Abstract | Research on 2D materials has been one of the fastest-growing fields in condensed matter and materials science research in the past 10 years. The low dimensionality and strong correlations of 2D systems give rise to electronic and structural properties, in the form of collective excitations, that do not have counterparts in ordinary 3D materials used in modern technology. These 2D materials present extraordinary opportunities for new technologies, such as in flexible electronics. In this Review, we focus on plasmons, excitons, phonons and magnons in 2D materials. We discuss the theoretical formalism of these collective excitations and elucidate how they differ from their 3D counterparts.

As a rule, the basis for the understanding of properties of materials relies on their non-interacting, or single-particle, properties. The ‘particle in a box’ model for 3D metals, proposed by Arnold Sommerfeld at the beginning of the twentieth century, is an example of how an oversimplified model of non-interacting electrons can provide explanations for physical behaviour and make predictions that have been confirmed experimentally innumerable times. It took many decades of study, and serious theoretical work, culminating in the Landau theory of the Fermi liquid, to understand why and how charged particles that interact through strong, long-range repulsive forces could behave as if they did not ‘feel’ each other. At the same time, Bloch’s theorem and the concept of band structure of crystals for non-interacting electrons have been at the core of virtually all modern technology based on semiconductors.

Driven either by pure scientific curiosity or by overwhelming technological demands, the study of new materials has been revolutionary, not evolutionary, in the past few decades. The isolation of graphene in 2004 and the emergence of a new class of 2D materials that became easily accessible experimentally has driven intense research in the past 16 years. It has become evident in the past decade that although several concepts in 3D solid-state physics still apply to 2D materials, the standard model of the theory of electrons in solids must be revised to explain the unique properties of 2D materials. We note that when referring to 2D materials, we mean crystalline systems in which one of their three dimensions is of atomic thickness, as opposed to quasi-2D systems, where the electrons are confined to a plane within a 3D material.

Electron interaction. The concepts of electric-field screening and non-interacting quasiparticles lie at the heart of the theory of 3D metals. In vacuum, two charged particles, such as electrons, interact with each other through the Coulomb potential, which decays as $e^2/r$, where $e$ is the charge of the electron and $r$ is the distance between the particles. The electric field generated by these charges permeates the whole space around them and creates a substantial repulsive force that keeps the charges apart. However, to quote Philip Anderson, “more is different”: when $10^{10}$ electrons in a material are forced to share the space of a 3D crystal, they can move around to shield, or screen, the strong 3D electric field. The final result of this electronic ‘dance’ is a much reduced interaction, $\exp \left( -r/l \right)/r$, where $l$ is a screening length that depends on the electronic density in the material. The new character that emerges from this scenario, the quasiparticle, can be seen as an electron ‘dressed’ by the cloud of all other electrons. Instead of a long-range interaction, the quasiparticles interact through weak short-range forces, validating the non-interacting model proposed by Sommerfeld. It is essential to note that this result comes about because electrons, and the electric field they carry with them, propagate in 3D.

In a 2D material, electrons are confined by the attractive potential of the lattice ions (with a positive charge) that halts their motion perpendicular to the plane of the material. Nevertheless, the electric field produced by the electrons still propagates in 3D space. Therefore, the electric field cannot be fully screened, because there are no free charges that can move around to shield the strong interaction between electrons. This lack of screening is one of the main characteristics of 2D materials.

Hence, 2D materials are strongly correlated by nature, and the concept of quasiparticles and single-particle excitations is debatable. Furthermore, collective electron modes in 2D metals such as plasmons become dispersive; that is, their energy becomes momentum-dependent, and their speed of propagation (or group velocity) diverges at long wavelengths and low energies. 2D plasmons cannot be observed experimentally in the same
way as their 3D counterparts, and new experimental techniques must be used.

In 2D semiconductors, this state of affairs is even more marked. In the 3D semiconductors used in modern-day technology, an electron in the conduction band and a hole in the valence band can bind in the form of a hydrogen-like atom called an exciton. These 3D excitons are very well described by the hydrogen model (H-model), which is one of the key theoretical achievements of quantum mechanics. In 2D, the situation is different. The lack of screening discussed previously invalidates many of the assumptions of the H-model. First, in 3D, only considering states close to the conduction- and valence-band edges is sufficient to describe the physics of exciton formation. The ‘effective mass’ approximation essentially retraces the Sommerfeld model of metals with non-interacting particles. In 2D semiconductors, the strong interactions require most of the states in the band structure to be accounted for to provide a precise description of the excitonic state and its binding energy. 2D excitons can be observed experimentally in photoluminescence (PL) experiments but cannot be calculated reliably from the H-model. Furthermore, because of the same strong interactions, many other excitations, such as biexcitons (bound state of two electron–hole pairs), trions (bound states of two electrons and one hole), quintons (bound states of three electrons and two holes) and so on, start to play an essential role in the interpretation of the experimental data.

**Structural properties.** Another essential characteristic of 2D materials is their inherent softness, in the sense that it costs much less energy to create collective states. Consider, for instance, collective lattice vibrations, or phonons. In 3D, an atom has its motion constrained by the presence of neighbouring atoms in all three directions (x,y,z). In 2D materials, one of the directions of motion, say z, is unconstrained. This leads to new phonon modes, ‘drum-head’ modes, whose energy, $E$, depends on the wavelength, $\lambda$, as $E \propto 1/\lambda^2$, becoming vanishingly small in energy at very long wavelengths. In thermal equilibrium, at a temperature $T$, the number of such modes becomes infinite, leading to structural instabilities such as crumpling, which is the basis of the Hohenberg–Mermin–Wagner theorem that states the absence of long-range order in 2D [Ref.1]. The same type of argument applies to 2D ferromagnetic systems where the collective magnetic states, or magnons, have the same dispersion relation. Therefore, 2D magnetic systems are prone to different types of magnetic collective states such as skyrmions and merons, increasing the zoo of collective states that live in 2D.

In this Review, we discuss various states in turn. We begin by addressing plasmons in 2D metals. Following this, we focus on excitonic states in 2D semiconductors. Next, we provide an overview of phonons and collective states in 2D magnets. We conclude with a perspective on future directions of the field.

**Plasmons**

Oscillations in electronic charge density give rise to plasmons. Even though, just like phonons, plasmonic modes can exist in systems of any dimensionality, 2D plasmons possess a set of characteristics that distinguish them from their 3D counterparts.

Subjecting an electronic system, regardless of its dimensionality, to an external perturbing potential $\phi_{ext}(r,t)$, where $r$ is a coordinate inside the material and $t$ is time, redistributes the charges and creates an induced electronic density $\rho_{ind}(r,t)$. This induced density, in turn, generates its own potential so that the total potential becomes

$$\phi_{tot}(r,t) = \phi_{ext}(r,t) + \int dr \rho_{ind}(r',t) V_{r-r'},$$

(1)

where $V_r$ is the Coulomb term, $r$ is the coordinate at which we want to know the total potential and $r'$ is the coordinate of the source density $\rho_{ind}(r',t)$. $V_{r-r'}$ is the Coulomb kernel. It is possible to relate $\rho_{ind}(r,t)$ to $\phi_{tot}(r,t)$ through the polarization function $\Pi(r,t)$ as

$$\rho_{ind}(r,t) = \int dr' \int dr \phi_{tot}(r',t') \Pi(r-r', t-t').$$

(2)

In the Fourier space, the convolutions in equations (1) and (2) are replaced by products so that

$$\phi_{tot}(q,\omega) = \phi_{ext}(q,\omega) + \Pi(q,\omega) V_q \phi_{tot}(q,\omega),$$

(3)

where $q$ is the wave vector and $\omega$ is the frequency of oscillations. The plasmonic modes are obtained from equation (3) by requiring that self-sustaining oscillations in $\phi_{tot}(q,\omega)$ exist in the absence of the external perturbation. In other words, $\Pi(q,\omega)V_q = 1$, which gives the relationship between the frequency of the oscillations $\omega$ and the corresponding wave vector $q$.

The polarization function can be computed from the Lindhard formula[17]. For a single-band parabolic dispersion $E_k$, the small-$q$ polarization is $\Pi(q,\omega) = m\omega^2/(\pi n_e)(\pi n_e)$ and $\Pi(q,\omega) = n_e n_e^*/(\pi n_e)$, where $\mu$ is the chemical potential, $n_e$ is the density of states at the Fermi momentum and $\gamma = \mu q^2/(2m^*)$ is the free electron dispersion with effective mass $m^*$. By comparing the 2D and 3D polarization functions, one can see that they both are proportional to $q^2/\alpha$. Even though these results were obtained for a massive isotropic dispersion, they also hold for Dirac11-13 and anisotropic bands14-16.

The key difference between the dimensionalities is in the Coulomb term: $V_{2D} \propto q^{-1}$, whereas $V_{3D} \propto q^{-2}$.
From this, $\Pi(q, \omega) V_q = 1$ states that, in 3D, the plasma frequency $\omega_p$ is a momentum-independent quantity. For real 3D materials, the magnitude of $\hbar \omega_p$ is of the order of electronvolts, making plasmons high-energy, weakly dispersive excitations. Although $\omega_p$ does depend on the carrier density, this quantity cannot be changed appreciably in a 3D sample so $\omega_p$ cannot be tuned experimentally. This large $\omega_p$ value means that substantial energy input is necessary to excite the charge oscillations in 3D.

As a consequence, one of the ways that the plasmons in 3D are confirmed experimentally is by sending energetic electrons through thin metallic films to observe that the energy lost by the electrons occurs in multiples of $\hbar \omega_p$ (REF 15).

The situation in 2D is very different with $\omega_p \propto \sqrt{q}$ (REF 16), meaning that the energy of the excitation $\hbar \omega$ can be made arbitrarily small by increasing the wavelength of the charge oscillations, resulting in 'soft modes.' A soft mode means that the energy of the excitation vanishes in the long-wavelength limit. This energy dependence is radically different from the 3D plasmon, where, even as $q \rightarrow 0$, $\hbar \omega_p$ remains finite. Also, unlike the 3D case, it is possible to change the charge density in a sample by gating, allowing one to tune the wavelength of the plasmonic modes.

Another important consequence of the $\sqrt{q}$ dispersion is to do with the plasmon–electron interaction. Because the speed of the plasmons varies with their momentum, it is possible to match this speed with the electronic Fermi velocity. This speed matching leads to a strong coupling between electrons and the collective modes, giving rise to plasmarons\(^\ref{16}\), which have been observed experimentally\(^\ref{17}\).

In a many-body electronic system, individual particle states can be labelled by their momentum $k$. The polarization function describes a process of transferring a particle from a filled state to an empty one. Crucially, the state in which the system ends up after such a transfer is not an eigenstate of the many-body Hamiltonian, being instead a superposition of many eigenstates. Because of this, the freshly moved particle will not remain in its new state $k'$ as the individual many-body states evolve in time. Another way to describe what happens is that the transferred particle scatters off other electrons. The Lindhard polarization function neglects this effect and assumes that the transferred particle stays in its new state. Although this might appear to be a serious shortcoming, we will see below that it works well for the systems of interest.

In addition to their tunability, 2D plasmons have the benefit of being observable in real space. To excite these modes, one needs to provide an external perturbation to a 2D system of the correct $\omega$ and $q$, determined by the solution of $\Pi_{2D}(q, \omega) V_q = 1$. It would seem reasonable that illuminating the system with the light of the desired frequency would excite the specific plasmon mode. The issue with this approach is that the wavelength of light with frequency $\omega$ is different from the wavelength of the plasma oscillations at the same $\omega$. Therefore, simply shining the light on the sample will not create plasmons because the momenta do not match.

To solve the momentum-mismatch problem, experimental groups use scanning near-field optical microscopy (REFS 18–22). In this technique, the electronic system is scanned by the tip of an atomic force microscope (AFM) which is being illuminated by a light beam of the desired frequency. The polarizable tip scatters the incident light and creates evanescent waves covering a range of different momenta, peaking at the inverse of the tip curvature radius, decoupling wavelength from frequency. A portion of these evanescent waves carries the correct momentum to excite the plasmons with frequency $\omega$.

The plasmons launched by the tip propagate radially and reflect off defects and edges. Continuously illuminating the system produces a non-equilibrium steady state with standing waves, and the tip detects the plasmonic signal at its location. If the reflected wave is in phase with the tip polarization, we get constructive interference at the tip position and observe an enhanced response. Similarly, if the reflected wave is out of phase with the tip, the signal will be diminished.

An experimental set-up used for plasmonic imaging is shown in FIG. 1. In FIG. 1a, one can see a schematic illustration of the device with the plasmon-hosting graphene encapsulated by hexagonal boron nitride (hBN), situated on top of the Si/SiO\(_2\) gate. Gold microstructures are deposited on top of the sample. FIGURE 1b shows an optical image of the sample.

By moving the illuminated AFM tip around, one can record the response throughout the sample to obtain images such as those shown in FIG. 1c. To reiterate: the data do not represent instantaneous snapshots. Instead, each point is a total signal arising from the tip perturbation plus the contribution due to reflected waves, explaining why the signal is essentially invariant as the tip moves along the edge of the sample.

Close to the edge, one can observe a series of fringes due to alternating constructive/destructive interference.

Fig. 1 | Experimental detection of plasmonic modes in graphene. a | A sketch showing the cross-section of the device. b | Optical image of a graphene sheet encapsulated by hexagonal boron nitride (hBN). c | Scanning near-field optical microscopy data collected at 60K showing plasmonic modes are launched by the tip and reflected by the sample edges and the gold microstructures. The arrows represent the propagation direction of the plasmon waves. Reprinted from REF\(^\ref{15}\), Springer Nature Limited.
As expected in an interference problem, the spacing between the fringes is one-half of the plasmonic wavelength, determined by the chemical potential and the light frequency. FIGURE 1c also demonstrates the ability of plasmons to ‘squeeze’ the electromagnetic field. The incident wave generated by an 886 cm⁻¹ laser has a wavelength of approximately 11.3 μm. The plasmon wavelength, however, is close to 200 nm while retaining the same frequency.

In the earliest work on real-space plasmon imaging⁴, it was observed that the amplitude of the signal oscillations gets weaker as the distance from the reflector (edge or defect) increases. Although one does expect this as a consequence of the circular wave spreading out, the experimentally observed decay is exponential rather than power-law, raising the question of what determines the plasmon lifetime.

Since plasmonic modes are composed of a periodically modulated electronic density, the decay of the modes implies that the constituent electrons stop moving in unison. The reason that an electron would change its trajectory is scattering, and, in the system of interest, there are three main scattering pathways: electron–electron, electron–phonon and electron–disorder. To understand which of these processes dominates the plasmonic decay, one should assess them individually. The situation with the electron–electron interaction is delicate, and we will discuss it shortly, after addressing the other two mechanisms.

To minimize the effects of the disorder, one can use ultraclean encapsulated samples. Lowering the temperature can also be used to freeze out the lattice vibrations. Both of these steps were taken in REF.⁵ (see FIG. 1). In BN-encapsulated graphene at cryogenic temperatures, plasmon oscillations retain their shape over multi-micrometre distances as the propagation length increases from a few wavelengths to 50. Increasing the temperature suppresses the plasmon propagation, indicating that, in clean samples, it is the electron–phonon scattering that plays the dominant role in plasmon decay at these frequencies.

The fact that the electron–phonon interaction plays the leading role in determining the plasmon lifetime does not mean that the electron–electron interaction is inconsequential. After all, it is the Coulomb repulsion that provides the restoring force for the charge oscillations. As is often the case in oscillating systems, it is all about the timescales. Electrons in a metal have a characteristic collision frequency that determines the time that it takes for a perturbation in a system to dissipate. If the perturbation is applied substantially more quickly than the electrons can equilibrate, one can neglect the equilibration process involving individual electrons and only consider the density–density interaction, which is at the heart of the Lindhard formalism. In other words, the external potential creates plasmons more quickly than they can decay through electron–electron scattering. As the frequency of the external perturbation is lowered, electron–electron interaction becomes relevant, and plasmons experience strong damping. To include this effect, we have to return to the polarization function and go beyond the Lindhard approximation.

**Excitons**

A layered semiconductor can be thinned down to a monolayer, becoming, essentially, a 2D material, where charge carrier motion (if any) is heavily restricted along the out-of-plane direction. The electrostatic field, however, remains 3D because the field lines can extend beyond the semiconducting layer, where the dielectric screening is limited. The resulting electrostatic interactions are therefore much stronger in 2D layers than in the 3D samples made of the same semiconducting material.

Optically excited electrons and holes [FIG. 2a], having opposite charge, experience Coulomb attraction. An electron–hole pair forming a bound state is known as...
an exciton (FIG. 2b). Excitonic features can be observed in absorbance, reflectance, photocurrent and PL spectra. In 2D semiconductors, however, the interactions are so strong that it is possible to resolve additional spectral features corresponding to much more complex electron–hole formations.

The physics of optically excited electron–hole complexes (which may comprise any number of electrons and holes) can be understood by using the same quasiparticle concept as for excitons. Each quasiparticle species can be characterized by properties such as electric charge, mass and binding energy. We start from the theoretically simplest case, the excitonic Hamiltonian given by

$$H_x = \frac{P_x^2}{2m_x} + \frac{P_y^2}{2\mu_x} + V_x (r_x),$$

(4)

where $P_x = -i\hbar \partial / \partial r_x$, $P_y = -i\hbar \partial / \partial r_y$ are momenta operators for the centre-of-mass and relative electron–hole motion, respectively, and $V_x (r_x)$ is the electron–hole potential. The relative and centre-of-mass coordinates are defined in a 2D plane as $r_y = r_y - r_x$, $R_x = (m_x r_x + m_y r_y) / m_{xy}$, where $m_{xy}$ and $r_{xy}$ are the electron (hole) masses and coordinates, respectively. The former determine the total mass $m_t = m_e + m_h$ as well as the reduced mass $\mu_x = m_e m_h / m_t$.

The simplest excitonic model is based on a hydrogen-like solution of the Schrödinger equation with the Coulomb potential $V_x (r_x) = -e^2 / (\epsilon r_x)$, where $\epsilon$ is the average relative dielectric permittivity of a semiconductor. The model works reasonably well in 3D semiconductors; however, it may fail in the 2D limit for several reasons.

The most important reason has been known for decades and has been revised recently. Once the thickness of a semiconductor layer becomes comparable to or smaller than the exciton size, the field lines connecting an electron and a hole extend beyond the semiconducting material, which reduces effective screening. Solving the Poisson equation for a 2D semiconductor with the in-plane dielectric polarizability $\chi$ and average dielectric permittivity of the environment $\bar{\epsilon}$, one can find the 2D Fourier transform of the electron–hole potential given by $V_x (q) = -2\pi e^2 / (\epsilon q)$, where $q = \epsilon (1 + r)\bar{q}$, and $r = 2\pi q / \epsilon$. The real-space potential can be expressed in terms of the Bessel and Struve functions. 2D excitons can be described by a hydrogen-like model as long as the excitonic ground-state radius is substantially larger than the screening length $r_s$, so that the quasiclassical radius falls onto the Coulomb-like tail of the potential. In the opposite limit, the electron–hole potential does not follow Coulombic $-1/r$ functional dependence, but turns out to be a much weaker function of the form $\ln (r/r_s)$. The binding energy is also a much weaker function of $\epsilon$ in this case, as compared with its hydrogenic version.

The less obvious reasons for the hydrogen-like model failure lie in limitations imposed by the effective mass approximation. First of all, the reduced electrostatic screening in the 2D limit makes the real-space exciton radius smaller but, at the same time, the exciton gets larger in the momentum space involving the states with higher momenta beyond the regions of parabolic band dispersion. This can easily happen for the valence band, which may have dispersion peculiarities not that far from its top in thin semiconducting layers. Hence, the tightly bound excitons may not always be described in terms of the electron and hole effective masses despite the real-space exciton size remaining larger than the lattice constant justifying the Wannier–Mott picture. In addition to the non-parabolicity of the bands, the effective mass approximation also does not take into account the Berry phase that can influence quantization of the relative electron–hole motion and change the exciton spectrum.

Hence, the excitons in 2D semiconductors are in general non-hydrogenic; however, the hydrogen-like model can still be used to understand the excitonic ground-state behaviour qualitatively. In this case, the Schrödinger equation can be solved explicitly, and the exciton binding energy reads $E_x = 2\mu_x e^2 / (\epsilon^2 \hbar^2)$. The binding energy $E_x$ is of the order of 100 meV in 2D semiconductors (for example, 0.32 eV in WS$_2$ (REF 44) and 0.37 eV in WSe$_2$ (REF 44)) when deposited on a SiO$_2$ substrate. Up to ~50% modulation of binding energy can be achieved by changing the dielectric substrate. Even greater modulation can be reached by making use of free-carrier screening in nearby graphene.

The screening in 2D semiconductors is so weak that even dipole–electron and dipole–dipole interactions turn out to be strong enough to form bound states between one exciton and one electron, resulting in a trion, between two excitons to form a biexciton, or even between one trion and one exciton to create a quinton. In what follows, we review the quasiparticle zoo (FIG. 2) on an equal theoretical footing.

**Trions.** Trion can be seen as a bound state between an electron (or a hole) and an exciton (FIG 2c). Trions are also known as charged excitons, which can be negative or positive depending on whether the excess charge is due to an electron or a hole. The majority of conventional 2D semiconductors are n-doped, so we focus on negatively charged trions here. (The sign of the trion charge could, however, be altered by electrostatic gating). The trionic binding energy is much lower than the excitonic energy because dipole–electron interactions are much weaker than electron–hole coupling. The formal description starts from the coordinate transformation that allows the relative and centre-of-mass exciton–electron motion to be separated. The trion Hamiltonian can be written as

$$H_T = \frac{P_x^2}{2m_x} + \frac{P_y^2}{2\mu_x} + V_x (r_x),$$

(5)

where $P_x = -i\hbar \partial / \partial R_x$, $P_y = -i\hbar \partial / \partial R_y$ are momenta operators for the centre-of-mass and relative electron–exciton motion, respectively, and $V_x (r_x)$ is the electron–exciton interaction described by the dipole potential. The relative and centre-of-mass coordinates are defined as $r_y = R_x - r_y$, $R_x = (m_x R_x + m_y r_y) / m_{xy}$, where $R_x$ is the outer electron coordinate. The total and reduced trion masses are $m_t = m_e + m_h$ and $\mu_x = m_e m_h / m_t$, respectively. Solving the Schrödinger equation with the Hamiltonian $H_T$,

Bessel and Struve functions
These are solutions of Bessel’s differential equation naturally occurring in problems with cylindrical symmetry.

Wannier–Mott picture
This is a model that assumes the exciton radius (or, in general, the mean electron–hole distance in exciton-like quasiparticles) to be much larger than the lattice constant, allowing for the effective mass description.
one can find the theoretical trion binding energy $E_p$, which is known to be of the order of 10 meV in 2D semiconductors\textsuperscript{46,52,53}. This is the energy we need to decompose a trion into an electron and an exciton.

Trions are in general less stable than excitons. Moreover, trionic decay may involve electron–hole recombination. In contrast to the electron–hole recombination in excitons, where the energy is released as a photon, the trionic recombination results in a free electron absorbing excess energy. The photon is then not emitted, which leads to a strong reduction in the PL quantum yield. Hence, electron–exciton scattering limits the performance of PL devices based on 2D semiconductors. The limitation can be lifted by appropriate chemical or electrostatic doping, which reduces excess electron concentration. Precluding trion formation and electron–exciton scattering is the way to reach nearly 100% PL efficiency\textsuperscript{45,54}. In the opposite regime of high excess electron concentrations, the two-body Hamiltonian (\(\Psi\)) becomes inapplicable owing to the truly many-body interactions involved\textsuperscript{46}. The resulting quasiparticle, known as a Fermi polaron, has been detected in 2D MoSe\textsubscript{2} (REF.\textsuperscript{55}).

**Bieexcitons.** Bieexcitons are quasiparticles created by dipole–dipole interaction between two excitons (FIG. 2d). The dipole–dipole potential is weaker than electron–dipole interactions, so the bieexcitonic binding energy is lower than the trionic one\textsuperscript{46,52,53}. Transforming coordinates in a similar way to before, we can write the bieexciton Hamiltonian as

$$H_{XX} = \frac{\mathbf{P}_{XX}^2}{2m_{XX}} + \frac{\mathbf{P}_{XX}^2}{2\mu_{XX}} + V_{XX}(\mathbf{r}_{XX}),$$  

where $\mathbf{P}_{XX} = -i\hbar \partial / \partial \mathbf{R}_{XX}$, $\mathbf{P}_{XX} = -i\hbar \partial / \partial \mathbf{r}_{XX}$ are momenta operators for the centre-of-mass and relative exciton motion, respectively, and $V_{XX}(\mathbf{r}_{XX})$ is the exciton–exciton interaction described by the dipole–dipole potential $\alpha r^{-3}$. The relative and centre-of-mass coordinates are respectively defined as $\mathbf{r}_{XX} = \mathbf{R}_{X} - \mathbf{R}_{X}$, $\mathbf{R}_{XX} = (\mathbf{R}_{X} + \mathbf{R}_{X})/2$, where $\mathbf{R}_{X}$ is the centre-of-mass coordinate for the second exciton, and $m_{XX} = 2m_{X}$ and $\mu_{XX} = m_{X}/2$.

The dipole–dipole potential vanishes rapidly when the distance between the two excitons, $r_{XX}$, is much larger than the exciton size itself, $r_{e}$. This hinders bieexciton formation, making it very challenging to detect the corresponding PL emission peak experimentally\textsuperscript{46}. Being a four-body process, the probability for exciton–exciton interactions increases quadratically with the excitonic density, and hence with the excitation power. Thus, parabolic dependence of PL intensity versus excitation power is a clear indication of bieexcitonic effects involved. Bieexcitonic scattering, off each other or off defects or edges, may lead to recombination of one exciton and ionization of another, a phenomenon known as exciton–exciton annihilation\textsuperscript{51}. This is a non-radiative process that strongly reduces the PL quantum yield at high excitation power\textsuperscript{51}. On the positive side, exciton–exciton annihilation produces high-energy free-electron–hole pairs that may be used in optically excited transport to increase internal quantum efficiency\textsuperscript{51}.

**Quintons.** A negatively charged quinton is a quasiparticle composed of three electrons and two holes (FIG. 2e). It can be described in two ways. First, a quinton can be seen as an electron bound to the quadruple core created by a biexciton\textsuperscript{58}. Second, it can be seen as a bound state between a trion and an exciton. Note, however, that the most stable state should have the lowest potential energy. The quadruple potential is weaker than the dipole one; hence, the trion–exciton configuration should have a lower ground state than the electron–biexciton one. The quinton Hamiltonian can be written as

$$H_{QT} = \frac{\mathbf{P}_{QT}^2}{2m_{QT}} + \frac{\mathbf{P}_{QT}^2}{2\mu_{QT}} + V_{QT}(\mathbf{r}_{QT}),$$  

where $\mathbf{P}_{QT} = -i\hbar \partial / \partial \mathbf{R}_{QT}$, $\mathbf{P}_{QT} = -i\hbar \partial / \partial \mathbf{r}_{QT}$ are momenta operators for the centre-of-mass and relative exciton–trion motion, respectively, and the exciton–trion interaction $V_{QT}(\mathbf{r}_{QT})$ is given by potential between a point-like charge and a dipole representing a trion and an exciton, respectively. The relative and centre-of-mass coordinates are defined as $\mathbf{r}_{QT} = \mathbf{R}_{X} - \mathbf{R}_{T}$, $\mathbf{R}_{QT} = (m_{X}\mathbf{R}_{X} + m_{T}\mathbf{R}_{T})/m_{XT}$ with the total and reduced quinton masses being $m_{XT} = m_{X} + m_{T}$ and $\mu_{XT} = m_{X}m_{T}/m_{XT}$, respectively.

Observation of quintons requires exceptional experimental conditions, including extreme purification of material components and cryogenic temperatures. The quasiparticle has been observed at 4 K in 2D WSe\textsubscript{2} encapsulated between two layers of hBN (REF.\textsuperscript{56}).

**Polaronic trions.** Interactions possible in 2D semiconductors are not limited to the Coulombic ones forming excitons and exciton-like quasiparticles reviewed above. Electrons can couple with phonons by means of Fröhlich interactions, forming polarons\textsuperscript{46}. Although 2D semiconductors possess intrinsic Fröhlich interactions\textsuperscript{46}, a much stronger polaronic coupling can be induced by an appropriate substrate, for example, by strontium titanium oxide (STO)\textsuperscript{56}. The outer electron in a trion and soft rotational optical (RO) phonon modes in the STO substrate both rotate with the frequencies equivalent to 10 meV matching each other’s symmetry and energy scale. The Fröhlich coupling should, therefore, be efficient for trions, hence creating a new quasiparticle — polaronic trion (FIG. 2f)\textsuperscript{57}. The polaronic trion Hamiltonian can be written as

$$H_{PT} = \frac{\mathbf{P}_{PT}^2}{2m_{PT}} + \frac{\mathbf{P}_{PT}^2}{2\mu_{PT}} + V_{PT}(\mathbf{r}_{PT}) + V_{P}(\mathbf{r}_{P}),$$  

where $\mathbf{P}_{PT}$, $\mathbf{P}_{P}$ and $V_{P}(\mathbf{r}_{P})$ have been defined below equation (\(\Psi\)). Here, we have introduced an additional interaction given by the polaronic term $V_{P}(\mathbf{r}_{P})$. The corresponding Fourier transform can be derived following the Fröhlich recipe\textsuperscript{46}, $V_{P}(q) = -i\hbar \omega / 2m_{P} \cos \beta / q$, where $r_{c} = \sqrt{2/\mu_{P}}$, $\omega$ is the interaction length, $\omega$ is the RO phonon frequency and $\alpha$ is the standard Fröhlich coupling constant\textsuperscript{46}. The striking difference between the standard 2D polaronic interaction and $V_{P}(q)$ is the $\sqrt{\cos \beta}$ pre-factor that occurs because of the special...
Moiré potential
A potential resulting from a superposition of two periodic potentials with slightly different lattice constants or misalignment, resulting in a pattern with a much larger periodicity.

Interlayer excitons. A pair of 2D semiconductors can be stacked together and excited optically. Because the Coulomb interactions are screened weakly, they may extend across the interface, providing coupling between electrons and holes excited in the neighbouring layers.

This forms interlayer excitons (fig. 2g). The interlayer exciton Hamiltonian can be written as

$$H_{\text{IX}} = \frac{p_x^2}{2m_X} + \frac{p_y^2}{2\mu_X} + V_{\text{IX}}(r_X),$$

where $p_x$ and $p_y$ have been introduced below equation (4), and $V_{\text{IX}}(r_X)$ is the interlayer electron–hole potential. The interlayer interaction is weaker than the intralayer one because of the constant interlayer distance $d$ always separating an electron and a hole. The interlayer exciton binding energy $E_{\text{IX}}$ is therefore expected to be lower than $E_X$. Note, however, that the interlayer bandgap $\Delta_i$ is strongly reduced because of the band offset between the two layers, so that the corresponding PL emission energy given by $\Delta_{\text{IX}} = \Delta_i - E_{\text{IX}}$ turns out to be redshifted\(^{29}\) with respect to the intralayer excitonic peak $\Delta_i$. The semiconducting layers can also be stacked with a twist that creates an additional moiré potential. This potential splits the single interlayer excitonic peak into several resonances, as observed in ref.\(^{71}\). The moiré excitons and exciton-like complexes are in the current focus of experimental\(^{72-74}\) and theoretical\(^{75-79}\) research.

The most important property of interlayer excitons is their extended recombination time due to the spatial separation of the electrons and holes\(^{29}\). This makes it possible to study strongly correlated electron–hole states such as exciton condensation\(^{79-80}\) and electron–hole liquids\(^{81}\). Recent papers on interlayer excitons have already been reviewed\(^{82}\).

Estimations and connections with experiments. To conclude the section, we evaluate the accuracy of the Hamiltonians introduced above when compared with data. We use the quasiparticle binding energies recently measured in high-quality 2D WSe\(_2\) (ref.\(^{84}\); $E_0 = 0.029\,\text{eV}$, $E_{\text{IX}} = 0.017\,\text{eV}$, $E_{\text{XT}} = 0.02\,\text{eV}$). The physical meaning of the energies is explained in fig. 3. A typical PL spectrum is schematically shown in fig. 4. Excitonic binding energy is difficult to figure out from PL spectroscopy data, as the fundamental electronic bandgap does not produce any PL peak. Using the band parameters for 2D WSe\(_2\), $m_r = 0.29m_0$, $m_\sigma = 0.34m_0$ (ref.\(^{85}\)) (here, $m_\sigma$ is free electron mass), and $\tau = 6.85$ ($\tau$ is the in-plane dielectric permittivity of hBN; ref.\(^{86}\)), we estimate $E_X$ as $2\mu_Xe^2/\left(\varepsilon\hbar^2\right) \sim 0.18\,\text{eV}$.

The quasiparticle size can then be estimated as a turning point of the relative quasiclassical motion in the corresponding potential at a given binding energy. For excitons, this results in $r_X = e^2/\left(\varepsilon E_X\right)$ being a little larger than 1 nm. The biexciton size can be estimated assuming the dipole–dipole potential that suggests $r_{\text{IX}} \sim \sqrt{e^2r_X^2/(\varepsilon E_{\text{IX}})}$ to be a little more than 2 nm. This is a reasonable value, as the biexciton comprises two excitons with 1 nm size each. Indeed, the biexciton size cannot be much larger than $r_X$, because the dipole–dipole potential is not able to bind at such distances. Moreover, a biexciton size smaller than $2r_X$ would be unreasonable because of electron–hole recombination annihilating the quasiparticle. Hence, the exciton and biexciton models are consistent.
Intrinsic features

| $\Delta_x$ | $\Delta_y$ | $\Delta_{xy}$ | $\Delta_a$ | $\Delta_x$ | $\Delta_y$ |
|-----------|------------|---------------|-----------|------------|------------|
| $E_{ax}$  | $E_{ay}$   | $E_{axy}$     | $E_a$     | $E_{ax}$   | $E_{ay}$   |

Extrinsic features

| $\Delta_a$ | $\Delta_y$ |
|------------|------------|
| $E_{ax}$   | $E_{ay}$   |

PL energy

**Fig. 4** | Relative positions of quasiparticle PL emission lines with the binding energies indicated. (See Figs 2 and 3 for notations.) Intrinsic features can be observed in isolated 2D semiconducting layers whereas extrinsic effects require either a special substrate or another 2D layer nearby. Presentation inspired by REF.11.

Following the same approach, we quasiclassically estimate the trion size from the electron–dipole potential as $r_T \approx \sqrt{r_x r_y / (\epsilon E_x)}$, resulting in about 3 nm. The quinton size can be then estimated in the same way as $r_{QT} \approx \sqrt{r_x^2 r_y^2 / (\epsilon E_{xx})}$, which is about 4 nm. Hence, the quinton size turns out to be equal to the sum of the trion and exciton characteristic lengths. This again confirms the feasibility of the framework.

We do not consider spin and valley degrees of freedom in our model Hamiltonians; however, the spin–valley locking intrinsic to optical excitations in 2D transition metal dichalcogenides may lead to circularly polarized excitonic PL emission, with the chirality depending on which valley is excited.85–88 Intervalley polarized excitonic PL emission, with the chirality intravalley and intervalley one91.

All the quasiparticles reviewed above can be seen as oscillators with the eigenfrequencies specific to each quasiparticle species. The large oscillator strength of exciton-like quasiparticles in 2D semiconductors makes it possible to strongly couple them with light in an optical cavity, creating exciton–polaritons. From the theoretical point of view, an exciton–polariton can be described as two strongly coupled, damped oscillators. The resulting eigenfrequency splits into two, and the Rabi splitting is observable in the emission spectra. The 2D exciton–polaritons were initially detected in 2D MoS$_2$ (REF.92) followed by similar observations in 2D WS$_2$ (REFS.93–95) and 2D MoSe$_2$ (REF.96). Interestingly, polaritons inherit the valley selectivity (spin–valley locking) from the parent excitons.97–99. Theoretically, not only excitons but any other of the quasiparticles listed above may couple with light in a cavity to form polaritons. This has already been experimentally confirmed for trions.100. Hence, the polariton effect potentially doubles the number of exciton-like quasiparticle species.

Note, however, that to make quantitative predictions we should consider more realistic potentials and use more exact solution techniques, such as the GW–Bethe–Salpeter equation101 or quantum Monte Carlo102,103.

Phonons

**Single layers.** In 2D space, a crystal with $N$ atoms per unit cell has $2N$ degrees of freedom, leading to the same number of vibrational modes (or phonons, in quasiparticle language). However, since they exist in 3D space, there are also out-of-plane vibrations, leading to a total of $3N$ vibrational modes. These vibrations include flexural modes104, which are responsible for breaking the flatness of suspended monolayer materials at finite temperatures. In graphene, which has two atoms per unit cell, there are two flexural modes, the acoustical flexural mode, which corresponds to the translation of the graphene at $k = 0$ along the direction normal to its plane, and the optical flexural mode, which corresponds to the out-of-phase oscillation of neighbouring atoms along the direction perpendicular to the graphene plane. For long wavelengths, the acoustic flexural phonon mode (or just the ‘flexural mode’) has a minimal restoring torque, and a distinct parabolic dispersion105,106.

**Multi-layer systems.** In multi-layer systems, the number of atoms is multiplied and thus so is the number of vibrational modes. FIGURE 5a illustrates this multiplication for the $A_2'$ mode of 2H-MoS$_2$. Each monolayer mode gives rise to $N$ modes in the $N$-layer system (the figure shows Raman-active modes only). Because of interlayer interactions, the frequency of the modes in the few-layer system is not the same as the monolayer system; in general, the bands split into different components; this is known as Davydov splitting. Additionally, few-layer systems may have different point group symmetries. Few-layer MoSe$_2$, with an even number of layers has inversion symmetry, whereas monolayer and other systems with an odd number of layers do not. This symmetry difference results in selection rules turning on and off different vibrational modes depending on the number of layers.

The splitting of vibrational bands with the number of layers has been observed, for example, in phosphorene107 and MoTe$_2$ (REF.108). The combination of acoustic (including flexural) vibrations in few-layer systems gives rise to shear modes and layer breathing modes when the layers move parallel to each other in different directions (FIG. 5b). Because interlayer interactions are very weak compared with covalent intralayer bonds, shear and interlayer breathing modes are very low in energy and therefore very difficult to observe directly in Raman spectroscopy, since they fall very close to the elastic scattering (Rayleigh) background. Their energy is also very sensitive to the layer number107–109.

**Observation of interlayer phonons.** In graphene, breathing modes have been observed by means of a process involving two phonons in doubly resonant Raman spectroscopy109 or using the coupling between the phonon and the optical transitions110. Shear modes have also been observed in bilayer graphene111. In few-layer phosphorene, the interlayer breathing modes are Raman-active and have been observed directly by Raman spectroscopy106,112. Transition metal dichalcogenides also have Raman-active shear and interlayer breathing modes that have been experimentally measured113,114. Such success is driven by recent advances in Raman spectroscopy techniques115. Theoretically, such low-frequency modes are difficult to model by density functional theory because of the small restoring forces involved.
compared with other crystal modes. However, they can be modelled by a simple linear chain model\textsuperscript{109,115}.

We mentioned that flexural phonons have dispersion $\omega \propto k^2$ (Ref\textsuperscript{4}), which has important consequences. Since the average population of phonons in mode $k$ is given by the Plank distribution,

$$\langle n(k) \rangle = [\exp(\beta Dk^2) - 1]^{-1},$$  

(10)

where $\beta$ is the $1/k_B T$ factor, the total number of phonons $n_k$ excited at temperature $T$ is given by:

$$n_k = \int_0^{\infty} g(k) \langle n(k) \rangle \, dk,$$  

(11)

where $g(k)$ is the density of states, $2\pi k$, in 2D. At long wavelengths, $\exp(\beta Dk^2) \sim 1 + \beta Dk^2$. Thus, at any temperature $T > 0$, the number of phonons diverges. This divergence would have catastrophic consequences for the stability of 2D materials — ruling out their existence altogether — were it not for the practical fact that 2D materials coexist with the surrounding 3D space (for example, the substrate they are on), which damps oscillations. As the number of layers increases, the dispersion of flexural modes becomes increasingly linear in $k$ (Ref\textsuperscript{103}).

However, free-standing monolayer membranes, such as graphene drum-heads\textsuperscript{116} or other suspended graphene configurations, are never really flat except at very low temperatures. This is a consequence of the dependence of the phonon density of states on $k$ in 2D. Mechanical vibrations are still damped even at resonance frequencies but can be excited by applied radio-frequency voltages\textsuperscript{117}. Graphene membranes are ideal for studying quantum motion. Because of their stiffness, the energies of the flexural modes are higher than cryogenic temperatures used in the experiments, ensuring the preservation of the quantum nature of these modes. The low mass of the membrane guarantees that the amplitude is high, allowing for easier detection of the zero-point motion\textsuperscript{118,119}.

**Magnons**

Magnons are collective excitations of the spins in magnetically ordered systems — ferromagnets and antiferromagnets. As the name indicates, they are quantized wave-like excitations, and therefore the term magnon is often used interchangeably with spin waves.

The ability to manipulate the magnetic state of 2D crystals is of great interest for the design of future ultra-high-density magnetic storage devices, spintronics devices, such as spin valves, and topological quantum computing. Although magnetism in 2D crystals has remained elusive until very recently, the study of magnetic excitations has already been greatly advanced by the ability to control the Fermi energy in 2D systems. In this section, we give a brief introduction to the controversy around the existence of ferromagnetism in 2D, topological excitations and spin density waves.

---

**Fig. 5 | Emergence of new vibrational modes in multi-layer 2D materials.** a | Raman-active out-of-plane vibrational modes in one to five layers (L) of MoSe$_2$. Dashed red (green) lines between the layers denote an increase (decrease) of the mode frequency compared with non-interacting layers. The horizontal dashed line and the dot indicate the mirror plane and inversion centre. b | Low-frequency Raman spectra from monolayer graphene (1LG, black line) and bilayer graphene (2LG, red line). The blue and red dots denote, respectively, the shear (C) mode and the layer breathing mode (LBM). The green dot denotes the combination mode (C + LBM). Panel a reprinted with permission from Ref\textsuperscript{146}, OSA. Panel b reprinted with permission from Ref\textsuperscript{147}, ACS.
2D ferromagnetism. The historical controversy around the existence of ferromagnetic order in 2D systems above zero temperature is intrinsically linked to the nature of 2D magnons. In the semi-classical Ising model of a 2D system of spins perpendicular to the material plane, a long-range ordered ferromagnetic state emerges below a critical temperature $T_c$. However, in the quantum mechanical Heisenberg model, where spins are allowed to point in any direction, low-energy magnons destroy such long-range order. More precisely, one can show that ferromagnetic or antiferromagnetic order is impossible above 0 K for isotropic 2D systems with only short-range interactions. This statement is known as the Mermin–Wagner theorem or Mermin–Wagner–Hohenberg theorem. The reason for the breakdown of long-range order can be readily understood from the following statistical reasoning for an isotropic medium. According to the 2D Heisenberg model, the dispersion of the lowest magnon band is quadratic:

$$\epsilon_k = D k^2.$$

(12)

Thus, following similar reasoning to that described for phonons in equations (10)–(11), we find that at any temperature $T > 0$, the number of magnons diverges.

Nevertheless, numerous experimental observations in 2D systems, starting with atomically thin epitaxial films of magnetic elemental systems (Co, Fe, Ni), have left no doubt that magnetic order is possible in 2D. Early observations in a Fe monolayer film on Ni(001) indicated that the magnetization is perpendicular to the monolayer, contrary to what would be expected from the continuous theory for surfaces of 3D ferromagnets. It was then understood that finite-size effects, anisotropy, spin–orbit coupling, magnetic dipole interactions and other interactions would break the assumptions of the Mermin–Wagner–Hohenberg theorem. In reality, it turns out that the finite size and anisotropy of real samples is too far removed from the conditions of the theorem.

Magnetic and ferromagnetic order has been observed in van der Waals materials and 2D sheets derived from van der Waals materials and 2D sheets derived from 3D materials. Examples of 2D crystals include CrI$_3$ [REFS (128,129)], Cr$_2$Ge$_2$Te$_6$ [REF. (130)] and FePS$_3$ [REF. (131)], with more predicted theoretically.

However, it should be noted that whereas such materials may be considered 2D from the point of view of structure or electronic interaction, they still have exchange interaction across layers. This is the case, for example, for CrI$_3$, one of the first ferromagnetic systems to be studied experimentally, where few-layer systems with an odd number of layers display ferromagnetism in the direction perpendicular to the plane, but with an even number of layers, the magnetization along this direction is suppressed owing to antiferromagnetic coupling between the layers. Another example is Cr$_2$Ge$_2$Te$_6$, for which $T_c$ vanishes in the monolayer limit, a result that can be reproduced by an ideal 2D Heisenberg model with negligible single-ion anisotropy.

The quadratic magnon dispersion near the $\Gamma$ point of the Brillouin zone (zero momentum), as given by the isotropic Heisenberg model, is a good approximation near the Brillouin zone centre for most 2D crystals [FIG. 6a–d]. Magnetic anisotropy, originating in the spin–orbit coupling, is a reason for the non-zero phonon energy at zero momentum (‘spin-wave gap’) [128,131], shown in FIG. 6b for CrI$_3$.

Magnon band structures for crystal lattices can be derived in a very similar way to phonon band structures. In honeycomb lattices, such as chromium trihalides CrX$_3$ ($X = F, Cl, Br, I$), two of the magnon bands form Dirac cones near the $K$ point, resembling the graphene
A statistical model in which the atomic spins are represented by vectors (classical Heisenberg model) or by spin or orbital angular momentum operators or their respective matrix representations (quantum Heisenberg model).

**Dzyaloshinskii–Moriya interaction**
A type of antisymmetric exchange interaction between neighbouring spins.

**Domain walls**
The boundaries between domains with different orientation of magnetization or other order parameter.

**Topological excitations.** In the 2D Heisenberg model, magnetic textures include vortices, and skyrmions and merons in magnets with a non-collinear spin ordering. Magnetic skyrmions are whirling configurations of magnetic momenta in which the orientation of spins rotates progressively from the up direction at the edge to the down direction at the centre, or vice versa (FIG. 7)\(^{136–138}\). More precisely, skyrmions (and antiskyrmions) are characterized by an integer topological charge, \(Q\):

\[
Q = \frac{1}{4\pi} \int d^2r (\partial_\alpha \mathbf{m} \times \partial_\beta \mathbf{m}) \cdot \mathbf{m},
\]

(13)

where \(\mathbf{m}\) is the order parameter, such as a unit vector along the direction of magnetization. \(Q\) describes how many times the magnetic moments wrap around a unit sphere upon application of stereographic projection\(^ {139}\). Merons have been described as half-skyrmions and have \(Q = -1/2\) (\(Q = 1/2\) for antimerons)\(^ {140}\). Although the lattice magnons considered in the previous section can be seen as quantized travelling waves and thus carry lattice momentum, skyrmions can carry angular momentum.

In 2D materials, skyrmions can be detected and manipulated with surface microscopy techniques. Some of the material characteristics to look for in 2D materials that may support skyrmions are a strong spin–orbit coupling, and additionally a Dzyaloshinskii–Moriya interaction, which favours spin canting and often seems to play an important role in stabilizing skyrmions\(^ {141}\).

A topological phase transition driven by the unbinding of vortex and antivortex pairs with the increase of temperature in XY ferromagnets has been proposed\(^ {142}\). Spin vortex phase transitions are still a matter of ongoing research.

Itinerant ferromagnetism in metals originates with the conduction (itinerant) electrons with an extended wavefunction. In a one-electron picture, this happens when the number of filled states for spin-up electrons and spin-down electrons is different (the difference can even be fractional). For example, SnO, GaS and GaSe have electron bands with Mexican hat dispersions, which have been shown to give rise to a ferromagnetic state if the Fermi level is chosen to lie in the nearly flat region\(^ {143}\). In twisted bilayer graphene moiré superlattices, close to a critical twist angle when the density of states at the Fermi energy becomes very large, ferromagnetic states emerge owing to increasing electron–electron interactions\(^ {144}\). Nevertheless, no ferromagnets of this type with Curie temperature close to room temperature are yet known.

**Spin density waves.** Spin density waves are broken-symmetry ground states of metals, with a periodic modulation of the charge density that is not necessarily commensurate with the lattice period. Unlike the spin waves (or magnons) considered in the preceding section, it is not the spin orientation that varies but the spin density. Charge density waves can have low-lying spin, charge and phase excitations. The spin density can vary in space in metallic systems in which there is a large density of states near the Fermi energy, particularly if there is nesting of the Fermi surface\(^ {145}\). For this reason, spin density waves, charge density waves and superconductivity often occur in the same 2D systems, and their interplay is still an object of research.

**Outlook**
In the decade and a half since the discovery of graphene, 2D materials have become an integral part of physics and materials science research driven by rapid scientific progress. This increased understanding has opened new frontiers in the study of 2D systems.

In the field of plasmonics, one of the most critical challenges has to do with augmenting the plasmon propagation length, which can be achieved by careful sample engineering. This improvement would allow the use of plasmons as signal carriers in plasmonic devices. Improved sample preparation techniques would also expand the plasmonic field from graphene to include materials such as phosphorene, predicted to host anisotropic plasmons. Finally, ultraclean samples make it possible to probe the hydrodynamic regime, where Lindhard treatment of electron–electron interaction becomes insufficient. The theoretical formalism of these hydrodynamic plasmons is currently an active area of research.

In the excitonic field, there are two main directions whose intertwining may result in non-trivial physics and applications. The directions are moiré excitons in twisted 2D semiconducting bilayers, and excitonic condensation.
in 2D homo- and heterostructures. Facilitating interactions by means of moiré excitons and making use of the increased interlayer exciton lifetime in excitonic condensation may produce synergistic effects yet to be discovered.

Phonon dispersion simulations in 2D materials are still very challenging, making it difficult to assess phonon-limited mobility and phonon-mediated heat transport. It will be necessary in future to quantify such effects in 2D materials on substrates and in heterostructures.

Since 2D magnetic materials have been realized, it is possible to study domain wall motion and topological excitations. There is ample room to find new physics; the barrier to applications seems to be the small Curie temperature of the materials that are known so far.

In general, the effect of electron–electron correlation and its impact in all these crystal excitations is still not well understood. New experiments on moiré lattices of twisted layers may be a way to vary correlation and understand its effects. The reason that spin density waves, magnetism and superconductivity coexist in some materials and the definition of their phase diagram is still an open problem.

Lastly, it can be said that any 2D material is its surface; this provides a window to observe in real space any type of collective excitation in an unprecedented way. Paraphrasing Isaac Newton, at this point in time, we are only children playing on a beach, while vast oceans of truth lie undiscovered before us.

Published online 7 September 2020
Mak, K. F., He, K., Shan, J. & Heinz, T. F. Control of valley polarization in monolayer MoS2 by optical pumping. Nat. Commun. 20, 201111 (2016).

Fröhlich, H. Electrons in lattice fields. Rev. Mod. Phys. 39, 14 (1967).

Ruiz-Tijerina, D. A. & Fal’ko, V. I. Interlayer coupling in polaronic trions. Rev. Mod. Phys. 91, 045002 (2019).

Kezerashvili, R. Y. & Tkiskauri, S. M. Trion and biexciton decoherence in monolayer MoS2 by spin-charge coupling. Phys. Rev. B 98, 085415 (2018).

Rivera, P. et al. Interlayer valley excitons in transition metal dichalcogenides. Nat. Nanotechnol. 14, 1004–1015 (2019).

Tiran, K. et al. Evidence for moiré excitons in two-dimensional materials. Nature 567, 76–80 (2019).

Selig, M. et al. Excitonic linewidth and coherence lifetime in monolayer transition metal dichalcogenides. Nat. Commun. 7, 1–6 (2016).

Pechinger, G. et al. Trion fine structure and coupled spin–valley–dynamical states in monolayer tungsten disulfide. Phys. Rev. Lett. 119, 1–6 (2017).

Liu, X. et al. Strong light–matter coupling in two-dimensional atomic crystals. Nat. Photonics 9, 50 (2015).

Gluza, M. et al. Exciton fine structure and spin decoherence in monolayers of transition metal dichalcogenides. Phys. Rev. B 89, 201302 (2014).

Zhang, X. et al. Raman spectroscopy of shear and layer breathing modes in few-layers black phosphorus. Nano Lett. 15, 3931–3935 (2015).

Lui, C. H. & Heinz, T. F. Measurement of layer breathing mode vibrations in few-layer graphene. Phys. Rev. B 87, 121404 (2013).

Song, X. et al. Observation of low-temperature Raman scattering of layered transition metal dichalcogenides. Nature 546, 75–80 (2017).

Mak, F. K., He, K., Shan, J. & Heinz, T. F. Control of valley polarization in monolayer MoS2 by optical helicity. Nat. Nanotechnol. 7, 494–498 (2012).

Brey, L., Fertig, H. A., Côté, R. & MacDonald, A. H. One of the earliest observations of ferromagnetism in a van der Waals crystal down to the monolayer limit. Nat. Phys. 5, 645–649 (2009).

Brey, L. & Fernández-Rossier, J. On the origin of magnetic anisotropy in two-dimensional CrI3. Nat. Mater. 4, 205–206 (2005).

Kezele, S. et al. Dirac magnons in honeycomb ferromagnet. Phys. Rev. X 8, 011010 (2018).

Fransson, J., Black-Schaffer, A. M. & Balatsky, A. V. Magnon Dirac fermions. Phys. Rev. B 89, 075401 (2015).

Lee, J.-U. et al. Isothreading magnetic ordering in few-layer MoS2 and WSe2. Nano Lett. 16, 7433–7438 (2016).

Lado, J. L. & Fernández-Rossier, J. Ferromagnetism in a van der Waals crystal down to the monolayer limit. Nat. Phys. 5, 645–649 (2009).

Brey, L., Fertig, H. A., Côté, R. & MacDonald, A. H. Strongly Correlated Magnetic and Superconducting Systems (eds Sierra, S. & Martin-Delgado, M. A.) 275–283, (Springer, 1997).

Fert, A., Cros, V. & Sampsa, J. Skyrmions on the track. Nat. Nanotechnol. 8, 152–156 (2013).

Kovalev, A. A. & Sandhofer, S. Skyrmions and antiskyrmions in quasi-two-dimensional magnets. Europhys. Lett. 89, 67008 (2010).

Yu, X. et al. Transformation between meron and skyrmion topological phases in a chiral magnet. Nature 564, 49 (2019).

Banerjee, S., Rowland, J., Erten, O. & Randeria, M. Enhanced stability of skyrmions in two-dimensional chiral magnets with Rashba spin–orbit coupling. Phys. Rev. X 4, 031045 (2014).

Sorokin, A. Critical density of topological defects upon a continuous phase transition. Ann. Phys. 411, 167952 (2017).

Seixas, L., Rodin, A., Carvalho, A. & Castro Neto, A. Multiferroic two-dimensional materials. Phys. Rev. Lett. 112, 206805 (2014).

Sharpe, A. L. et al. Emergent ferromagnetism near three-quarters filling in twisted bilayer graphene. Science 365, 605–608 (2019).
Acknowledgements
This work was supported by the National Research Foundation, Prime Minister Office, Singapore, under its Medium Sized Centre Programme and CRP award ‘Novel 2D materials with tailored properties: beyond graphene’ (grant no. R-144-000-295-281). A.R. acknowledges support by Yale-NUS College (through grant no. R-607-265-380-121). M.T. is supported by the Director’s Senior Research Fellowship from the Centre for Advanced 2D Materials.

Author contributions
The authors contributed equally to all aspects of the article.

Competing interests
The authors declare no competing interests.

Peer review information
Nature Reviews Physics thanks Ursula Wurstbauer and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher’s note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© Springer Nature Limited 2020