, P. Maldonado, G. Eilers, J. Nötzold, S. Mährlein, V. Zbarsky, F. Freimuth, Y. Mokrousov, S. Blügel, M. Wolf, I. Radu, P. M. Oppeneer, and M. Münzenberg, *Terahertz spin current pulses controlled by magnetic heterostructures*, Nature Nanotechnology **8**, 256–260 (2013).

Kan18. T. Kang, R. H. Y. Kim, G. Choi, J. Lee, H. Park, H. Jeon, C. H. Park, and D. S. Kim, *Terahertz rectification in ring-shaped quantum barriers*, Nature Communications **9**, 1–8 (2018).

Kar22. O. Karni, E. Barré, V. Pareek, J. D. Georgaras, M. K. Man, C. Sahoo, D. R. Bacon, X. Zhu, H. B. Ribeiro, A. L. O'Beirne, J. Hu, A. Al-Mahboob, M. M. Abdelrasoul, N. S. Chan, A. Karmakar, A. J. Winchester, B. Kim, K. Watanabe, T. Taniguchi, K. Barmak, J. Madéo, F. H. da Jornada, T. F. Heinz, and K. M. Dani, *Structure of the moiré exciton captured by imaging its electron and hole*, Nature **603**, 247–252 (2022).

Kas19. C. Kastl, R. J. Koch, C. T. Chen, J. Eichhorn, S. Ulstrup, A. Bostwick, C. Jozwiak, T. R. Kuykendall, N. J. Borys, F. M. Toma, S. Aloni, A. Weber-Bargioni, E. Rotenberg, and A. M. Schwartzberg, *Effects of Defects on Band Structure and Excitons in WS_2 Revealed by Nanoscale Photoemission Spectroscopy*, ACS Nano **13**, 1284–1291 (2019).

Kei15. B. Keimer, S. A. Kivelson, M. R. Norman, S. Uchida, and J. Zaanen, *From quantum matter to high-temperature superconductivity in copper oxides*, Nature **518**, 179–186 (2015).

Kei16. S. Keiber, S. Sederberg, A. Schwarz, M. Trubetskov, V. Pervak, F. Krausz, and N. Karpowicz, *Electro-optic sampling of near-infrared waveforms*, Nature Photonics **10**, 159–162 (2016).

Kel65. L. Keldysh, *Ionization in the field of strong electromagentic wave*, Soviet Physics JETP **20**, 1307–1314 (1965).

Kie15. D. Kienzler, H.-Y. Lo, B. Keitch, L. de Clercq, F. Leupold, F. Lindenfelser, M. Marinelli, V. Negnevitsky, and J. P. Home, *Quantum harmonic oscillator state synthesis by reservoir engineering*, Science **347**, 53–56 (2015).

Kip23. L. Kipczak, A. O. Slobodeniuk, T. Woźniak, M. Bhatnagar, N. Zawadzka, K. Olkowska-Pucko, M. Grzeszczyk, K. Watanabe, T. Taniguchi, A. Babiński, and M. R. Molas, *Analogy and dissimilarity of excitons in monolayer and bilayer of $MoSe_2$*, 2D Materials **10**, 025014 (2023).

Kno18. M. Knorr, P. Steinleitner, J. Raab, I. Gronwald, P. Merkl, C. Lange, and R. Huber, *Ultrabroadband etalon-free detection of infrared transients by van-der-Waals contacted sub-μm GaSe detectors*, Optics Express **26**, 19059–19066 (2018).

Kra09. F. Krausz and M. Ivanov, *Attosecond physics*, Reviews of Modern Physics **81**, 163–234 (2009).

Krü11. M. Krüger, M. Schenk, and P. Hommelhoff, *Attosecond control of electrons emitted from a nanoscale metal tip*, Nature **475**, 78–81 (2011).

Küb04. C. Kübler, R. Huber, S. Tübel, and A. Leitenstorfer, *Ultrabroadband detection of multi-terahertz field transients with GaSe electro-optic sensors: Approaching the near infrared*, Applied Physics Letters **3360**, 10–13 (2004).

Kub07. A. Kubo, N. Pontius, and H. Petek, *Femtosecond microscopy of surface plasmon polariton wave packet evolution at the silver/vacuum interface*, Nano Letters **7**, 470–475 (2007).

Kum12. A. Kumar and P. K. Ahluwalia, *Electronic structure of transition metal dichalcogenides monolayers 1H-MX_2 (M = Mo, W; X = S, Se, Te) from ab-initio theory: New direct band gap semiconductors*, European Physical Journal B **85**, 18–22 (2012).

Kun22. S. Kundu and N. Makri, *Intramolecular Vibrations in Excitation Energy Transfer: Insights from Real-Time Path Integral Calculations*, Annual review of physical chemistry **73**, 349–375 (2022).

Lae53. D. Laertius, *Lives and Opinions of Eminent Philosophers*, (Translated by C. D. Yonge), 1853.

Lan86. N. D. Lang, *Spectroscopy of single atoms in the scanning tunneling microscope*, Physical Review B **34**, 5947–5950 (1986).

Lan14. C. Lange, T. Maag, M. Hohenleutner, S. Baierl, O. Schubert, E. R. J. Edwards, D. Bougeard, G. Woltersdorf, and R. Huber, *Extremely Nonperturbative Nonlinearities in GaAs Driven by Atomically Strong Terahertz Fields in Gold Metamaterials*, Physical Review Letters **113**, 227401 (2014).

Lan18. F. Langer, C. P. Schmid, S. Schlauderer, M. Gmitra, J. Fabian, P. Nagler, C. Schüller, T. Korn, P. G. Hawkins, J. T. Steiner, U. Huttner, S. W. Koch, M. Kira, and R. Huber, *Lightwave valleytronics in a monolayer of tungsten diselenide*, Nature **557**, 76–80 (2018).

Lee19. J. Lee, K. T. Crampton, N. Tallarida, and V. A. Apkarian, *Visualizing vibrational normal modes of a single molecule with atomically confined light*, Nature **568**, 78–82 (2019).

Lei23. A. Leitenstorfer, A. S. Moskalenko, T. Kampfrath, J. Kono, E. Castro-Camus, K. Peng, N. Qureshi, D. Turchinovich, K. Tanaka, A. G. Markelz, M. Havenith, C. Hough, H. J. Joyce, W. J. Padilla, B. Zhou, K.-Y. Kim, X.-C. Zhang, P. U. Jepsen, S. Dhillon, M. Vitiello, E. Linfield, A. G. Davies, M. C. Hoffmann, R. Lewis, M. Tonouchi, P. Klarskov, T. S. Seifert, Y. A. Gerasimenko, D. Mihailovic, R. Huber, J. L. Boland, O. Mitrofanov, P. Dean, B. N. Ellison, P. G. Huggard, S. P. Rea, C. Walker, D. T. Leisawitz, J. R. Gao, C. Li, Q. Chen, G. Valušis, V. P. Wallace, E. Pickwell-MacPherson, X. Shang, J. Hesler, N. Ridler, C. C. Renaud, I. Kallfass, T. Nagatsuma, J. A. Zeitler, D. Arnone, M. B. Johnston, and J. Cunningham, *The 2023 terahertz science and technology roadmap*, Journal of Physics D: Applied Physics **56**, 223001 (2023).

Li17. S. Li, S. Chen, J. Li, R. Wu, and W. Ho, *Joint Space-Time Coherent Vibration Driven Conformational Transitions in a Single Molecule*, Physical Review Letters **119**, 1–5 (2017).

Li22. Y. Li, G. Kuang, Z. Jiao, L. Yao, and R. Duan, *Recent progress on the mechanical exfoliation of 2D transition metal dichalcogenides*, Materials Research Express **9**, 122001 (2022).

Liu12. M. Liu, H. Y. Hwang, H. Tao, A. C. Strikwerda, K. Fan, G. R. Keiser, A. J. Sternbach, K. G. West, S. Kittiwatanakul, J. Lu, S. A. Wolf, F. G. Omenetto, X. Zhang, K. A. Nelson, and R. D. Averitt, *Terahertz-field-induced insulator-to-metal transition in vanadium dioxide metamaterial*, Nature **487**, 345–348 (2012).

Lot10. S. Loth, M. Etzkorn, C. P. Lutz, D. M. Eigler, and A. J. Heinrich, *Measurement of fast electron spin relaxation times with atomic resolution*, Science **329**, 1628–1630 (2010).

Luo20. Y. Luo, V. Jelic, G. Chen, P. H. Nguyen, Y. J. R. Liu, J. A. Calzada, D. J. Mildenberger, and F. A. Hegmann, *Nanoscale terahertz STM imaging of a metal surface*, Physical Review B **102**, 205417 (2020).

Mad20. J. Madéo, M. K. L. Man, C. Sahoo, M. Campbell, V. Pareek, E. L. Wong, A. Al-Mahboob, N. S. Chan, A. Karmakar, B. M. K. Mariserla, X. Li, T. F. Heinz, T. Cao, and K. M. Dani, *Directly visualizing the momentum-forbidden dark excitons and their dynamics in atomically thin semiconductors*, Science **370**, 1199–1204 (2020).

Mak13. K. F. Mak, K. He, C. Lee, G. H. Lee, J. Hone, T. F. Heinz, and J. Shan, *Tightly bound trions in monolayer MoS_2*, Nature Materials **12**, 207–211 (2013).

Mak22. K. F. Mak and J. Shan, *Semiconductor moiré materials*, Nature Nanotechnology **17**, 686–695 (2022).

Man15. A. Manchon, H. C. Koo, J. Nitta, S. M. Frolov, and R. A. Duine, *New perspectives for Rashba spin-orbit coupling*, Nature Materials **14**, 871–882 (2015).

McE93. M. McEllistrem, G. Haase, D. Chen, and R. J. Hamers, *Electrostatic sample-tip interactions in the scanning tunneling microscope*, Phys. Rev. Lett. **70**, 2471–2474 (1993).

Mei22. C. Meineke, M. Prager, J. Hayes, Q. Wen, L. Z. Kastner, D. Schuh, K. Fritsch, O. Pronin, M. Stein, F. Schäfer, S. Chatterjee, M. Kira, R. Huber, and D. Bougeard, *Scalable high-repetition-rate sub-half-cycle terahertz pulses from spatially indirect interband transitions*, Light: Science and Applications **11**, 1–8 (2022).

Mer08. J. Merlein, M. Kahl, A. Zuschlag, A. Sell, A. Halm, J. Boneberg, P. Leiderer, A. Leitenstorfer, and R. Bratschitsch, *Nanomechanical control of an optical antenna*, Nature Photonics **2**, 230–233 (2008).

Mer19a. P. Merkl, F. Mooshammer, P. Steinleitner, A. Girnghuber, K. Q. Lin, P. Nagler, J. Holler, C. Schüller, J. M. Lupton, T. Korn, S. Ovesen, S. Brem, E. Malic, and R. Huber, *Ultrafast transition between exciton phases in van der Waals heterostructures*, Nature Materials **18**, 691–696 (2019).

Mer19b. P. Merkl, M. Knorr, C. Meineke, L. Kastner, D. Peller, and R. Huber, *Multibranch pulse synthesis and electro-optic detection of subcycle multiterahertz electric fields*, Optics Letters **44**, 5521–5524 (2019).

Mer20. P. Merkl, F. Mooshammer, S. Brem, A. Girnghuber, K.-Q. Lin, L. Weigl, M. Liebich, C.-K. Yong, R. Gillen, J. Maultzsch, J. M. Lupton, E. Malic, and R. Huber, *Twist-tailoring Coulomb correlations in van der Waals homobilayers*, Nature Communications **11**, 2167 (2020).

Mer21. P. Merkl, C. K. Yong, M. Liebich, I. Hofmeister, G. Berghäuser, E. Malic, and R. Huber, *Proximity control of interlayer exciton-phonon hybridization in van der Waals heterostructures*, Nature Communications **12**, 1–7 (2021).

Mit03. D. M. Mittleman, *Sensing with Terahertz Radiation*, vol. 85, Springer Series in Optical Sciences, Springer Berlin Heidelberg, 2003.

Mit21. E. Mitterreiter, B. Schuler, A. Micevic, D. Hernangómez-Pérez, K. Barthelmi, K. A. Cochrane, J. Kiemle, F. Sigger, J. Klein, E. Wong, E. S. Barnard, K. Watanabe, T. Taniguchi, M. Lorke, F. Jahnke, J. J. Finley, A. M. Schwartzberg, D. Y. Qiu, S. Refaely-Abramson, A. W. Holleitner, A. Weber-Bargioni, and C. Kastl, *The role of chalcogen vacancies for atomic defect emission in MoS_2*, Nature Communications **12**, 1–8 (2021).

Mog22. H. Mogi, Y. Arashida, R. Kikuchi, R. Mizuno, J. Wakabayashi, N. Wada, Y. Miyata, A. Taninaka, S. Yoshida, O. Takeuchi, and H. Shigekawa, *Ultrafast nanoscale exciton dynamics via laser-combined scanning tunneling microscopy in atomically thin materials*, npj 2D Materials and Applications **6:72**, 1–10 (2022).

Mou82. P. Moulton, *Ti-doped sapphire: tunable solid-state laser*, Optics News **8**, 9–9 (1982).

Mül20. M. Müller, N. Martín Sabanés, T. Kampfrath, and M. Wolf, *Phase-Resolved Detection of Ultrabroadband THz Pulses inside a Scanning Tunneling Microscope Junction*, ACS Photonics **7**, 2046–2055 (2020).

Na19. M. X. Na, A. K. Mills, F. Boschini, M. Michiardi, B. Nosarzewski, R. P. Day, E. Razzoli, A. Sheyerman, M. Schneider, G. Levy, S. Zhdanovich, T. P. Devereaux, A. F. Kemper, D. J. Jones, and A. Damascelli, *Direct determination of mode-projected electron-phonon coupling in the time domain*, Science **366**, 1231–1236 (2019).

Nai22. M. H. Naik, E. C. Regan, Z. Zhang, Y. H. Chan, Z. Li, D. Wang, Y. Yoon, C. S. Ong, W. Zhao, S. Zhao, M. I. B. Utama, B. Gao, X. Wei, M. Sayyad, K. Yumigeta, K. Watanabe, T. Taniguchi, S. Tongay, F. H. da Jornada, F. Wang, and S. G. Louie, *Intralayer charge-transfer moiré excitons in van der Waals superlattices*, Nature **609**, 52–57 (2022).

Nam21. M. Namboodiri, C. Luo, G. Indorf, T. Golz, I. Grguraš, J. H. Buss, M. Schulz, R. Riedel, M. J. Prandolini, and T. Laarmann, *Optical properties of Li-based nonlinear crystals for high power mid-IR OPCPA pumped at 1 µm under realistic operational conditions*, Optical Materials Express **11**, 231 (2021).

Nat17. F. D. Natterer, K. Yang, W. Paul, P. Willke, T. Choi, T. Greber, A. J. Heinrich, and C. P. Lutz, *Reading and writing single-atom magnets*, Nature **543**, 226–228 (2017).

Nee18. A. Neef, *Master's Thesis: Simulating Ultrashort Light Pulses in STM Tunnel Junctions*, Universität Regensburg, 2018.

Nie18. I. Niehues, R. Schmidt, M. Drüppel, P. Marauhn, D. Christiansen, M. Selig, G. Berghäuser, D. Wigger, R. Schneider, L. Braasch, R. Koch, A. Castellanos-Gomez, T. Kuhn, A. Knorr, E. Malic, M. Rohlfing, S. Michaelis de Vasconcellos, and R. Bratschitsch, *Strain Control of Exciton-Phonon Coupling in Atomically Thin Semiconductors*, Nano Lett. **18**, 1751–1757 (2018).

Nob86. *The Nobel Prize in Physics 1986*, The Royal Swedish Academy of Sciences, Press release, October 1986, https://www.nobelprize.org/prizes/physics/1986/press-release/.

Nob14. *The Nobel Prize in Chemistry 2014*, The Royal Swedish Academy of Sciences, Press release, September 2014, https://www.nobelprize.org/prizes/chemistry/2014/summary/.

Nob17. *The Nobel Prize in Chemistry 2017*, The Royal Swedish Academy of Sciences, Press release, October 2017, https://www.nobelprize.org/prizes/chemistry/2014/summary/.

Nob18. *The Nobel Prize in Physics 2018*, The Royal Swedish Academy of Sciences, Press release, October 2018, https://www.nobelprize.org/prizes/physics/2018/summary/.

Nun93. G. J. Nunes and M. Freeman, *Picosecond Resolution in Scanning Tunneling Microscopy*, Science **262**, 1029–1032 (1993).

Pan99. S. H. Pan, E. W. Hudson, and J. C. Davis, *^{3}He refrigerator based very low temperature scanning tunneling microscope*, Review of Scientific Instruments **70**, 1459–1463 (1999).

Par00. H. Park, J. Park, A. K. L. Lim, E. H. Anderson, A. P. Alivisatos, and P. L. McEuen, *Nanomechanical oscillations in a single-C_{60} transistor*, Nature **407**, 57–60 (2000).

Pat19. L. L. Patera, F. Queck, P. Scheuerer, N. Moll, and J. Repp, *Accessing a Charged Intermediate State Involved in the Excitation of Single Molecules*, Physical Review Letters **123**, 16001 (2019).

Pav17. N. Pavliček, A. Mistry, Z. Majzik, N. Moll, G. Meyer, D. J. Fox, and L. Gross, *Synthesis and characterization of triangulene*, Nature Nanotechnology **12**, 308–311 (2017).

Pel20a. D. Peller, *Dissertation: Femtosecond videography and control of single-molecule dynamics with atom-scale lightwaves*, Universität Regensburg, 2020.

Pel20b. D. Peller, L. Z. Kastner, T. Buchner, C. Roelcke, F. Albrecht, N. Moll, R. Huber, and J. Repp, *Sub-cycle atomic-scale forces coherently control a single-molecule switch*, Nature **585**, 58–62 (2020).

Pel21. D. Peller, C. Roelcke, L. Z. Kastner, T. Buchner, A. Neef, J. Hayes, F. Bonafé, D. Sidler, M. Ruggenthaler, A. Rubio, R. Huber, and J. Repp, *Quantitative sampling of atomic-scale electromagnetic waveforms*, Nature Photonics **15**, 143–147 (2021).

Pep17. M. Peplow, *The next big hit in molecule Hollywood*, Nature **544**, 408–410 (2017).

Per96. J. P. Perdew, K. Burke, and M. Ernzerhof, *Generalized gradient approximation made simple*, Physical Review Letters **77**, 3865–3868 (1996).

Per16. F. Perakis, L. De Marco, A. Shalit, F. Tang, Z. R. Kann, T. D. Kühne, R. Torre, M. Bonn, and Y. Nagata, *Vibrational Spectroscopy and Dynamics of Water*, Chem. Rev. **116**, 7590–7607 (2016).

Pla21. M. Plankl, P. E. Faria Junior, F. Mooshammer, T. Siday, M. Zizlsperger, F. Sandner, F. Schiegl, S. Maier, M. A. Huber, M. Gmitra, J. Fabian, J. L. Boland, T. L. Cocker, and R. Huber, *Subcycle contact-free nanoscopy of ultrafast interlayer transport in atomically thin heterostructures*, Nature Photonics **15**, 594–600 (2021).

RA18. S. Refaely-Abramson, D. Y. Qiu, S. G. Louie, and J. B. Neaton, *Defect-Induced Modification of Low-Lying Excitons and Valley Selectivity in Monolayer Transition Metal Dichalcogenides*, Physical Review Letters **121**, 167402 (2018).

Rác17. P. Rácz, Z. Pápa, I. Márton, J. Budai, P. Wróbel, T. Stefaniuk, C. Prietl, J. R. Krenn, and P. Dombi, *Measurement of Nanoplasmonic Field Enhancement with Ultrafast Photoemission*, Nano Letters **17**, 1181–1186 (2017).

Raf21. S. Rafiq, B. Fu, B. Kudisch, and G. D. Scholes, *Interplay of vibrational wavepackets during an ultrafast electron transfer reaction*, Nature Chemistry **13**, 70–76 (2021).

Rei18. J. Reimann, S. Schlauderer, C. P. Schmid, F. Langer, S. Baierl, K. A. Kokh, O. E. Tereshchenko, A. Kimura, C. Lange, J. Güdde, U. Höfer, and R. Huber, *Subcycle observation of lightwave-driven Dirac currents in a topological surface band*, Nature **562**, 396–400 (2018).

Rep05. J. Repp, G. Meyer, S. M. Stojković, A. Gourdon, and C. Joachim, *Molecules on insulating films: Scanning-tunneling microscopy imaging of individual molecular orbitals*, Physical Review Letters **94**, 1–4 (2005).

Rep06. J. Repp, G. Meyer, S. Paavilainen, F. E. Olsson, and M. Persson, *Imaging bond formation between a gold atom and pentacene on an insulating surface*, Science **312**, 1196–1199 (2006).

Ric94. A. Rice, Y. Jin, X. F. Ma, X. C. Zhang, D. Bliss, J. Larkin, and M. Alexander, *Terahertz optical rectification from* $\langle 110 \rangle$ *zinc-blende crystals*, Applied Physics Letters **64**, 1324–1326 (1994).

Rie00. E. Riedle, M. Beutter, S. Lochbrunner, J. Piel, S. Schenkl, S. Spörlein, and W. Zinth, *Generation of 10 to 50 fs pulses tunable through all of the visible and the NIR*, Applied Physics B **71**, 457–465 (2000).

Rom13. N. Romming, C. Hanneken, M. Menzel, J. E. Bickel, B. Wolter, K. von Bergmann, A. Kubetzka, and R. Wiesendanger, *Writing and Deleting Single Magnetic Skyrmions*, Science **341**, 636–639 (2013).

Rüc92. M. Rücker, W. Knoll, and J. P. Rabe, *Surface-plasmon-induced contrast in scanning tunneling microscopy*, Journal of Applied Physics **72**, 5027–5031 (1992).

Ryb16. T. Rybka, M. Ludwig, M. F. Schmalz, V. Knittel, D. Brida, and A. Leitenstorfer, *Sub-cycle optical phase control of nanotunnelling in the single-electron regime*, Nature Photonics **10**, 667–670 (2016).

Sar21. S. Sarkar and P. Kratzer, *Signatures of the Dichalcogenide-Gold Interaction in the Vibrational Spectra of MoS_2 and $MoSe_2$ on Au(111)*, Journal of Physical Chemistry C **125**, 26645–26651 (2021).

Sav12. K. J. Savage, M. M. Hawkeye, R. Esteban, A. G. Borisov, J. Aizpurua, and J. J. Baumberg, *Revealing the quantum regime in tunnelling plasmonics*, Nature **491**, 574–577 (2012).

Sch14. O. Schubert, M. Hohenleutner, F. Langer, B. Urbanek, C. Lange, U. Huttner, D. Golde, T. Meier, M. Kira, S. W. Koch, and R. Huber, *Sub-cycle control of terahertz high-harmonic generation by dynamical Bloch oscillations*, Nature Photonics **8**, 119–123 (2014).

Sch16. J. Schulte, T. Sartorius, J. Weitenberg, A. Vernaleken, and P. Russbüldt, *Nonlinear pulse compression in a multi-pass cell*, Optics Letters **41**, 2070 (2016).

Sch17a. G. D. Scholes, G. R. Fleming, L. X. Chen, A. Aspuru-Guzik, A. Buchleitner, D. F. Coker, G. S. Engel, R. Van Grondelle, A. Ishizaki, D. M. Jonas, J. S. Lundeen, J. K. McCusker, S. Mukamel, J. P. Ogilvie, A. Olaya-Castro, M. A. Ratner, F. C. Spano, K. B. Whaley, and X. Zhu, *Using coherence to enhance function in chemical and biophysical systems*, Nature **543**, 647–656 (2017).

Sch17b. Z. Schumacher, A. Spielhofer, Y. Miyahara, and P. Grutter, *The limit of time resolution in frequency modulation atomic force microscopy by a pump-probe approach*, Applied Physics Letters **110** (2017).

Sch19a. S. Schlauderer, C. Lange, S. Baierl, T. Ebnet, C. P. Schmid, D. C. Valovcin, A. K. Zvezdin, A. V. Kimel, R. V. Mikhaylovskiy, and R. Huber, *Temporal and spectral fingerprints of ultrafast all-coherent spin switching*, Nature **569**, 383–387 (2019).

Sch19b. B. Schuler, D. Y. Qiu, S. Refaely-Abramson, C. Kastl, C. T. Chen, S. Barja, R. J. Koch, D. F. Ogletree, S. Aloni, A. M. Schwartzberg, J. B. Neaton, S. G. Louie, and A. Weber-Bargioni, *Large Spin-Orbit Splitting of Deep In-Gap Defect States of Engineered Sulfur Vacancies in Monolayer WS_2*, Physical Review Letters **123**, 76801 (2019).

Sch22. D. Schmitt, J. P. Bange, W. Bennecke, A. A. AlMutairi, G. Meneghini, K. Watanabe, T. Taniguchi, D. Steil, D. R. Luke, R. T. Weitz, S. Steil, G. S. Jansen, S. Brem, E. Malic, S. Hofmann, M. Reutzel, and S. Mathias, *Formation of moiré interlayer excitons in space and time*, Nature **608**, 499–503 (2022).

Sed20. S. Sederberg, D. Zimin, S. Keiber, F. Siegrist, M. S. Wismer, V. S. Yakovlev, I. Floss, C. Lemell, J. Burgdörfer, M. Schultze, F. Krausz, and N. Karpowicz, *Attosecond optoelectronic field measurement in solids*, Nature Communications **11**, 1–8 (2020).

Sei16. T. Seifert, S. Jaiswal, U. Martens, J. Hannegan, L. Braun, P. Maldonado, F. Freimuth, A. Kronenberg, J. Henrizi, I. Radu, E. Beaurepaire, Y. Mokrousov, P. M. Oppeneer, M. Jourdan, G. Jakob, D. Turchinovich, L. M. Hayden, M. Wolf, M. Münzenberg, M. Kläui, T. Kampfrath, S. Jaiswal, P. M. Oppeneer, I. Radu, A. Kronenberg, L. Braun, F. Freimuth, U. Martens, J. Henrizi, T. Kampfrath, M. Münzenberg, Y. Mokrousov, L. M. Hayden, M. Kläui, T. Seifert, G. Jakob, J. Hannegan, and E. Beaurepaire, *Efficient metallic spintronic emitters of ultrabroadband terahertz radiation*, Nature Photonics **10**, 483–488 (2016).

Sei22. M. Seidel, P. Balla, C. Li, G. Arisholm, L. Winkelmann, I. Hartl, and C. M. Heyl, *Factor 30 Pulse Compression by Hybrid Multipass Multiplate Spectral Broadening*, Ultrafast Science **2022** (2022).

Sel08. A. Sell, R. Scheu, A. Leitenstorfer, and R. Huber, *Field-resolved detection of phase-locked infrared transients from a compact Er:fiber system tunable between 55 and 107 THz*, Applied Physics Letters **93** (2008).

Seo09. M. A. Seo, H. R. Park, S. M. Koo, D. J. Park, J. H. Kang, O. K. Suwal, S. S. Choi, P. C. Planken, G. S. Park, N. K. Park, Q. H. Park, and D. S. Kim, *Terahertz field enhancement by a metallic nano slit operating beyond the skin-depth limit*, Nature Photonics **3**, 152–156 (2009).

She22. S. Sheng, A. C. Oeter, M. Abdo, K. Lichtenberg, M. Hentschel, and S. Loth, *Launching Coherent Acoustic Phonon Wave Packets with Local Femtosecond Coulomb Forces*, Physical Review Letters **129**, 43001 (2022).

Sin20. A. Singh, A. Pashkin, S. Winnerl, M. Welsch, C. Beckh, P. Sulzer, A. Leitenstorfer, M. Helm, and H. Schneider, *Up to 70 THz bandwidth from an implanted Ge photoconductive antenna excited by a femtosecond Er:fibre laser*, Light: Science & Applications **9**, 30 (2020).

Sør14. S. G. Sørensen, H. G. Füchtbauer, A. K. Tuxen, A. S. Walton, and J. V. Lauritsen, *Structure and electronic properties of in situ synthesized single-layer MoS_2 on a gold surface*, ACS Nano **8**, 6788–6796 (2014).

Spe92. M. Specht, J. D. Pedarnig, W. M. Heckl, and T. W. Hänsch, *Scanning plasmon near-field microscope*, Phys. Rev. Lett. **68**, 476–479 (1992).

Ste14. W. Steurer, L. Gross, and G. Meyer, *Local thickness determination of thin insulator films via localized states*, Applied Physics Letters **231606** (2014).

Ste17. P. Steinleitner, P. Merkl, P. Nagler, J. Mornhinweg, C. Schüller, T. Korn, A. Chernikov, and R. Huber, *Direct Observation of Ultrafast Exciton Formation in a Monolayer of WSe_2*, Nano Letters **17**, 1455–1460 (2017).

Sti11. J. Stigler, F. Ziegler, A. Gieseke, J. C. M. Gebhardt, and M. Rief, *The complex folding network of single calmodulin molecules*, Science **334**, 512–516 (2011).

Sti21. F. Stilp, A. Bereczuk, J. Berwanger, N. Mundigl, K. Richter, and F. J. Giessibl, *Very weak bonds to artificial atoms formed by quantum corrals*, Science **372**, 1196–1200 (2021).

Str04. J. A. Stroscio and R. J. Celotta, *Controlling the Dynamics of a Single Atom in Lateral Atom Manipulation*, Science, 242–248 (2004).

Ter83. J. Tersoff and D. R. Hamann, *Theory and Application for the Scanning Tunneling Microscope*, Physical Review Letters **50**, 1998–2001 (1983).

Ter85. J. Tersoff and D. R. Hamann, *Theory of the scanning tunneling microscope*, Physical Review B **31**, 805–813 (1985).

Tho01. J. J. Thomson, *About bodies smaller than atoms*, The popular science montly (1901).

Tog15. A. Togo and I. Tanaka, *First principles phonon calculations in materials science*, Scripta Materialia **108**, 1–5 (2015).

Uge14. M. M. Ugeda, A. J. Bradley, S. F. Shi, F. H. Da Jornada, Y. Zhang, D. Y. Qiu, W. Ruan, S. K. Mo, Z. Hussain, Z. X. Shen, F. Wang, S. G. Louie, and M. F. Crommie, *Giant bandgap renormalization and excitonic effects in a monolayer transition metal dichalcogenide semiconductor*, Nature Materials **13**, 1091–1095 (2014).

Vel18. M. Velický, G. E. Donnelly, W. R. Hendren, S. McFarland, D. Scullion, W. J. I. DeBenedetti, G. C. Correa, Y. Han, A. J. Wain, M. A. Hines, D. A. Muller, K. S. Novoselov, H. D. Abruna, R. M. Bowman, E. J. G. Santos, and F. Huang, *Mechanism of Gold-Assisted Exfoliation of Centimeter-Sized Transition-Metal Dichalcogenide Monolayers*, ACS Nano **12**, 10463–10472 (2018).

vSL12. E. M. van Schrojenstein Lantman, T. Deckert-Gaudig, A. J. G. Mank, V. Deckert, and B. M. Weckhuysen, *Catalytic processes monitored at the nanoscale with tip-enhanced Raman spectroscopy*, Nature nanotechnology **7**, 583–586 (2012).

Wal63. C. T. Walker and R. O. Pohl, *Phonon Scattering by Point Defects*, Phys. Rev. **131**, 1433–1442 (1963).

Wal21. R. Wallauer, M. Raths, K. Stallberg, L. Münster, D. Brandstetter, X. Yang, J. Güdde, P. Puschnig, S. Soubatch, C. Kumpf, F. C. Bocquet, F. S. Tautz, and U. Höfer, *Tracing orbital images on ultrafast time scales*, Science **371**, 1056–1059 (2021).

Wan12. Q. H. Wang, K. Kalantar-Zadeh, A. Kis, J. N. Coleman, and M. S. Strano, *Electronics and optoelectronics of two-dimensional transition metal dichalcogenides*, Nature Nanotechnology **7**, 699–712 (2012).

Wan18. G. Wang, A. Chernikov, M. M. Glazov, T. F. Heinz, X. Marie, T. Amand, and B. Urbaszek, *Colloquium: Excitons in atomically thin transition metal dichalcogenides*, Reviews of Modern Physics **90**, 21001 (2018).

Wei93. S. Weiss, D. F. Ogletree, D. Botkin, M. Salmeron, and D. S. Chemla, *Ultrafast scanning probe microscopy*, Applied Physics Letters **63**, 2567–2569 (1993).

Wei17a. J. Weitenberg, T. Saule, J. Schulte, and P. Rußbüldt, *Nonlinear Pulse Compression to Sub-40 fs at 4.5 µJ Pulse Energy by Multi-Pass-Cell Spectral Broadening*, IEEE Journal of Quantum Electronics **53** (2017).

Wei17b. J. Weitenberg, A. Vernaleken, J. Schulte, A. Ozawa, T. Sartorius, V. Pervak, H.-D. Hoffmann, T. Udem, P. Russbüldt, and T. W. Hänsch, *Multi-pass-cell-based nonlinear pulse compression to 115 fs at 75 µJ pulse energy and 300 W average power*, Optics Express **25**, 20502 (2017).

Wil69. J. A. Wilson and A. D. Yoffe, *The transition metal dichalcogenides discussion and interpretation of the observed optical, electrical and structural properties*, Advances in Physics **18**, 193–335 (1969).

Wim17. L. Wimmer, O. Karnbach, G. Herink, and C. Ropers, *Phase space manipulation of free-electron pulses from metal nanotips using combined terahertz near fields and external biasing*, Physical Review B **95**, 1–13 (2017).

Wol15. E. E. Wollman, C. U. Lei, A. J. Weinstein, J. Suh, A. Kronwald, F. Marquardt, A. A. Clerk, and K. C. Schwab, *Quantum squeezing of motion in a mechanical resonator*, Science **349**, 952–955 (2015).

Xu14. X. Xu, W. Yao, D. Xiao, and T. F. Heinz, *Spin and pseudospins in layered transition metal dichalcogenides*, Nature Physics **10**, 343–350 (2014).

Yan11. H. Yanagisawa, M. Hengsberger, D. Leuenberger, M. Klöckner, C. Hafner, T. Greber, and J. Osterwalder, *Energy distribution curves of ultrafast laser-induced field emission and their implications for electron dynamics*, Physical Review Letters **107**, 1–5 (2011).

Ye19. J. Ye, Y. An, H. Yan, and J. Liu, *Defects and strain engineering the electronic structure and magnetic properties of monolayer WSe_2 for 2D spintronic device*, Applied Surface Science **497**, 143788 (2019).

Yos15. K. Yoshida, K. Shibata, and K. Hirakawa, *Terahertz Field Enhancement and Photon-Assisted Tunneling in Single-Molecule Transistors*, Physical Review Letters **115**, 1–5 (2015).

Yos16. K. Yoshioka, I. Katayama, Y. Minami, M. Kitajima, S. Yoshida, H. Shigekawa, and J. Takeda, *Real-space coherent manipulation of electrons in a single tunnel junction by single-cycle terahertz electric fields*, Nature Photonics **10**, 762–765 (2016).

Yos18. K. Yoshioka, I. Katayama, Y. Arashida, A. Ban, Y. Kawada, K. Konishi, H. Takahashi, and J. Takeda, *Tailoring Single-Cycle Near Field in a Tunnel Junction with Carrier-Envelope Phase-Controlled Terahertz Electric Fields*, Nano Letters **18**, 5198–5204 (2018).

Yos19. S. Yoshida, H. Hirori, T. Tachizaki, K. Yoshioka, Y. Arashida, Z. H. Wang, Y. Sanari, O. Takeuchi, Y. Kanemitsu, and H. Shigekawa, *Subcycle Transient Scanning Tunneling Spectroscopy with Visualization of Enhanced Terahertz Near Field*, ACS Photonics **6**, 1356–1364 (2019).

Yos21. S. Yoshida, Y. Arashida, H. Hirori, T. Tachizaki, A. Taninaka, H. Ueno, O. Takeuchi, and H. Shigekawa, *Terahertz Scanning Tunneling Microscopy for Visualizing Ultrafast Electron Motion in Nanoscale Potential Variations*, ACS Photonics **8**, 315–323 (2021).

Yur90. B. Yurke and G. P. Kochanski, *Momentum noise in vacuum tunneling transducers*, Phys. Rev. B **41**, 8184–8194 (1990).

Zha12. Y. Zhang, Y. Cheng, Q. Pei, C. Wang, and Y. Xiang, *Thermal conductivity of defective graphene*, Physics Letters A **376**, 3668–3672 (2012).

Zha13. W. Zhao, Z. Ghorannevis, L. Chu, M. Toh, C. Kloc, P.-H. Tan, and G. Eda, *Evolution of Electronic Structure in Atomically Thin Sheets of WS_2 and WSe_2*, ACS Nano, 791–797 (2013).

Zha16. Z. Zhang, S. Sheng, R. Wang, and M. Sun, *Tip-Enhanced Raman Spectroscopy*, Analytical Chemistry **88**, 9328–9346 (2016).

Zho04. Z. Zhou, K. Kato, T. Komaki, M. Yoshino, H. Yukawa, M. Morinaga, and K. Morita, *Effects of dopants and hydrogen on the electrical conductivity of ZnO*, Journal of the European Ceramic Society **24**, 139–146 (2004).

Zho16. H. Zhong, R. Quhe, Y. Wang, Z. Ni, M. Ye, Z. Song, Y. Pan, J. Yang, L. Yang, M. Lei, J. Shi, and J. Lu, *Interfacial Properties of Monolayer and Bilayer MoS_2 Contacts with Metals: Beyond the Energy Band Calculations*, Scientific reports **6**, 21786 (2016).

Zho21. Y. Zhou, J. Sung, E. Brutschea, I. Esterlis, Y. Wang, G. Scuri, R. J. Gelly, H. Heo, T. Taniguchi, K. Watanabe, G. Zaránd, M. D. Lukin, P. Kim, E. Demler, and H. Park, *Bilayer Wigner crystals in a transition metal dichalcogenide heterostructure*, Nature **595**, 48–52 (2021).

Acknowledgements

The past years as a PhD student have been truly memorable, thanks to the amazing colleagues and lab companions who made the journey so enjoyable. Working at the interface of ultrafast photonics and scanning probe microscopy, I am incredibly grateful to have had an extraordinary duo of mentors in Prof. Dr. Rupert Huber and Prof. Dr. Jascha Repp, guiding and inspiring me.

I would like to express my sincerest gratitude to Rupert for supporting me in my projects in all respects – be it by sharing his profound knowledge in stimulating discussions or by providing helpful guidance in the lab. Having enjoyed the exciting challenge and huge opportunity to learn, I am immensely thankful for the trust Rupert put in me with the ambitious MIR-STM project. I earnestly look up to his wholehearted excitement for our research, which deeply inspired me in my daily work. Thank you very much for your unwavering encouragement especially during challenging times and for enabling invaluable experiences such as conference trips, which greatly benefited both my scientific and personal growth.

I would like to extend my deepest thankfulness to Jascha for generously sharing his extensive expertise, teaching me the captivating physics and implementation of scanning tunnelling microscopy. Your precise and equitable approach to science has left a lasting impression on me. Thank you very much for your unhesitating and constructive assistance with any lab issues, which greatly contributed to my enthusiasm for nanoscopy. The fruitful exchanges throughout my project have been truly invaluable.

Moreover, I want to convey my genuine thanks to the board of examiners includ-

ing Prof. Dr. Klaus Richter, Prof. Dr. Rupert Huber, Prof. Dr. Jascha Repp, and Prof. Dr. Christian Schüller. Many thanks for taking the time to assess my work and conduct my oral examination.
In addition, I am very grateful to Dr. Dominik Peller for his invaluable guidance in navigating the complexities of operating the LW-STM setup and in applying scientific methods in general. The discussions during evening walks and bike rides were incredibly helpful throughout the ups and downs of my scientific pursuits.
Furthermore, I would like to thank Carmen Roelcke for the indispensable devoted teamwork, for always being ready to address issues, and becoming a true confidant and friend throughout the years. Our constructive discussions and your unwavering support have made our collaboration even more meaningful and fulfilling.
Many thanks, Dr. Yaroslav Gerasimenko, for enriching the office and lab with your invaluable knowledge on both physics-related topics and pearls of wisdom for many situations in life. Thank you for always finding the time for insightful discussions and sharing your snacks.
A big thank you goes to Christian Meineke and Johannes Hayes for the fruitful exchange on our optical setups, sharing ideas and issues we have encountered as well as recommendations for optimisation. I will always cherish the memories of the bike rides we have shared and am looking forward to many more to come.
Our collaboration with the chair of Prof. Dr. Jascha Repp has provided me with the opportunity to work closely with experts in scanning probe microscopy. I am deeply grateful to Dr. Thomas Buchner, Raffael Spachtholz, and Dr. Tzu-Chao Hung for their valuable insights, assistance, and the enjoyable moments we shared. Thank you in particular to Raffael for the cooperative efforts on the new setup. May it provide you with numerous Nature publications!
Additionally, I am truly appreciative of Johannes Hayes, Peter Menden, and Samuel Balonier for your support in setting up the NOPA. Your help has greatly aided in completing the MIR-STM setup prior to the completion of my PhD.
I extend my deepest gratitude to Dr. Markus Huber, who provided a captivating introduction to ultrafast nanoscopy during my Bachelor's degree. Your inspiration and encouragement have played a pivotal role in my decision to pursue further research in this dynamic and groundbreaking field of physics.
Throughout the years, numerous students have joined us in the lab and made sig-

nificant contributions to our research. I want to extend my appreciation to Andreas Biereder, Lukas Böhm, Jennifer Lehner, and Sarah Philips for their dedication and input.

Besides, I am particularly indebted to Ulla Franzke, Martin Furthmeier, Christoph Rohrer, Imke Gronwald, and Ignaz Läpple for their invaluable contributions. Their unwavering reliability and precision have played a crucial role in enabling my research.

I would also like to acknowledge all the individuals with whom I have had the privilege of collaborating, even though I have not mentioned them by name here. This encompasses all *Huberbuam:innen* & *STM-guys/gals*. Thank you for having been exceptional colleagues and for sharing your expertise and camaraderie. I am highly appreciative not only of the vibrant, friendly atmosphere at the workspace, but especially the enjoyable group events bringing us closer together.

Moreover, I owe a great debt of gratitude to my parents, my sister, and my girlfriend for their unshakable support and love. Their presence and unconditional backing have been invaluable throughout my journey as a student and young researcher.

Last but not least, I would like to acknowledge the contributions of Dr. Sabine Huber, Carmen Roelcke, Dr. Yaroslav Gerasimenko, and Johannes Hayes for proofreading my thesis. Their careful attention to detail and insightful suggestions have greatly improved the quality of my work.