For ten years, the ExoMol database of molecular line lists has provided vital spectroscopic data for the study of hot atmospheres. Now, after 700 billion molecular transitions, Jonathan Tennyson and Sergei N Yurchenko reflect on what the project has achieved and what the future may hold.

ExoMol at 10

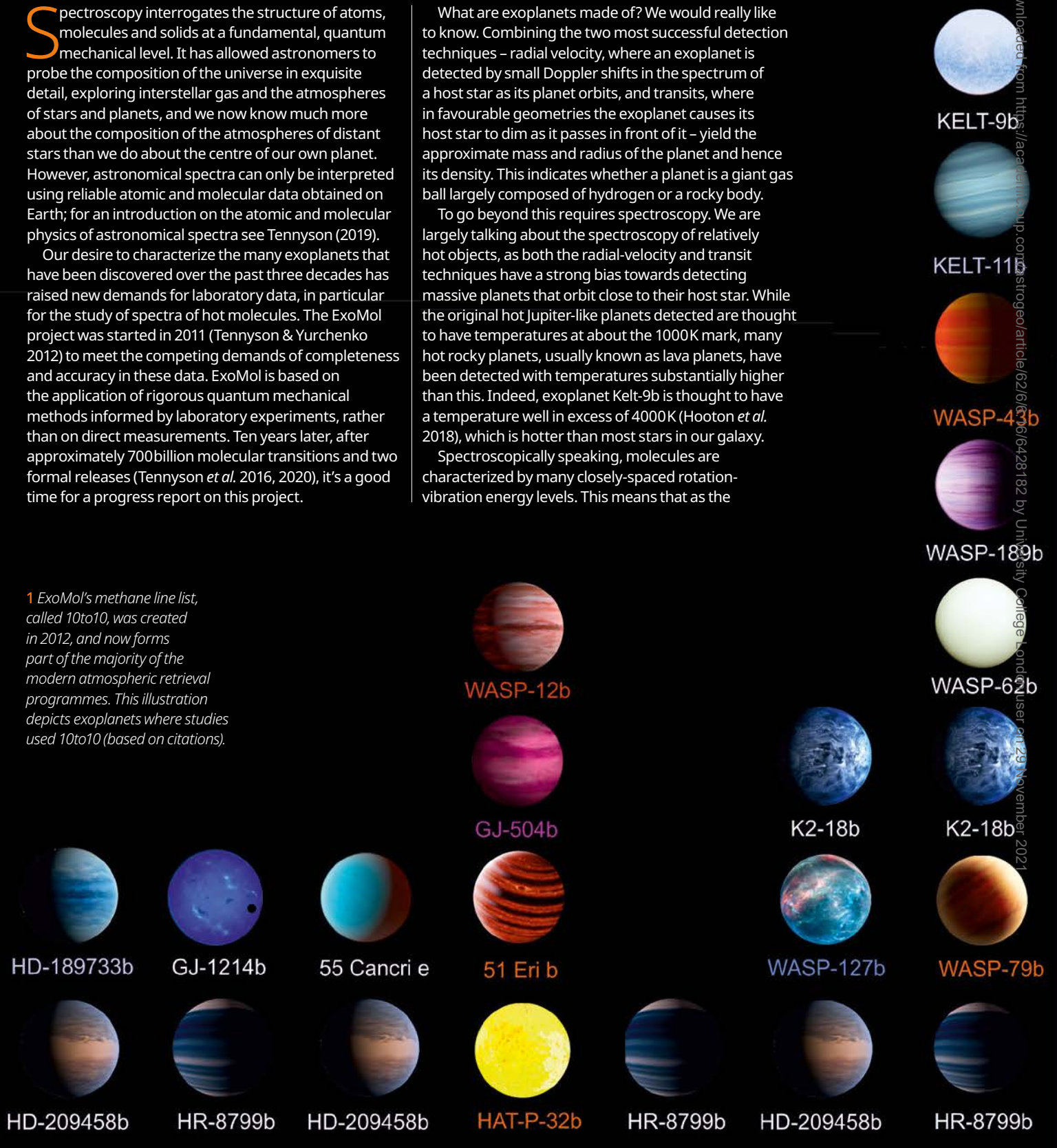
Spectroscopy interrogates the structure of atoms, molecules and solids at a fundamental, quantum mechanical level. It has allowed astronomers to probe the composition of the universe in exquisite detail, exploring interstellar gas and the atmospheres of stars and planets, and we now know much more about the composition of the atmospheres of distant stars than we do about the centre of our own planet. However, astronomical spectra can only be interpreted using reliable atomic and molecular data obtained on Earth; for an introduction on the atomic and molecular physics of astronomical spectra see Tennyson (2019).

Our desire to characterize the many exoplanets that have been discovered over the past three decades has raised new demands for laboratory data, in particular for the study of spectra of hot molecules. The ExoMol project was started in 2011 (Tennyson & Yurchenko 2012) to meet the competing demands of completeness and accuracy in these data. ExoMol is based on the application of rigorous quantum mechanical methods informed by laboratory experiments, rather than on direct measurements. Ten years later, after approximately 700 billion molecular transitions and two formal releases (Tennyson *et al.* 2016, 2020), it's a good time for a progress report on this project.

What are exoplanets made of? We would really like to know. Combining the two most successful detection techniques – radial velocity, where an exoplanet is detected by small Doppler shifts in the spectrum of a host star as its planet orbits, and transits, where in favourable geometries the exoplanet causes its host star to dim as it passes in front of it – yield the approximate mass and radius of the planet and hence its density. This indicates whether a planet is a giant gas ball largely composed of hydrogen or a rocky body.

To go beyond this requires spectroscopy. We are largely talking about the spectroscopy of relatively hot objects, as both the radial-velocity and transit techniques have a strong bias towards detecting massive planets that orbit close to their host star. While the original hot Jupiter-like planets detected are thought to have temperatures at about the 1000 K mark, many hot rocky planets, usually known as lava planets, have been detected with temperatures substantially higher than this. Indeed, exoplanet Kelt-9b is thought to have a temperature well in excess of 4000 K (Hooton *et al.* 2018), which is hotter than most stars in our galaxy.

Spectroscopically speaking, molecules are characterized by many closely-spaced rotation-vibration energy levels. This means that as the



1 *ExoMol's methane line list, called 10to10, was created in 2012, and now forms part of the majority of the modern atmospheric retrieval programmes. This illustration depicts exoplanets where studies used 10to10 (based on citations).*

temperature increases, a greater number of these levels can be occupied, which, in turn, leads to an explosion in the number of transitions. Experience shows that, as in figure 2, reproducing the observed spectrum of a hot (in a molecular sense) astronomical object can require billions of transitions. Experimental high-resolution spectroscopy is a mature subject that has produced a wealth of data, such as that required to characterize the many molecules observed in our own atmosphere or at very low temperatures in the interstellar medium. However, it struggles to provide the volume of data needed to represent either the spectrum or radiative transport properties in hotter objects. Hence the need for a theoretical solution to this problem.

Solution strategy

It is relatively straightforward to formulate the quantum mechanical equations that must be solved to predict the spectrum of a molecule as function of temperature. The problem is that these equations result in sets of high-dimensional coupled second-order differential equations that can only be solved approximately for most systems, even using the largest computers available.

The ExoMol philosophy has therefore been to make a rigorous quantum mechanical formulation of the spectroscopic problem for each molecule, to use the best practicable method for solving this problem, and to then use any available experimental data to improve our spectroscopic model. We describe this as "first principles theory informed by experiment". We will here give a broadbrush picture of the approach (see figure 3); a more technical description can be found in Tennyson (2012).

Working within the standard (Born-Oppenheimer) approximation of molecular physics involves solving separately for the electronic and nuclear motions of a given molecule. For the electronic structure problem – easy to formulate but very difficult to solve accurately – we generally use standard quantum chemistry electronic structure codes. Our favoured code is Molpro (Werner *et al.* 2012). Solving the electronic structure problem gives a potential energy surface (PES) upon which the nuclei move, as well as dipole moment surfaces (DMS) which are used to compute transition probabilities.

The second step of the procedure is to solve the appropriate Schrödinger equation for nuclear motion on the potential energy surface. For this we use a range of codes we have helped to develop (Tennyson & Yurchenko 2017) which employ the powerful variational principle to provide near-exact solutions for the given nuclear motion problem. This makes the underlying potential energy surface the source of most of the uncertainty in our calculations, although small corrections due to the failure of the Born-Oppenheimer approximation or the neglected relativistic motions of the electrons often also have to be accounted for.

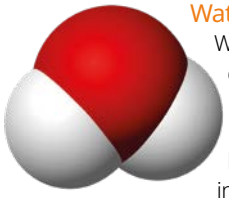
PESs can be improved by comparing the results of a calculation with laboratory observations and adjusting potential energy parameters to improve agreement. This can be time-consuming and fiddly but gives results that are valid over a wide range of energy levels, and hence transitions. Interestingly, for most systems we have studied, the *ab initio* dipole moments have often yielded transition intensities as accurate as state-of-the-art experimental determinations (Polyansky *et al.* 2015).

Even with improved PESs, our calculations usually do not reproduce the accuracy of laboratory measurements. For many astronomical applications this may not matter but the increasing use of high-resolution studies has led to the demand for high accuracy. To put it crudely, we have no difficulty providing accurate data for observations with resolving power (R) of a few hundred needed for upcoming exoplanet characterization space missions such as Ariel (Tinetti *et al.* 2018) or Twinkle (Savini *et al.* 2018). Similarly, in most cases our provision of data is good enough for the $R\approx3000$ observations anticipated for the James Webb Space Telescope (JWST). However, the requirement for spectroscopic data for observations with $R\geq100000$ provide much more of a challenge. Such data are required for the new breed of ground-based telescopes such the Extremely Large Telescope (ELT) under construction at the European Southern Observatory. To meet this demand we are increasingly incorporating measured energy levels directly into our spectroscopic models, which allows wavelengths to be predicted for many lines – including those which have yet to be observed in the laboratory – to high accuracy.

Of course ExoMol isn't unique in providing spectroscopic data for studies of exoplanets. The TheoRets project (Rey *et al.* 2016) and NASA Ames have produced a number of line lists for key polyatomic molecules using techniques similar to those described above. In addition, the MoLLIST project (Bernath 2020) has produced line lists, available via the ExoMol website (exomol.com), mainly for diatomic species, produced largely by combining experimental wavelength measurements with dipole moments calculated using sophisticated electronic structure methods similar to those used by ExoMol.

For temperate planets with Earth-like temperatures, the HITRAN database (Gordon *et al.* 2017) provides the first port of call. HITRAN has been worked on for over 50 years and provides comprehensive and accurate coverage including line lists and line broadening parameters for some 50 small molecules that occur in the Earth's atmosphere, plus cross sections for larger species, as well as treatments of collision-induced absorption. HITRAN makes no claim to be comprehensive at higher temperatures although we are participating in a new release of the associated HITEMP database, designed to extend HITRAN to higher temperatures underway, (for example Hargreaves *et al.* 2019).

The ExoMol database contains extensive line lists for approximately 80 molecules, in most cases for several isotopically substituted variants ("isotopologues") (Tennyson *et al.* 2020). Rather than list them all, let's consider a few important species in detail.

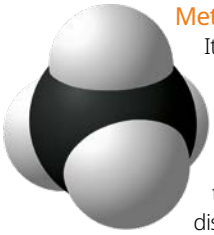


Water

Water was the first molecule detected in an exoplanet (Tinetti *et al.* 2007). The reliance on an extensive calculated water line list to make this detection, an initial analysis with a less complete line list having proved unsatisfactory, provided much of the inspiration for launching the ExoMol project. Water is widely detected in exoplanets (Tinetti *et al.* 2013, Madhusudhan 2019) and its detection on planet K2-18b, which orbits within its star's so-called habitable zone (Tsiaras *et al.* 2019), sparked worldwide media interest. The ExoMol project has provided a comprehensive water line list, "Pokazatel" (Polyansky *et al.* 2018), which considers all transitions between rotation-vibration states that lie below the limit where the water molecule dissociates.

Figure 4 gives the predicted absorption cross section of the Pokazatel line list as function of temperature. The change in shape of the cross section is caused by the rapid increase in the number of transitions as the water molecule becomes hotter. This increase is due to both an increase in the number of rotational transitions, which makes each band broader, and the occurrence

of so-called hot bands, which result from absorption by thermally populated vibrational states. These hot bands have similar structures to the rotation-vibration bands arising from the vibrational ground state but with shifted band origins; they become increasingly important as temperature rises and causes the flattening of the overall absorption cross sections.

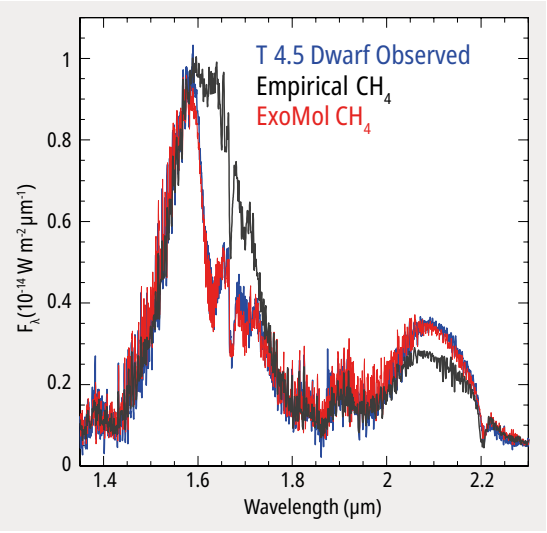


Methane

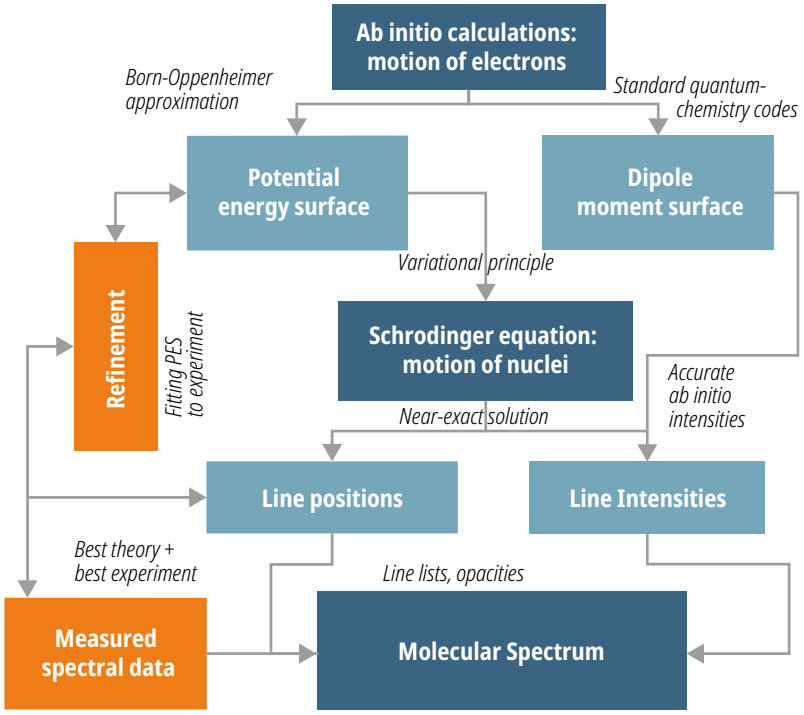
It was clear at the outset of the ExoMol project in 2011 that the laboratory data on methane (CH_4) were inadequate for exoplanetary atmospheric studies. At that time, the majority of exoplanet discoveries were short-period hot Jupiters requiring spectroscopic data covering high temperatures, 1000–2000 K and even higher; the existing laboratory spectra of methane came almost entirely from the room-temperature measurements. One common outcome of the atmospheric retrievals from that time was higher C/O ratios because the amounts of methane detected were overestimated. An illustration of the problem of interpreting hot spectra with room-temperature spectroscopic data is given in figure 2, which shows an attempt to model the spectrum of a T-dwarf using the best then-available (room-temperature) laboratory data. The comparison with the actual observations demonstrates the lack of the underlying opacity, which leads to the requirement of unrealistic amounts of methane being inferred.

Yurchenko & Tennyson (2014) produced ExoMol's line list for CH_4 using the computational methodology described above; this calculation required 4.5 million CPU hours on the UK National HPC facilities. Figure 2 also shows a synthetic spectrum of a T-dwarf atmosphere calculated using ExoMol's line list for methane ("10to10"). This line list contains almost 10^{10} transitions constructed to provide an adequate description for hot atmospheric conditions. As a comparison, the then-existing laboratory spectroscopic data on CH_4 (as provided by the 2012 edition of the HITRAN database) contained about 340 000 lines, 300 000 times fewer than were provided by ExoMol.

The 10to10 line list quickly became popular among scientists studying exoplanetary atmospheres. Figure 1 illustrates exoplanetary atmospheric studies where the 10to10 line list was used (based on the citations); cross sections and other data generated from 10to10 form a now-integral part of the majority of the modern atmospheric retrieval programs: ARCiS, Taurex, petitRADTRANS, HELIOS, ATMO, NEMESIS, CHIMERA, PLATON, Aurora, Pyrat-Bay, EXOPLINES etc.



2 *Spectrum of CH_4 in methane-rich T 4.5 dwarf star as observed with NASA's Infrared Telescope Facility. The model fits use the empirical STDS database of methane spectra and the ExoMol 10to10 CH_4 line list. To get a good fit to the spectrum required retaining 3.2 billion of the almost 10 billion transitions in the 10to10 line line list; see Yurchenko et al. (2014) for further details.* (Tennyson and Yurchenko)



3 *Schematic diagram of the ExoMol approach to generating molecular line lists; the approach can loosely be described as using first principles quantum mechanics informed by experiment.* (Tennyson and Yurchenko)

The computation of such an extensive line list required technological and theoretical advances. The computational programs had to be developed and optimized for efficient numerical quantum-mechanical solutions and the production of billions of spectroscopic lines. When it comes to methane there is the additional step that allows for the high symmetry of the molecule and saves computer time, but results in a subtly different structure to the many rotation-vibration Hamiltonian matrices that needed to be diagonalized. The project also relied on the support of powerful high-performance computing (HPC) centres; the work at 10to10 used the UK National Cosmology Supercomputer COSMOS (1.5 million CPU hours) and The Cambridge HPC Cluster Darwin (3.0 million CPU hours). Other studies have made heavy use of the UCL high-performance computing clusters Legion, Myriad and Grace, particularly for calculating very extensive potential energy surfaces. More recent nuclear motion calculations have used machines run by the STFC DiRAC service at Cambridge, and we have also moved to make extensive use of GPUs where possible.

The opacity of CH_4 is practically solved for low-resolution atmospheric retrievals. However, for the high-resolution (HR) technique the line positions have to be improved, especially in the near-IR regions. Figure 5 illustrates why it is a challenging problem to experiment with analyzing the congested methane spectrum even at low temperatures. Such analysis requires good-quality predictions, efficient software, skilful people, sweat, tears and time.

The spectroscopic detection of molecules such as methane in exoplanets can be performed using a number of techniques. Figure 6 gives examples of methane detections using the three most common methods. The original detection of methane in exoplanet HD 189733b by Swain *et al.* (2008) used transit spectroscopy, which can provide a broad overview of how a planet's atmosphere either absorbs (primary transit) or emits/reflects (secondary) light. Transit spectroscopy remains widely employed and is the basis for upcoming space missions, but other techniques are now also available.

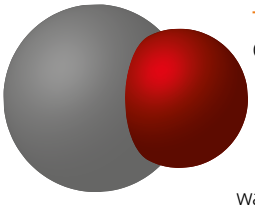
With the new powerful telescopes, more directly imaged planets are now being studied, often targeting spectroscopic regions where CH_4 can be seen absorbing

or emitting light. An example of the detection of CH_4 in a directly imaged exoplanet, 51 Eri b, is illustrated in figure 6.

The new HR technique pioneered by Snellen (2014) changed attitudes to the analysis of the atmospheres of exoplanets as well as the data requirement. It is important for HR molecular retrievals that the underlying data are provided at high resolution ($R \geq 100\,000$). The success of the ExoMol line list for CH_4 came from its completeness. However, the HR standards also demand that accuracy increases, establishing new challenges for the modern computational spectroscopists and the experimental spectroscopic community. ExoMol is part of the global effort to improve methane data to meet these standards, involving theoreticians, computational scientists and experimentalists. Figure 6 shows a final example of the detection of CH_4 in an exoplanetary atmosphere using a cross-correlated HR signal for planet HD 102195b.

The importance of methane for atmospheric studies of exoplanets motivated other key players from the spectroscopic community to measure and calculate methane spectra, thus providing more accurate data with better coverage of wavelengths and temperatures. A high-quality CH_4 line list from the TheoReTS group (Rey *et al.* 2016) subsequently provided an accurate description of the experimental spectroscopy of CH_4 as high as 1000 K and covering up to the near IR (Wong *et al.* 2019) region.

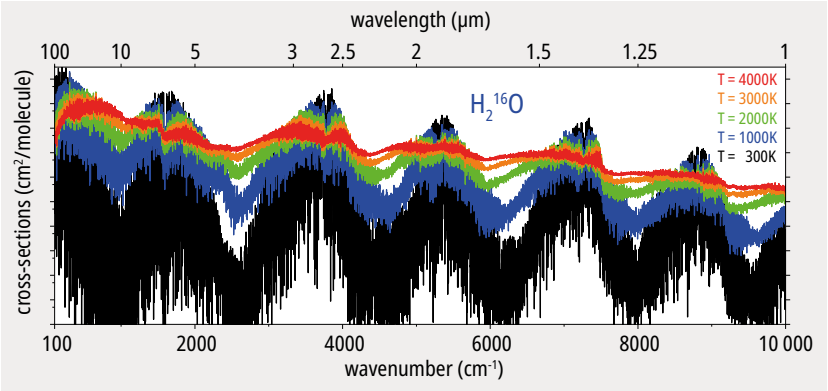
"Modern quantum chemistry has not fully realized how important molecules such as TiO are for astrophysics"



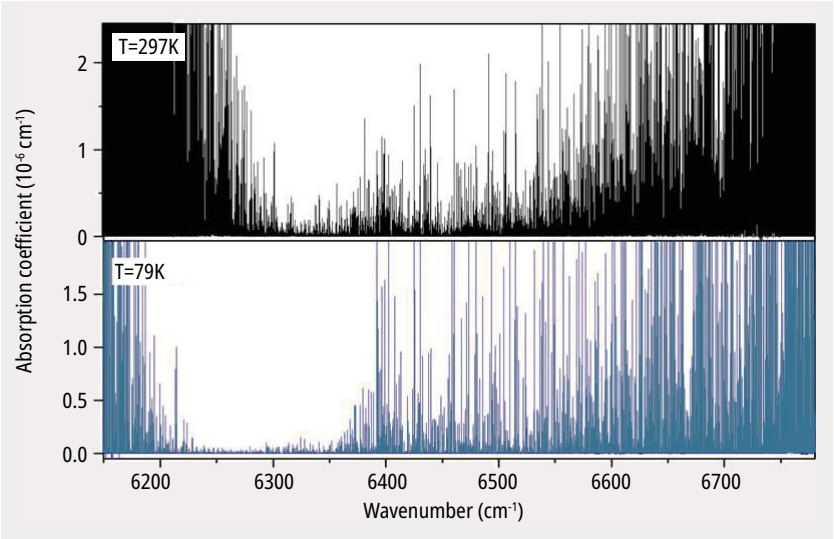
Titanium monoxide

Once the methane opacity data became available for exoplanetary studies, this molecule lost its status as the molecule for which everyone wanted data, and was replaced by TiO. In many exoplanets TiO is thought to have an atmospheric importance similar to that played by ozone around Earth.

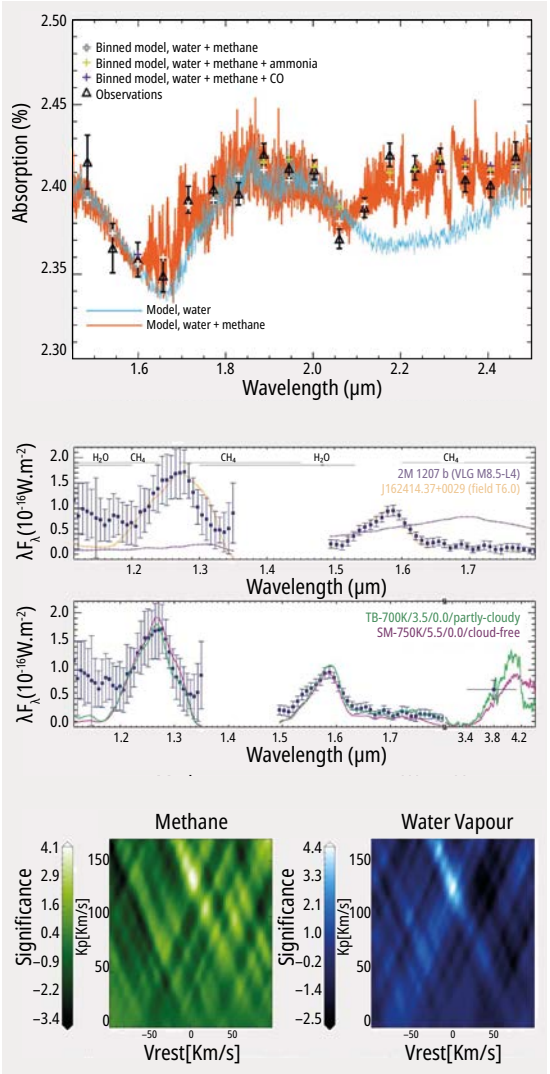
TiO_2 is a stable (closed-shell) molecule and a major ingredient of white paint. Astronomically, however, it is titanium monoxide (TiO) that is important. TiO has as number of low-lying electronic states that absorb strongly at near-infrared and red wavelengths, which is at the peak flux of many cool stars. Absorption by TiO is thought to be a crucial piece of the puzzle for atmospheric studies of exoplanets with thermal inversion. However it is a real challenge, both for theory and experiment, to produce good TiO opacity data. The spectrum of TiO must be measured at temperatures higher than 1500 K for the molecule to be in the gas phase. Only a very few labs in the world are capable of HR spectroscopic studies at these temperatures under precisely controlled conditions, and even then it is very hard to place any observed intensities on an absolute scale. The spectrum of TiO



4 *Predicted absorption cross section of the Pokazatel water line list as function of temperature (Polyansky et al. 2018). Note the strong temperature dependence of the shape of the cross section: the relatively flat hotter cross sections require very extensive line lists to be modelled correctly. The Pokazatel line list contains over 5 billion transitions which should be compared with just over 140 000 $H_2{}^{16}O$ transitions provided by databases such as HITRAN (Gordon et al. 2017) used to model radiative transport in our own atmosphere.*



5 *An example of a congested laboratory spectrum of methane from Campargue et al. (2012). These spectra, recorded at room and liquid nitrogen temperatures, are challenging to analyze. Raising the temperature leads to significantly increased congestion.* (Elsevier)



6 *Three spectroscopic detections of methane in exoplanets. Upper panel: transit spectroscopy of HD 189733b by Swain et al. (2008). Middle panel: directly imaged exoplanet 51 Eri b observed by Macintosh et al. (2015). Lower panel: high-resolution Doppler shift detection of water and methane in HD 102195b by Guilluy et al. (2019).* (Springer Nature, Science, ESO)

consists of multiple electronic bands; the most important are the visible bands. The spectral analysis is not straightforward for such a system with many crossings and couplings, but neither is it hopeless.

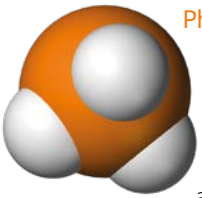
Theoretically, the challenge is in the solution of the electronic part of the problem. The electronic structure of TiO is largely determined by the character of the partially occupied Ti 3*d* orbital, which is not well enough described by any current quantum-chemistry software, not least because this character changes between electronic states. We are still using quantum chemistry methods developed over 30 years ago; modern quantum chemistry has not fully realized how important molecules such as TiO are for astrophysics.

While it is common to use the TiO as a test system for new methods, such calculations mostly explore a single internuclear separation corresponding to the equilibrium geometry, and give properties at this point. This is a long way from the curves and couplings that are required to build a full spectroscopic model for the system, even one that only explores states absorbing in the infrared and visible. This situation is true for many diatomic molecules containing a transition metal of interest to astrophysics such as FeH, CrH and FeO.

The most practical approach must therefore be semiempirical. The spectroscopic model of TiO, which contains potential energy curves and different couplings, is refined by fitting to accurate experimental data (energies or line positions). This means that the more experimental data are available, the better the empirical model. The long-awaited ExoMol TiO line list, called Toto, was published in 2019 (McKemmish *et al.* 2019), and immediately found an application in multiple studies of exoplanets, although results remain a little patchy.

Attempts to detect TiO in explanets have been a saga. Take the planet WASP-121b. In 2016, evidence of TiO was obtained via detection of water (Evans *et al.* 2016), which it was thought would not be there unless the UV radiation of the star is blocked by an ozone-like stratospheric layer populated by absorbers where the stellar radiation peaks, with TiO and VO being the main suspects. In 2017, Nugroho *et al.* (2017) detected TiO using HR spectroscopy and an inaccurate line list by Kurucz. However, in 2020 no detection of TiO was made using HR dispersion spectroscopy and our new ToTo line list.

In 2017 TiO was detected in WASP-19b (Sedaghati *et al.* 2017) using data from before ExoMol's Toto. Figure 7 shows a transmission spectrum in visible light of WASP-19b with clear TiO features. Toto has been used subsequently in many exoplanet studies. Attempts to detect TiO using HR spectroscopy, which should give the most secure detections, have been patchy with successes, failures and concern about getting false positives due to the sheer number of possible lines. The saga continues.



Phosphine

Phosphine (PH_3) is well-known as a constituent of the atmosphere of Jupiter and, despite being black-listed on Earth as a chemical weapon, it has been suggested as a possible biosignature in anoxic (oxygen-lacking) exoplanets (Sousa-Silva *et al.* 2020). So far, phosphine has yet to be detected in the atmosphere of an exoplanet; however, the recent claimed observation of phosphine in the clouds of Venus by Greaves *et al.* (2020) and the suggestion that this may be a signature of some life form has sparked considerable controversy.

The ExoMol line list for phosphine, known as SAITY (Sousa-Silva *et al.* 2015), suggests that the molecule shows a strong spectral feature at about 4.3μm, as shown in figure 8. This region of the infrared will soon be accessible to space telescopes which will, we hope, lead to detections of phosphine in exoplanet atmospheres. Any association of such a detection with life would take longer to establish.

Novel species

To achieve reliable atmospheric retrievals it is important to include comprehensive coverage of molecular absorbers. Therefore any new line lists provided by ExoMol are included in atmospheric analyses and compared with existing and new observations, sometimes motivating new discoveries. Examples of discoveries of molecules in atmospheres of exoplanets triggered by new ExoMol or MoLList data sets include detections of AlO, SiH, TiO, VO, TiH, OH, FeH and C_2H_2 molecules.

One further interesting example is calcium oxide (CaO). Our calculated line list (Yurchenko *et al.* 2016) suggests that CaO has particularly strong electronic transitions at visible wavelengths but the molecule has yet to be detected in space. However, Berezhnoy *et al.* (2018) used our line list to identify transitions in the so-called CaO orange band in meteor spectra.

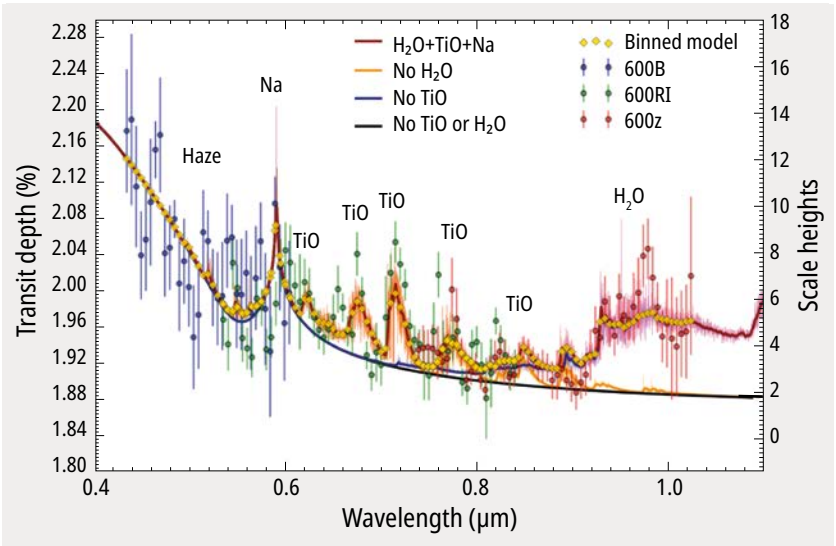
Data products

Spectroscopy is an underpinning technique widely applicable to many areas of science, not just astronomy and geophysics. For example, our line list for BeH, one of the first ExoMol line lists, was improved and extended to the BeD and BeT isotopologues to allow the analysis of spectra recorded in fusion plasma experiments (Darby-Lewis *et al.* 2018), which are nowadays experimenting with beryllium walls.

The ExoMol project provides transition data, partition functions, lifetimes of individual states, cross sections (k-tables) as standard products to astronomers. These data are also used by scientists studying combustion and plasma processes as well as design studies ranging from attempts to make new THz lasers to the detection of dark matter. The transition dipoles that underpin the transition probabilities at the heart of the ExoMol database have been used for a range of theoretical studies on molecular steering (Owens *et al.* 2017). We continue to develop our data products in response to requests from scientists.

Present and future work

So, after a decade of ExoMol, what are we doing and what are our plans? Computing more line lists is one obvious continuing activity, as requests for these are still coming thick and fast. However, there is strong



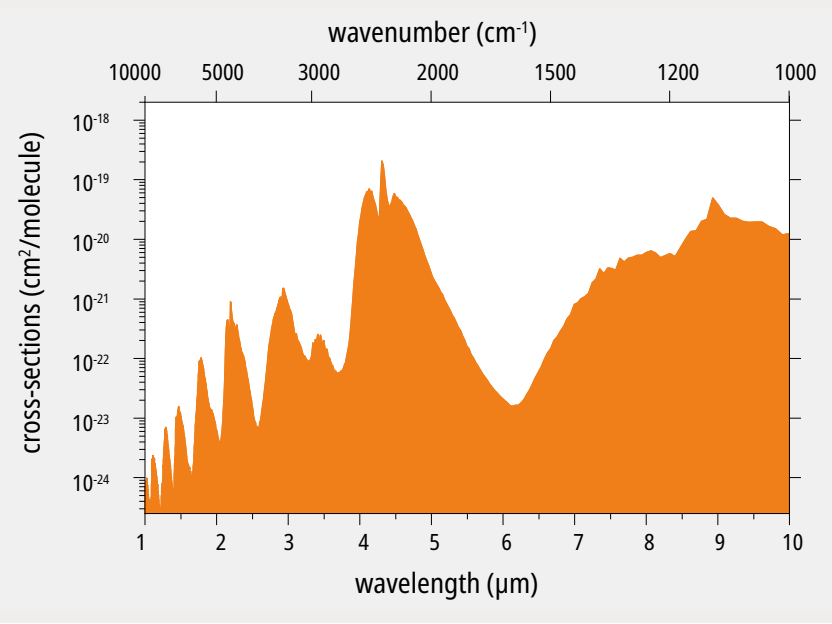
7 *Detection of TiO in Wasp-19b by Sedaghati et al. (2017).* (Springer Nature)

demand for HR spectra to aid molecular detections using HR Doppler shift spectroscopy. The required accuracy can only be achieved using frequencies obtained from laboratory HR spectroscopy experiments. Our procedure for this is called MARVEL (Measured Active Vibration Rotation Energy Levels) (Furtenbacher *et al.* 2007) and involves inverting the laboratory measurements to give a set of empirical energy levels with spectroscopic accuracy. These energies are then used to substitute for those calculated in our line list. This approach not only guarantees that we reproduce the observed transitions accurately, but results in many other transitions being predicted with HR accuracy. For example, a recent MARVEL study of formaldehyde (CH_2O) used 16403 non-redundant observed transitions to give 5029 empirical energy levels; these in turn gave 367779 transition frequencies with accuracy good enough to be used in HR astronomical studies (Al-Derzi *et al.* 2021).

MARVEL projects are labour-intensive because they require retrieving all available HR spectroscopic data from the literature. Several MARVEL projects have been performed as part of the ORBYTS (Original Research By Young Twinkle Scientists) outreach project, including on TiO and methane. ORBYTS involves school children in active scientific research with the outcome that they become co-authors on scientific papers, see Sousa-Silva *et al.* (2018). ORBYTS won the UK's 2021 NEON (National Education Opportunities Network), award for Widening Access Initiative (Outreach). So far we have only completed MARVEL projects for about one-fifth of the molecules studied by ExoMol, so there is plenty more to do.

One spin-off from MARVEL studies involves isotopically substituted species, which are now being detected in exoplanetary atmospheres (Zhang *et al.* 2021). In general these isotopologues require high resolution to distinguish their spectra from that of their parent. We have developed methods of improving the predicted transition wavelengths of isotopologues based on observed minus calculated differences found in our calculations on the parent (Polyansky *et al.* 2017). Our original study was for water but the method has since been successfully used for TiO (McKemmish *et al.* 2019), and will be applied to many more systems in due course.

The original ExoMol project concentrated heavily on wavelengths longer than 0.4μm, i.e. in the infrared and visible parts of the spectrum. However, it is clear that there is a demand for molecular data in the ultraviolet range. Our response has been both to start to extend our line lists into the UV and to start a separate set of calculations



8 *Infrared spectrum of PH_3 at T = 1000 K simulated using the SAITY line list (Sousa-Silva et al. 2015); note the strong, characteristic 4.3 μm absorption band, which could provide a signature for this molecule in exoplanetary atmospheres.*
(Tennyson & Yurchenko)

to study molecular photodissociation. This has required a slightly different approach to solving the nuclear motion Schrödinger equation (Pezzella *et al.* 2021). Besides any effect this process might have on the radiative transport through an exoplanet atmosphere, photodissociation at the top of the atmosphere, which is usually bathed in starlight, has a profound influence on the atmospheric chemistry of the planet.

One important missing link in spectroscopic data provision is line broadening. Spectroscopic transitions are characterized by line profiles whose shapes and width are affected by different physical effects, such as broadening caused by the natural lifetimes, temperature-dependent Doppler effects or collisions with other atmospheric species. Our existing experimental and theoretical approaches can provide only very limited information on line-broadening effects: accurate information on pressure and temperature-dependent line shapes are limited to a small number of lines, small molecules or low temperatures. There exist promising semi-empirical methods capable of providing an extensive coverage of line shape data, which however require full integration into the modern variational methodology presented above. This is one of the directions of ExoMol in the near future.

Looking back over 10 years and 700 billion transitions, we can see that we have achieved many of our original goals. But there remains much to be done. ●

AUTHOR

Jonathan Tennyson is Massey Professor of Physics at the Department of Physics and Astronomy, University College London London WC1E 6BT, UK.

Sergei N Yurchenko is Professor of Physics, Department of Physics and Astronomy, University College London, London WC1E 6BT, UK

ACKNOWLEDGEMENTS

We owe a huge debt to the many talented scientists who have worked with us during the ExoMol project. The project has been supported by ERC Advanced Investigator Awards 267219 (2011-16) and 883830 (2020-25), as well as STFC as part of the UCL consolidated grants (ST/M001334/1 and ST/R000476/1) in the intervening years.

REFERENCES

Al-Derzi A R *et al.* 2021 *J. Quant. Spectrosc. Radiat. Transf.* **266** 107563
Berezhnoy A A *et al.* 2018 *Planet Space Sci.* **151** 27
Bernath P F 2020 *J. Quant. Spectrosc. Radiat. Transf.* **240** 106687
Campargue A *et al.* 2012 *Icarus* **219** 110
Darby-Lewis D *et al.* 2018 *J. Phys. B:At. Mol. Opt. Phys.* **51** 185701
Evans T M *et al.* 2016 *Astrophys. J.* **822** L4
Furtenbacher T *et al.* 2007 *J. Mol. Spectrosc.* **245** 115
Gordon I E *et al.* 2017 *J. Quant. Spectrosc. Radiat. Transf.* **203** 3
Greaves J S *et al.* 2020 *Nature Astron.* **5** 655
Guilluy G *et al.* 2019 *Astron. & Astrophys.* **625** A107
Hargreaves R J *et al.* 2019 *J. Quant. Spectrosc. Radiat. Transf.* **232** 35
Hooton M J *et al.* 2018 *Astrophys. J. Letts* **869** L25
Macintosh B *et al.* 2015 *Science* **350** 64
Madhusudhan N 2019 *Ann. Revs. Astron. Astrophys.* **57** 617
McKemmish L K *et al.* 2019 *Mon. Not. Roy. Astron. Soc.* **488** 2836
Nugroho S K *et al.* 2017 *Astrophys. J.* **154** 221
Owens A *et al.* 2017 *Sci. Rep.* **45068** 7
Pezzella M *et al.* 2021 *Phys. Chem. Chem. Phys.* **23** 16390
Polyansky O L *et al.* 2015 *Phys. Rev. Lett.* **114** 243001
Polyansky O L *et al.* 2017 *Mon. Not. Roy. Astron. Soc.* **466** 1363
Polyansky O L *et al.* 2018 *Mon. Not. Roy. Astron. Soc.* **480** 2597
Rey M *et al.* 2016 *J. Mol. Spectrosc.* **327** 138
Savini G *et al.* 2018 *Proc. SPIE* **9904** 99044M
Sedaghati E *et al.* 2017 *Nature* **549** 238
Snellen I 2014 *Phil. Trans. Royal Soc. London A* **372** 20130075
Sousa-Silva C *et al.* 2015 *Mon. Not. Roy. Astron. Soc.* **446** 2337
Sousa-Silva C *et al.* 2018 *Phys. Educ.* **53** 015020
Sousa-Silva C *et al.* W 2020 *Astrobiology* **20** 235
Swain M R *et al.* 2008 *Nature* **452** 329
Tennyson J 2012 *WIREs Comput. Mol. Sci.* **2** 698
Tennyson J 2019 *Astronomical spectroscopy : an introduction to the atomic and molecular physics of astronomical spectra* **3** edn. (World Scientific Singapore)
Tennyson J Yurchenko S N 2012 *Mon. Not. Roy. Astron. Soc.* **425** 21
Tennyson J Yurchenko S N 2017 *Int. J. Quantum Chem.* **117** 92
Tennyson J *et al.* 2016 *J. Mol. Spectrosc.* **327** 73
Tennyson J *et al.* 2020 *J. Quant. Spectrosc. Radiat. Transf.* **255** 107228
Tinetti G *et al.* 2018 *Exp. Astron.* **46** 135
Tinetti G *et al.* 2007 *Nature* **448** 169
Tinetti G *et al.* 2013 *Astron. Astrophys. Rev.* **21** 1
Tsiaras A *et al.* 2019 *Nature Astron.* **3** 1086
Werner H-J *et al.* 2012 *WIREs Comput. Mol. Sci.* **2** 242
Wong A *et al.* 2019 *Astrophys. J. Supp.* **240** 4
Yurchenko S N & Tennyson J 2014 *Mon. Not. Roy. Astron. Soc.* **440** 1649
Yurchenko S N *et al.* 2014 *Proc. Nat. Acad. Sci.* **111** 9379
Yurchenko S. N *et al.* 2016 *Mon. Not. Roy. Astron. Soc.* **456** 4524
Zhang Y *et al.* 2021 *Nature* **595** 370