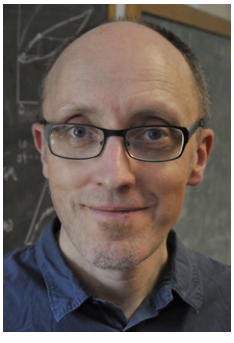
Science Progress (2016), 99(1), 1–58
Paper 1500152 doi:10.3184/003685016X14509452393033

Puzzles in bonding and spectroscopy: the case of dicarbon†

RODERICK M. MACRAE



Dr. Roderick Macrae is Lechleiter Endowed Professor of Chemistry at Marian University in Indianapolis. He received his Ph.D. in chemistry from the University of Glasgow in 1990. During his subsequent travels, he spent six years in Japan, at KEK (the High-Energy Accelerator Research Organization) and RIKEN (the Institute for Physical and Chemical Research). His research work has mainly been divided between experimental studies using the muon spin rotation method on materials ranging from fullerenes to semiconductors, and the application of computational chemistry to a variety of systems. Additionally, he is co-founder of the Institute for Green and Sustainable Science at Marian University, a research, education, and outreach centre engaging in the study of sustainability issues. He may be contacted at the School of Mathematics and Sciences, Marian University, Indianapolis, IN 46222, USA. E-mail: rmacrae@marian.edu

ABSTRACT

The unstable molecule C_2 has been of interest since its identification as the source of the "Swan band" features observable in the spectra of flames, carbon arcs, white dwarf stars, and comets, and it continues to serve as a focal point for experimental and theoretical discovery. Recent spectroscopic work has identified a quintet state of the molecule for the first time, while new insights into the bond order of C_2 in its ground state have been provided by sophisticated computational methods based on valence bond theory. This article gives a review of spectroscopic and computational work on C_2 including both historical background and the most recent discoveries.

Keywords: *Chemical bonding, quantum chemistry, spectroscopy, dicarbon*

1. Introduction.

Dicarbon, C_2, sounds like an elusive, unstable, highly reactive molecule, and indeed it is, yet practically everyone has seen it. It is the molecule largely responsible for the pale blue colour of premixed hydrocarbon flames, found on gas cookers, Bunsen burners, and oxy-acetylene cutting torches. This colour originates in the *Swan bands*, among the earliest molecular bands ever observed spectroscopically, by the Scottish physicist William Swan in 1856[1]. The role of

†Dedicated to the memory of my parents.

C_2 in flame chemistry is still not clearly understood[2]: the bands are seen even in the spectrum of methane (with one carbon atom), and in acetylene flames the two C atoms often come from different acetylene molecules[3]. (Recent kinetic modelling studies suggest that the most likely mechanism of C_2 formation in acetylene flames is the reaction between the methylene diradical and a free carbon atom[4].) Dicarbon is also an astronomically important molecule, with the majority of interstellar carbon being found in the form of free atomic carbon, C_2, or C_3[5]. The molecule plays an important role in the spectra of DQ-type white dwarf stars[6], and is a key diagnostic molecule in the study of carbon arcs[7] and comets[8].

Rather surprisingly, the bonding in this very simple molecule is still a matter of some controversy. From a naïve Lewis electron pair perspective, the eight valence electrons in the C_2 molecule might be thought capable of producing a quadruple bond. (Indeed, this is the only way in which the molecule can satisfy the "octet rule".) However, the simplest qualitative molecular orbital picture makes quite a different prediction. With the carbon 2s and 2p orbitals combining to yield σ and π molecular orbitals, the ground state valence electron configuration of C_2 is predicted to be $(\sigma_{2s})^2(\sigma^*_{2s})^2(\pi_{2p})^4$, yielding a bond order of 2 and, unusually, a double π bond with no accompanying σ bond. Figure 1 shows a qualitative MO diagram for C_2. A third qualitative approach to the bonding would be to adopt a maximum orbital overlap principle based on orbital hybridisation concepts, in which case the structure of C_2 might be predicted to resemble the diradical conceptually formed by removing the two hydrogen atoms from acetylene, with a σ bond between two sp hybrid carbons supplemented by two π bonds to produce an overall bond order of 3; spin pairing of the odd electrons (in the outward-pointing sp hybrids) then leads to a singlet ground state. Thus, seemingly reasonable arguments can be made for a bond order of 2, 3, or 4.

The experimentally-observed bond length in C_2 is 124.24 pm[9], rather shorter than an alkene double bond, but a little longer than a typical alkyne triple bond (and, indeed, than the triple bond in acetylene, which weighs in at 120.3 pm). The bond length of C_2 lies between that of the nominally singly-bonded B_2 (159.0 pm) and that of N_2 (109.768 pm), the archetypal triply-bonded molecule[9]; similarly, its bond dissociation energy of 602 kJ mol^{-1} lies between the B_2 and N_2 values, which are respectively 293 kJ mol^{-1} and 942 kJ mol^{-1} [10]. These arguments suggest a bond order somewhere between two and three. Nevertheless, a recent paper in *Nature Chemistry*[11] claims on the basis of valence bond (VB) theory and full configuration interaction (FCI) calculations that C_2 may indeed be considered to be quadruply bonded, in the sense that the theoretical analysis implies that there are four separate sets of interactions all of which have a positive contribution to the bond energy of the molecule: the σ and (two) π bonds alluded to above, *and* the interaction of the electrons in the outward-pointing sp hybrids.

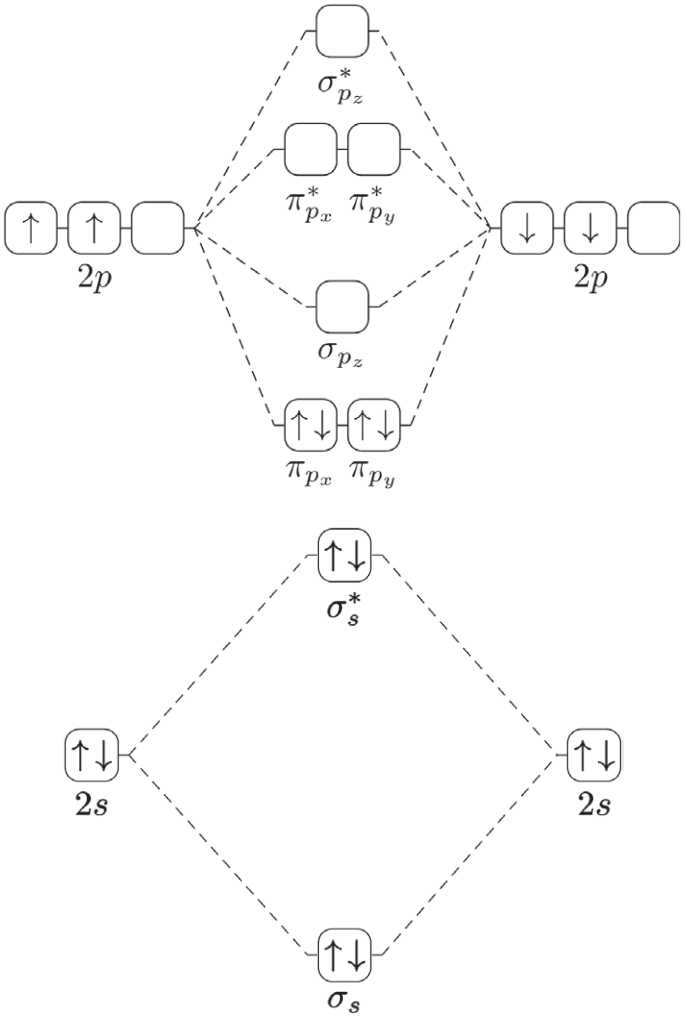


Figure 1 *Molecular orbital energy diagram for C_2. In the simple MO picture the formal bond order is 2, consisting of two π bonds and no σ bond.*

2. Experimental insights into electronic structure

2.1 The Swan bands

The electronic structure of small molecules such as C_2 is known entirely through very exacting analysis of spectroscopic data, particularly in the ultraviolet and visible regions where most electronic transitions take place. In the gas phase, molecular absorption and emission spectra exhibit fine structure corresponding to transitions between quantised vibrational and rotational substates of the electronic states participating in the transition. This can lead to great complexity in the appearance of the spectra, but it is in this complexity that the information required to reconstruct an understanding of the molecule's structure resides.

The first detailed spectral observations of C_2 – among the earliest molecular spectra ever observed – were by William Swan[1], who used a theodolite and flint-glass prism to resolve the light from "carbohydrogen" flames into its constituent wavelength components. By the term "carbohydrogen" he mainly meant hydrocarbon, though he included some compounds containing oxygen in his study. (The same spectrum had, in fact, previously been observed by Wollaston in light from the base of a candle flame[12], but Swan's was the first systematic study.) Thanks to the work of Sir Humphry Davy and John William Draper, Swan was already quite well aware of several aspects of the anatomy of such a flame: that the brightly luminescent yellow upper part of flames obtained when gases mix with air by diffusion are due to the presence of carbon particles, and possess a continuous spectrum, that flames resulting from gases premixed with air are quite different in character and exhibit spectra containing bright and dark bands, and that the innermost non-luminous portion of a flame consists of gases "not yet ignited", while the outermost envelope of the flame contains the combustion products ("carbonic acid and aqueous vapour"). In his paper, Swan described in some detail the flame of the then-new Bunsen burner (with coal gas as fuel), the spectrum of which he used as a reference point in his studies; this flame consists of a bluish-green inner cone, a diffuse outer mantle of lavender colouration, together with a "perpetual scintillation of yellow sparks" due to foreign matter suspended in the gas stream. It is the blue–green inner cone of the Bunsen flame that furnishes the line spectrum Swan used to compare with his hydrocarbon flame spectra; the latter were obtained using a stream of air from a "table blowpipe", at that time an indispensable tool for all laboratory chemists[13].

A premixed acetylene–air flame and the Swan band spectra obtained from it using a modern instrument of modest resolution can be seen in Figures 2 and 3. The Swan bands are labelled using both Swan's original notation and the modern

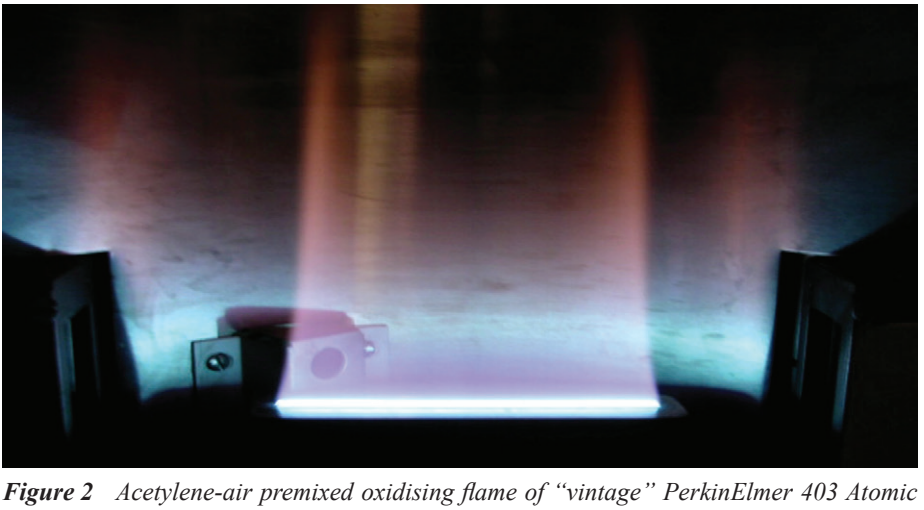


Figure 2 *Acetylene-air premixed oxidising flame of "vintage" PerkinElmer 403 Atomic Absorption spectrometer at Marian University.*

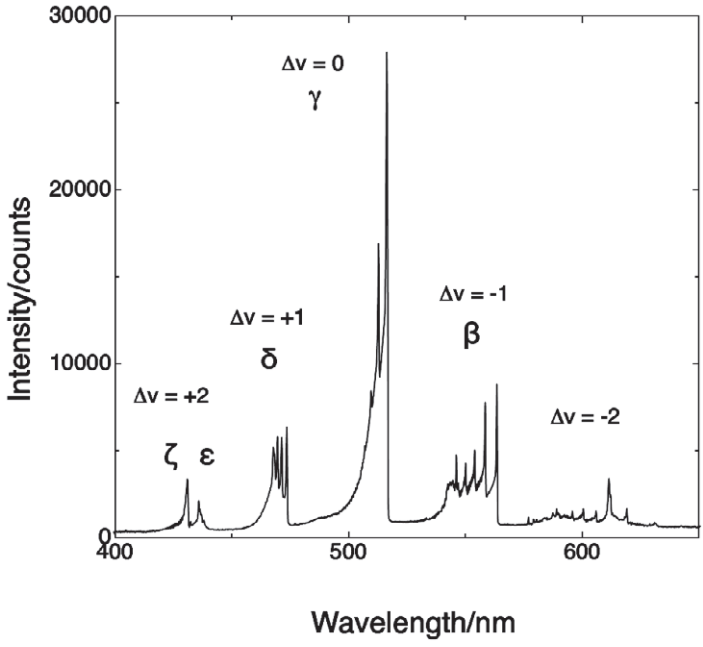


Figure 3 *Swan band spectra obtained from the flame shown in Figure 2 using an Ocean Optics HR 4000 spectrometer. Bands are labelled using both Swan's notation and modern assignments.*

assignments. The spectrum shows quite sharp features which appear in clusters; these are typically referred to as "bands" (based on their appearance when recorded on photographic plates), as opposed to the "lines" seen in atomic spectra.

Remarkably, in all the samples Swan studied, which ranged from "light carburetted hydrogen" (methane) and "olefiant gas" (ethylene) to heavier hydrocarbons such as paraffin and oil of turpentine, and drew in materials of broader composition such as methyl and ethyl alcohols and ethers, camphor, tallow, wax, and spermaceti (a wax found in the head of the sperm whale), the band spectrum observed was in every case practically identical, seemingly indicating the presence of some prominent universal species in the combustion process.

Swan's data are presented in the form of deviation angles of the light refracted by the prism, but thanks to his inclusion of the corresponding deviation angles for the Fraunhofer lines of the sun's spectrum[14], it is straightforward to convert his values into wavelengths in nanometres. Swan listed his observed wavelengths in groups labelled from α through ζ, and while the α group at 589 nm is undoubtedly the sodium D line appearing as an artefact due to foreign matter in the gas stream, and the ζ group at 431 nm belongs to a different molecule (namely CH), the β, γ, δ and ε series originating near

564 nm, 518 nm, 475 nm and 436 nm respectively are four of the bands now recognised as originating from transitions of C_2.

The first investigator to assign these spectral bands to a species of pure carbon was John Attfield, Director of the Laboratory of the Pharmaceutical Society[15]. (Swan himself attributed the spectrum to a hydrocarbon species, as all of his samples contained the elements carbon and hydrogen.) Attfield showed that the same bands were produced in the combustion not only of hydrocarbons but also of cyanogen $(CN)_2$, generated by the thermal decomposition of mercuric cyanide, and by electrical discharges in carbon monoxide and carbon disulfide; as these species have only carbon in common, Attfield argued, the spectral features must originate in some form of elemental carbon.

This did not, however, immediately settle the question of the origin of the spectral bands. For one thing, it was always *possible* that Attfield's spectra were susceptible to the presence of impurities, in particular to impurity hydrogen from water vapour. For another, Attfield made no specific claim as to whether the bands originated in atomic carbon or some molecular form of the element, as at that time the distinction between line spectra (from atoms) and band spectra (from molecules) had not been clarified. While in the years that followed the majority of spectroscopists came to conclusions that agreed with Atfield's, in 1875 a heterodox claim, that the bands were due to acetylene, was proposed by Ångström and Thalén[16] (in a paper published after Ångström's death). Their arguments were based on evidence from spark discharges created using carbon electrodes in various gases: while in a hydrogen atmosphere the Swan bands appear, in oxygen the bands seen are those of carbon monoxide and in nitrogen those of a compound of carbon and nitrogen. This work reopened the topic to debate, joined vigorously by outspoken individuals such as Charles Piazzi Smyth, professor of Practical Astronomy at the University of Edinburgh, who launched colourful polemics against the carbon theory[17]. The story of this debate is retold in entertaining detail by J.C.D. Brand in his excellent work on the history of dispersive spectroscopy[18], and he notes that "a list of those who contributed to the study of the Swan bands between 1850 and 1890 was practically a rollcall of contemporary spectroscopists." Indeed, the "last gasp" of the acetylene theory was a paper published by R.C. Johnson as late as 1927[19].

The identification of the Swan bands as originating in a *molecular* carbon species on the basis of its band spectrum had been made at least by 1914, when W. Marshall Watts asserted the position of the carbon theory proponents with the claim that the spectrum originates in a "compound of *carbon with carbon*" (italics his own)[20]. The definitive identification of the particular compound as C_2 took a little longer, and involved the participation of theory as well as experiment.

By the 1920s, improvements in dispersive instrumentation and experimental methods as well as the development of techniques for recording spectra on photographic plates had led to the availability of extraordinarily detailed Swan

band spectra: for example, in the 1927 paper mentioned above Raynor Carey Johnson (who had the unusual and perhaps dubious distinction of recognition both as a physicist and as a parapsychologist) tabulated approximately 2,000 lines in the Swan band series[19]. (This also happens to have been the last paper supporting the acetylene theory.) To achieve this degree of resolution, the spectra were generated using carbon arcs in low-pressure (less than 1,500 Pa) hydrogen, and in vacuum tubes containing trace quantities of carbon gases, diluted with inert gases to a total pressure of 2,500 – 6,500 Pa. (The dominant broadening effect on the spectra as observed in flames is due to molecular collisions, which can be limited by reducing the total pressure.)

At the same time, developments in the theoretical understanding of molecular spectra had progressed considerably. While in the late 19th century quantitative investigations of molecular band spectra were limited to the extraction of empirical mathematical relationships from the data, by 1926 it had become possible to extract information about molecular structure – such as bond lengths – from an analysis of the spectral features. Early attempts to find quantitative relationships in molecular band spectra of the kind found by Balmer in the line spectra of the hydrogen atoms were limited in their effectiveness by the presence in these spectra of *band heads* or *band edges*, which seemed to form a natural starting point for enumeration of lines but are (*vide infra*) to some extent accidental features of molecular electronic spectra. The first person to identify arithmetic relationships between spectral line positions was Henri-Alexandre Deslandres in the mid-1880s, but the significance of the band heads remained unknown for about 30 years until the underlying physical relationships began to be revealed through the work of Torsten Heurlinger, mostly done as a graduate student at Lund University, and others[18].

2.2 Vibronic spectra

The Swan band spectrum of C_2 falls into the category of molecular spectrum known as *vibronic*: the fact that the spectrum lies in the visible region owes to the characteristic energy range (from the perspective of the Planck relationship $\Delta E = hc/\lambda$, with λ between about 400 nm and 700 nm, or $\tilde{\nu}$ between about 14,000 and 25,000 cm^{-1}) of transitions between electronic energy states, while the structural features of the spectrum result from transitions involving vibrational substates of these electronic states, the energy separation between which tends to be about 10 times smaller (with typical infrared spectrometers having sensitivity over the range 400 – 4,000 cm^{-1}). At higher spectral resolution, these individual vibronic bands show finer structure originating in rotational transitions – pure rotational spectra appear in the microwave or far infrared region, so such levels are separated by amounts less than about 10 cm^{-1}.

The vibrational structure in a vibronic spectrum typically exhibits one of two basic appearances, either that of *sequences* or that of *progressions*[21], depending on the relative values of the upper and lower state vibrational

constants. Sequences and progressions are most easily characterised in terms of the vibrational quantum number, v. (In spectroscopic terminology, v' is often used to denote vibrational levels of the upper electronic state, while v'' denotes those of the lower state.) A sequence designates a group of bands having the same value of Δv: thus, for example, if the bands are labelled using the convention $v'-v''$ (that is, with upper state written first), the set $0-1$, $1-2$, $2-3$, and so forth form the $\Delta v=-1$ sequence. The clustering of spectra into sequences is typical of transitions where the upper and lower state vibrational constants are similar. A progression is a series of bands all originating (or terminating) in the same vibrational level v; for example, the set of transitions $0-0$, $1-0$, $2-0$, $3-0$, … represents a progression of absorption transitions from the $v''=0$ state. Such progressions appear, for example, in the well-known visible absorption spectrum of I_2 vapour (responsible for its purple colour), where the upper and lower electronic states (which carry the spectroscopic labels $B^3\Pi_{0u}^+$ and $X^1\Sigma_g^+$) have rather different bond lengths (of 3.016 Å and 2.667 Å respectively), leading to a room temperature spectrum dominated, thanks to the Franck–Condon principle, by transitions from the $v''=0$ and $v''=1$ substates of the $X^1\Sigma_g^+$ground state into relatively closely-spaced excited substates of $B^3\Pi_{0u}^+$ with v' ranging from about 15 to about 25. The Swan band spectrum of C_2 is a classic example of a vibronic spectrum built on sequences, with the features labelled β, γ, and δ by Swan corresponding to $\Delta v=-1$, $\Delta v=0$, and $\Delta v=+1$ respectively. (See Figure 3 for an overview of the spectrum, and Figures 4 and 5 for an explanation of the origin of the $\Delta v=-1$ band.) The Swan bands are due to transitions between states labelled $d^3\Pi_g$ and $a^3\Pi_u$.

The fine structure of vibronic spectra originates in rotational transitions, with some potential complicating factors due to spin–orbit coupling, spin–rotation interactions, and other small terms in the Hamiltonian. In the simplest case (singlet–singlet transitions) the vibronic bands have P and R branches analogous to those seen in vibrational absorption spectra in the infrared region; these correspond to transitions where $J'=J''-1$ and $J'=J''+1$ respectively, with J being the rotational quantum number. The spacing between the lines in the P and R branches is dependent upon the rotational constant, which in energy units takes the form

$$B=\frac{h^2}{8\pi^2 I}$$

where h is Planck's constant and I is the molecule's moment of inertia, defined as $I=\mu r^2$, with r the equilibrium internuclear distance (bond length) and $\mu=m_1m_2/(m_1+m_2)$ the reduced mass (equal in the case of C_2 to half of the mass of a carbon atom). If the bond length in the upper state is longer than that in the lower state (that is, if $B'<B''$) the line spacings in the R branch (higher energy) will grow progressively smaller, and the branch will eventually double back on itself forming a *band head*; if the bond length in the upper state is shorter, the band head occurs in the P branch. (The former is more common, but the

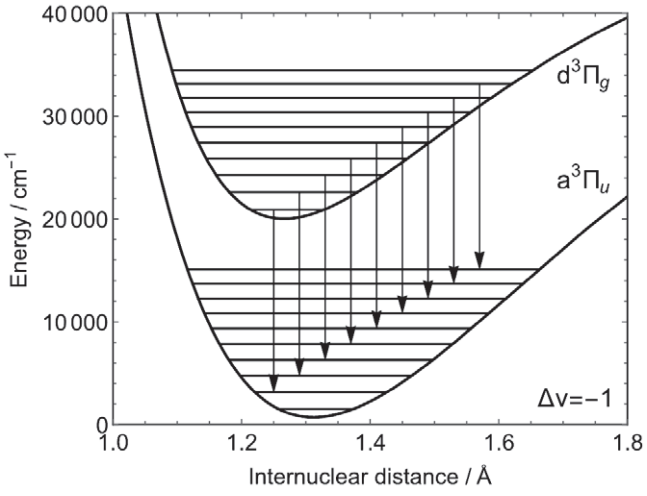


Figure 4 *Potential energy curves for the two states involved in the Swan band system, showing the set of transitions belonging to the* $\Delta v = -1$ *band. Only the first 10 vibrational levels are shown in each well. The energy curves are Morse potentials based on spectroscopic data.*

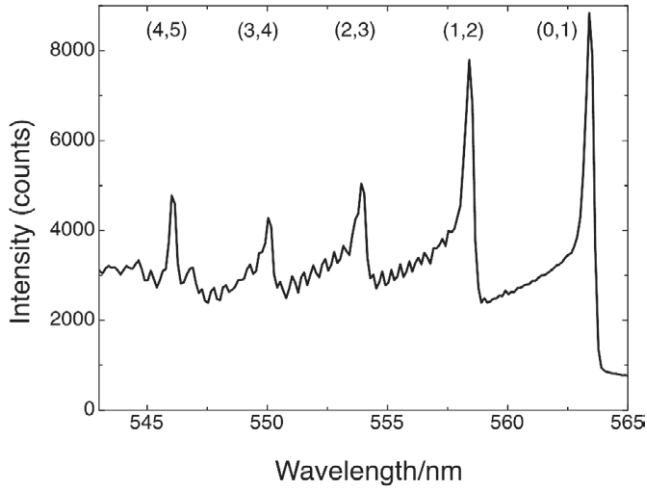


Figure 5 *Enlarged* $\Delta v = -1$ *band from Figure 3 showing labelled band heads.*

latter is the case in the Swan bands of C_2.) For convenience of analysis, the line positions of the *P* and *R* branches can be rolled together into a single data set by the adoption of a new index m, defined as $m = J'' + 1$ for the *R* branch and $m = -J''$ for the *P* branch; with this definition, the rotational fine structure lines of a vibronic band approximately obey (neglecting centrifugal distortion effects, and in wavenumber units)

$$\tilde{\nu}_{P,R} = \tilde{\nu}_0 + (B' + B'')m + (B' - B'')m^2$$

where $\tilde{\nu}_0$ is the band origin. This is the equation of a parabola (known as a Fortrat parabola), which has a minimum at $m = -(B' + B'')/2(B' - B'')$. Using tabulated values for the appropriate states of C_2 (from Huber and Herzberg[3]), the Swan band heads occur at around $J'' = 14$ in the P branch. (Band heads do not occur in infrared vibration–rotation spectra because the initial and final states both have the same rotational constant.) A simulated spectrum showing the vibrational fine structure of a simplified model electronic transition appears in Figure 6, together with the corresponding Fortrat parabola. A more realistic simulation of a Swan band transition appears in Figure 7.

The reader familiar with the main narratives in the historical development of quantum mechanics may be surprised to learn that analysis of molecular spectra using a version of quantum theory began as early as 1912, even before the publication of Bohr's model of the hydrogen atom. The prime mover in this field was Niels Bjerrum[22], compatriot (and, later in life, sailing companion) of Bohr's. While Bohr's model of the hydrogen atom is still taught in schools and universities as laying the foundations for the Heisenberg–Schrödinger theory that followed it, Bjerrum's contribution to molecular spectroscopy seems to have been largely forgotten. Curiously, both Bohr and Bjerrum had the ideas leading to their contributions while working at laboratories overseas, Bohr in Rutherford's laboratory in Manchester, UK, and Bjerrum in Walther Nernst's laboratory in Berlin, Germany.

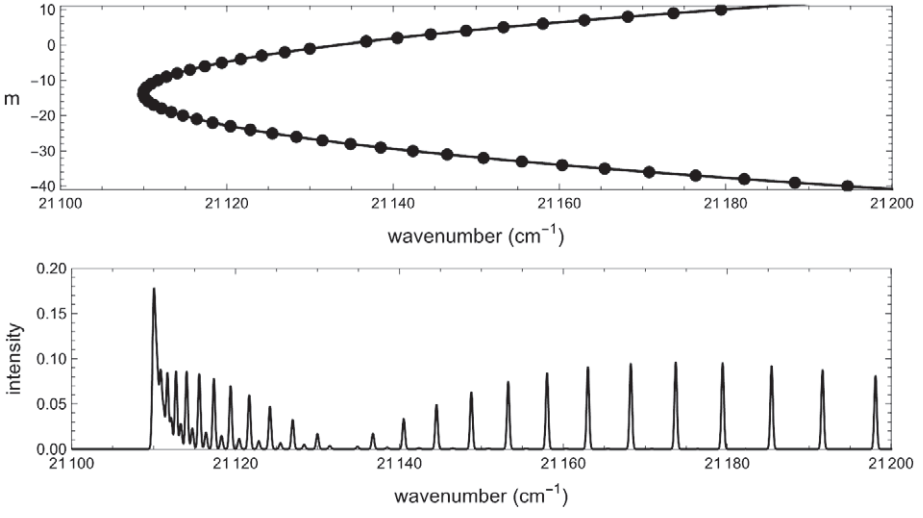


Figure 6 *(Upper) Fortrat diagram and (lower) simulated spectrum for a model singlet–singlet electronic transition. Note the band head at 21,110 cm^{-1}, the "zero gap" showing the position of the band origin at 21,133 cm^{-1}, and the "blue-degraded" character of the band, indicating that the rotational constant is larger (bond length is shorter) in the upper state. The exactly parabolic character of the Fortrat diagram is due to the neglect of centrifugal distortion.*

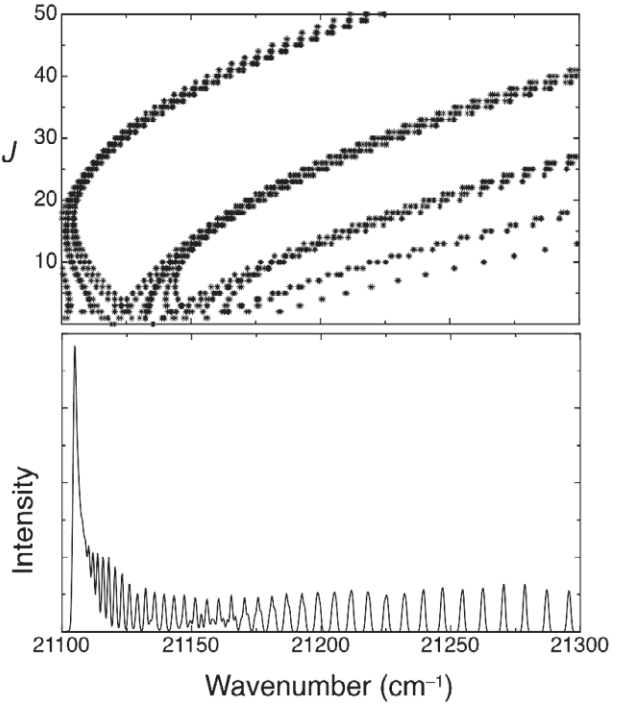


Figure 7 *(Lower) Simulation of the (1,0) Swan band transition carried out using PGOPHER[48], with parameters from Prasad and Bernath[51]. An arbitrary linewidth of 1 cm^{-1} has been applied for presentation purposes. (Upper) Substructure of the various spectral branches in the form of a Fortrat diagram.*

That molecular band spectra show structure is very clear, and by 1886 Deslandres had codified observations on visible light spectra into a set of empirical rules; among these was the observation that both the positions (in frequency units) of the successive bands and those of the lines composing the fine structure follow a pattern which can be expressed in the form of a quadratic dependence on an integer[23]. For a tidy summary (in French) of Deslandres' "laws" see the 1924 paper of R. Fortrat[24]. For a derivation of Deslandres' laws from the theory of a quantised rotating oscillator, see Chapter IX of the second English translation of Sommerfeld's *Atombau und Spektrallinien*, the "bible" of molecular spectroscopists of the time[25]. The quadratic character of the rotational fine structure is a direct consequence of the dependence of the rigid rotor energy on the rotational quantum number, while that of the band origins arises from the first anharmonic terms in the expansion of the vibrational energy in powers of $v''+1/2$ and $v'+1/2$. (Deslandres' original observations, made before the recognition of the importance of the band origins, were based upon measurements of the positions of band *heads*.) This behaviour is in sharp contrast to atomic spectra, where the lines – as in the Balmer series and similar

cases – often show series exhibiting a $1/n^2$ dependence. The approximately quadratic behaviour often observed spanning the *P* and *R* branches has been noted above. In the spectra of real molecules, where centrifugal distortion can be quite significant, the quadratic approximation breaks down and the Fortrat diagram departs from parabolic behaviour.

In Nernst's laboratory, in the wake of the first Solvay conference, the main research activity was the effort to develop a means of applying quantum theory to matter in the manner Planck had successfully done for blackbody radiation, with the goal of relating thermodynamic concepts such as heat to the atomic structure of matter. Bjerrum, as a visiting scientist from Copenhagen, devoted himself to the problem of applying dynamical concepts based on the quantum hypothesis, developed for the analysis of heat capacity problems, to the understanding of molecular (infrared) spectra. At this time the chemist's concept of a molecule was static and rigid, while physicists, inspired by Clausius's 1857 treatise on the nature of heat, were more willing to consider molecules as lively, dynamical objects, though without much of a concrete conception of molecular structure.

At the Solvay meeting, Hendrik Lorentz had already proposed a form of quantisation for a mechanical "rotator" that might be an appropriate model for a diatomic molecule in the gaseous state, later modified slightly by Paul Ehrenfest[26] (in the context of heat capacity calculations on gases) into a form almost identical to the one used today,

$$\epsilon_n = \frac{n^2 h^2}{8\pi^2 I}$$

with *h* Planck's constant, *I* the moment of inertia, and *n* the appropriate quantum number. (The modern expression for the quantum-mechanical rigid rotor replaces n^2 with $J(J+1)$, where *J* is referred to as the angular momentum quantum number.) Application of the same type of quantisation to "oscillators" (the type of model Planck used for radiation, but applicable here to molecular vibrations) led to an expression

$$\epsilon_n = nh\nu$$

again very similar to the modern quantum-mechanical expression for the energy of the harmonic oscillator, $\epsilon_v = \left(v + \frac{1}{2}\right) h\nu$, differing only by the omission of the vibrational zero-point energy term. (In fact, the question of the existence or otherwise of quantum-mechanical zero-point energy was a major thread of this period in scientific history, but it would be out of place to discuss it further here; see, for example, the article by Assmus[23] or Kragh[27].) The different dependences on the quantum number arise because the frequency of the rotator depends upon the energy, while that of the oscillator does not.

Thus Bjerrum had the tools at his disposal for the construction of a theory of molecular spectra based on a dynamical molecule, and was the first to introduce what is essentially the modern picture of a vibrating-rotating

diatomic molecule[28]. He was able to deduce a value for the average interatomic distance from the moment of inertia (not quite correct because of some invalid assumptions, but representing the first time a deduction about the physical structure of molecules had been made from spectra), and accurately modelled the fine structure of von Bahr's 1913 vibrational spectrum of HCl (the first to resolve rotational structure in the *P* and *R* branches) using a value for *I* deduced from the shape of the band envelope. In addition to this, he correctly interpreted the progressions of equally spaced lines observed in microwave spectra as being due to molecular rotation, and (with von Bahr) used the information from the infrared spectrum of HCl to assign several lines in the microwave spectrum. (See Brand's book for a more detailed discussion[18].)

A fundamental weakness of Bjerrum's theory, and the reason it is typically referred to as a semiclassical theory rather than a true quantum theory, is that the frequencies observed in spectra were directly identified with frequencies of molecular vibrational or rotational motion, rather than being recognised as energy differences between quantised states related by the emission or absorption of a photon. This distinction was finally noted in 1920 by Edwin Kemble[29], thus adding a further component of the modern picture of molecular spectra. (In the intervening period, Karl Schwarzschild[30] proposed a "dual" picture in which the vibronic spectra in the visible and ultraviolet regions obeyed a Bohr energy condition – in other words, behaved quantum-mechanically – while the spectra observed in the infrared and microwave regions directly corresponded to the vibrational and rotational frequencies of the molecule in the manner suggested by Bjerrum.) Eventually, and with some awkwardness, the Bohr theory was applied to the whole molecular spectrum[18,23], helped by the notion of transitions limited to adjacent states (later justified by *selection rules* in modern quantum theory).

As a result of these theoretical developments, an analysis of the Swan bands by Shea appearing in 1927[31] uses terminology quite similar to current terminology and is relatively straightforward to interpret. Thus the vibrational quantum number assignments given in his Table I (reproduced as Table 1 here) are fully correct (even down to labelling the lowest state with a quantum number of zero). The Swan band series ($\Delta v = 0, \pm 1, \pm 2, \ldots$) can be reconstructed by reading wavelengths diagonally in this table.

Shea's detailed analysis of the *P* and *R* branches of the Swan bands (founded on work of Birge, Kemble, and others) is based on a power series expansion of the force between the two atoms, and an assumed form for the energy levels expressed – again as a power series – in terms of an angular momentum quantum number *m*. The most probable values of the unknown parameters were then obtained by a method of successive approximation – a challenging manual computation at the time. This analysis resulted in values for the moment of inertia of the upper and lower states corresponding to internuclear distances of 126.1 pm and 130.7 pm respectively, close to the values tabulated by Huber and Herzberg[9].

Table 1 *Band assignment table for the Swan system of C_2, adapted from ref. 31; wavelengths are in nm*

v'	$v''=0$	1	2	3	4	5	6
0	516.5	563.5	618.8				
1	473.7	512.9	558.5	612.0			
2	438.1	471.6	509.6	554.1	606.0		
3		437.1	469.8		550.2	600.5	
4			436.4	468.3		547.3	595.8
5					467.3		

2.3 Modern quantum theory of diatomic molecular spectra

The key works on the spectroscopic analysis of diatomic molecules are those by Herzberg[32], Lefebvre-Brion and Field[33], and Brown and Carrington[34], all of which have gone through many editions. A useful brief introduction to the electronic spectroscopy of diatomic molecules can be found in the chapter by Peter Bernath[35] in the *Handbook of molecular physics and quantum chemistr*y.

A discussion of spectroscopic analysis of diatomic molecules should begin with *term symbols* which are used to label electronic states. These are based on the symmetry of the occupied orbitals, and have their origin in group theory, as well as in the quantum theory of angular momentum. (A similar system exists in the spectroscopy of atoms and ions.) Homonuclear diatomic molecules belong to the point group $D_{\alpha h}$, and so the electrons occupying the molecular orbitals may be labelled according to the irreducible representations of that group. These labels are closely related to the usual σ and π MO designations – see Figure 1 – with σ bonding and antibonding orbitals having σ_g and σ_u symmetry, and π bonding and antibonding orbitals having π_u and π_g symmetry. (The symbols σ and π are molecular analogues of the atomic orbital labels s and p, and the subscripts *g* and *u* refer to symmetry or antisymmetry with respect to inversion through the molecule's centre.)

A given electron *configuration* can give rise to multiple *states* depending on the manner in which the spin and orbital angular momenta of the electrons combine. These can be distinguished through the use of term symbols. For example, the ground state electron configuration of O_2 can be written

$$(\sigma_g 1s)^2 (\sigma_u^* 1s)^2 (\sigma_g 2s)^2 (\sigma_u^* 2s)^2 (\sigma_g 2p)^2 (\pi_u 2p)^4 (\pi_g^* 2p)^2$$

but this configuration does not unambiguously define an electronic *state*, as the two most energetic electrons can be singlet-paired in one of the two degenerate π^* orbitals, or can occupy separate π^*orbitals with spins either paired or unpaired. The ground state is known to be that with spins unpaired. (O_2 is paramagnetic.) The overall term symbol for a given state has the form $^{2S+1}\Lambda$, where *S* is the total spin, and Λ is a Greek capital letter representing the total orbital angular momentum around the bond axis, which is 0 for σ-orbitals, ± 1 for π-orbitals, *etc*. A total orbital angular momentum of 0 corresponds to a Σ

state, 1 to a Π state, 2 to a Δ state, and so forth. To obtain the term symbol for a given state one need only consider the incompletely filled orbitals[36]. In the ground state of O_2 the two electrons have parallel spin, leading to a triplet state ($2S+1=3$), and are in separate π^* orbitals with orbital angular momenta of $+1$ and -1 (in units of ħ) around the internuclear axis, leading to $\Lambda=0$ and a state designated $^3\Sigma$. To obtain the full symbol a subscript and superscript are added indicating the symmetry under inversion ($(+1)\times(+1)=(+1)$, leading to a g subscript) and reflection in a particular plane that contains the internuclear axis ($(+1)\times(-1)=(-1)$, leading to a "–" superscript), yielding $^3\Sigma_g^-$ as the overall term symbol.

Applying a similar reasoning to an electron configuration of C_2 constructed according to the *Aufbau* principle (based on the MO energy level diagram illustrated in Figure 1) leads to a ground state term of $^1\Sigma_g^+$. The observant reader will note that this is *not* the symbol of the lower state involved in the transition leading to the Swan bands, which connect states with term symbols $^3\Pi_g$ (upper) and $^3\Pi_u$ (lower). It should be emphasised that the *Aufbau* principle is mainly an empirical rule of thumb, and that it is quite possible that electron–electron repulsion might be lower in a state that violates the *Aufbau* principle[37]; however, in this case it is true that the Swan bands, which dominate the emission spectrum of C_2 under most conditions, originate in a transition between two excited states of the molecule. Before this was eventually noted, in 1959, by Edward Ballik (who was at the time of the research an undergraduate summer student) and Donald Ramsay[38] in a letter to the editors of the *Journal of Chemical Physics*, the Swan band transitions were incorrectly labelled $A^3\Pi_g - X^3\Pi_u$ in textbooks and papers; their correct designation is $d^3\Pi_g - a^3\Pi_u$. (The label X is assigned to the ground state. Excited states of the same spin as the ground state are labelled with uppercase letters A, B, C, …, while those with different spin – triplet if the ground state is singlet, singlet if it is triplet – are labelled with lowercase letters a, b, c, …. The change in labelling of the states of C_2 indicates a change from a picture in which the ground state was thought to be a triplet to one in which it is recognised to be a singlet.)

Identification of the electronic configurations corresponding to the upper and lower Swan band states relied on quantum-mechanical electronic structure theory, to be discussed in a later section of this article. The electronic configurations and term symbols of the excited states (showing valence electrons only) are

$$(\sigma_g 2s)^2\,(\sigma_u^* 2s)^2\,(\pi_u 2p)^3\,(\pi_g^* 2p)^1 \qquad d\,^3\Pi_g$$

$$(\sigma_g 2s)^2\,(\sigma_u^* 2s)^2\,(\pi_u 2p)^3\,(\sigma_g 2p)^1 \qquad a\,^3\Pi_u$$

Within the Born–Oppenheimer approximation separating nuclear and electronic motions, each molecular electronic state corresponds to a different internuclear potential, leading to a different equilibrium internuclear distance r_e

(internuclear distance at the minimum of the potential energy curve), different harmonic and anharmonic force constants, and a different set of quantised vibrational and rotational energy levels.

In general, the mathematical form of the internuclear potential is not known. Approximate forms such as the Morse potential[39] can be used (with an additional term for the centrifugal potential added to accommodate molecular rotation), or more generally the potential can be expanded in a power series around its minimum in the manner introduced by James Lawrence Dunham[40] (only about a year before his sudden death at the age of 29 as he was recovering at home after surgery)

$$U(x) = a_0 x(1 + a_1 x + a_2 x^2 + \ldots)$$

where $x = r - r_e$, and the Schrödinger equation solved approximately to yield a power series expression for the vibration–rotation energy levels. For the simplest case (a $^1\Sigma^+$ state) this takes the form

$$E_{vJ} = \sum_{ij} Y_{ij} \left(v + \tfrac{1}{2}\right)^i [J(J+1)]^j$$

where the $v + \frac{1}{2}$ and $J(J+1)$ terms represent the quantum number dependences of the solutions of the Schrödinger equation for the harmonic oscillator and rigid rotor. The Dunham coefficients Y_{ij} form a tidy notation for expression of the traditional terms used in vibration–rotation analysis; for example, Y_{01} is the rotational constant B_e, Y_{10} is the harmonic frequency ω_e, while other terms represent vibrational anharmonicity, deviations from rigid rotor behaviour, and various orders of vibration–rotation coupling. In the simplest case where all states are $^1\Sigma^+$ states, within the Dunham analysis, the energy of a given "rovibronic" state with respect to the bottom of the potential energy curve of the ground state is

$$E_{nvJ} = T_e(n) + \sum_{ij} Y_{ij} \left(v + \tfrac{1}{2}\right)^i [J(J+1)]^j$$

where $T_e(n)$ represents the shift of the potential energy curve of electronic state n with respect to the ground state. In cases other than $^1\Sigma^+$, a variety of other terms contribute to the energy of the rovibronic state. In singlet states of nonzero Λ, the phenomenon of "lambda-doubling", originating in the coupling of the rotational motion of the nuclei with the electronic orbital angular momentum $\mathbf{\Lambda}$, introduces a small splitting of the form $\Delta\upsilon = qJ(J+1)$. In states of nonzero spin, the situation is more complicated, with contributions from electronic spin, electronic orbital angular momentum, and angular momentum due to nuclear rotation as well as vibration–rotation interactions all contributing to the multiplet structure of the rovibronic levels in a manner dependent on the relative magnitudes and orientations of these motions. The priority order of the different sources of angular momentum in the system determines the choice

of "good quantum numbers" for analysis. Several idealised extreme cases of coupling were described by Friedrich Hund in 1933[41], and are still known as "Hund's cases". In realistic cases, all relevant terms must be considered and the Hamiltonian matrix diagonalised numerically. Detailed treatments of several cases are given in Herzberg[32], Lefebvre-Brion and Field[33], and Brown and Carrington[26].

While the Dunham analysis is systematic, it can be unwieldy for practical purposes of spectroscopic line fitting. An alternative approach, using another common type of notation, is to consider the energy of a given molecular state as a sum

$$E(v,J) = T_e + G(v) + F(J)$$

where T_e is the energy of the electronic state (at the bottom of the relevant potential energy curve), and $G(v)$ and $F(J)$ are the vibrational and rotational contributions to the state's energy, usually expressed (as in the Dunham analysis) as expansions in terms of $v + \frac{1}{2}$ and $J(J+1)$. The situation is complicated by vibration–rotation coupling (centrifugal force in higher rotational states leads to an increase in the bond length), expressed in the Dunham analysis through Y_{ij} with both i and j nonzero, and so a typical approximate energy expression might be expressed as

$$E(v,J) = T_e + \omega_e\left(v+\tfrac{1}{2}\right) - \omega_e x_e\left(v+\tfrac{1}{2}\right)^2 + B_e J(J+1) - D_e[J(J+1)]^2 + \alpha_e\left(v+\tfrac{1}{2}\right)J(J+1)$$

in the most common spectroscopic notation. Higher terms can be added to the expansion if needed. The variables (and state "energies" themselves) are usually expressed in units of cm^{-1}. Transitions between states then occur at "frequencies" (in cm^{-1})

$$\nu(v',J',v'',J'') = E(v',J') - E(v'',J'')$$

where the single and double primes refer to the upper and lower state, respectively, and with intensities that are subject to the appropriate selection rules.

In the analysis of high-resolution spectra containing rotational and other fine structure (Λ-doubling, spin–orbit coupling, spin–spin coupling, hyperfine coupling, *etc*.), it is common to consider a vibrational term energy T_v for a given vibrational quantum number v; the "band origin" of a given vibronic band is then given by the difference between T_v values for the upper and lower states, while the fine structure of the upper and lower states can be expressed in terms of spectroscopic constants valid for that particular value of v, for example B_v, D_v, and H_v (linear, quadratic, and cubic rotational constants), A_v and A_{Dv} (spin–orbit coupling terms), λ_v (spin–spin coupling), various Λ-doubling constants (o, p, q, and higher terms), and so on. As the most appropriate Hund's coupling case to describe the molecular system varies as a function of

J (because of the changing relationship between the contributions to angular momentum from overall rotation and from electronic spin and orbital angular momentum), it is typical to construct a Hamiltonian in the basis appropriate to one particular case and then allow for off-diagonal terms. For the Swan bands of C_2, Hund's (a) case ($A \gg BJ$, *i.e.* spin–orbit coupling energy larger than rotational energy) is appropriate for small values of *J*, with a gradual transition to a situation intermediate between Hund's (a) and (b) cases (where in the (b) case $BJ \gg A$) occurring with increasing *J*. (The other Hund cases, (c), (d), and (e), are not relevant here.)

For the $^3\Pi$ states involved in the Swan band transitions, approximate rotational term energy expressions were first derived by A. Budó[42] in 1936. A slightly modified version of these was used in the analysis of the Swan band system by John G. Phillips and Sumner Davis published, in book form, with a list of spectral line positions and assignments, by the University of California press[43]. The quantum numbers in this analysis are based on the Hund (b) case in which electronic orbital angular momentum **Λ** (around the intermolecular axis) and molecular rotational angular momentum **R** (of the axis, perpendicular to **Λ**) couple to yield a "total angular momentum without spin" *N*, which then couples with electron spin S to produce the total angular momentum **J**. Thus each value of *N* leads to three components with $J=N+1$, $J=N$, and $J=N-1$ respectively. The term expressions of Phillips and Davis (subject to a a typographical correction from the original printed form) are

$$F_3(J) = B_v \left(J(J+1) + [\, y_1 + 4J(J+1)]^{1/2} - \frac{2}{3} \frac{y_2 - 2J(J+1)}{y_1 + 4J(J+1)} \right) + D_v \left(J + \frac{3}{2} \right)^4$$

$$F_2(J) = B_v \left(J(J+1) + \frac{4}{3} \frac{y_2 - 2J(J+1)}{y_1 + 4J(J+1)} \right) + D_v \left(J + \frac{1}{2} \right)^4$$

$$F_1(J) = B_v \left(J(J+1) - [\, y_1 + 4J(J+1)]^{1/2} - \frac{2}{3} \frac{y_2 - 2J(J+1)}{y_1 + 4J(J+1)} \right) + D_v \left(J - \frac{1}{2} \right)^4$$

where

$$y_1 = Y(Y-4) + \frac{4}{3}$$

$$y_2 = Y(Y-1) - \frac{4}{9}$$

and $Y = A/B_v$, where *A* is the spin–orbit coupling constant and B_v is the rotational constant. The cases $A>0$ and $A<0$ (which is the case here) lead to "normal" and "inverted" energy order of the states F_1, F_2 and F_3, corresponding to $J=N+1$, $J=N$, and $J=N-1$. This expression omits both spin–spin and spin–rotation coupling, as well as Λ-doubling. Treatments of spin–spin and spin–rotation coupling can be found, for example, in Brown and Carrington[34], while expressions for Λ-doubling in the case of $^3\Pi$ states were derived in 1979 by Brown and Merer[44].

In the case of a homonuclear diatomic molecule such as C_2, the overall molecular wavefunction possesses a definite *parity*; that is, it is either unchanged or undergoes a sign change upon inversion of the spatial coordinates through the origin. Rotational wavefunctions (based upon the quantum-mechanical rigid rotor model) have even parity (remain unchanged) for even values of J and odd parity (change sign) for odd values of J. Additionally, in the case where $\Lambda \neq 0$ (as in the Π states here), Λ -doubling leads to a contributing factor which is respectively odd or even for each J. The overall parity is given by the product of these individual parities in accordance with the rules: odd × odd = even, even × even = even, odd × even = odd. Overall even and odd parity are designated respectively by + and –. The parity selection rule for dipole transitions connects states of overall "+" parity only with states of overall "–" parity. Thus, as noted by Phillips and Davis, the Swan bands of C_2 exhibit a "stagger" pattern in alternate members of branches because all the lines of even J belong to one Λ-doubling sub-branch while all the lines of odd J belong to the other.

While, in principle, P, Q, and R branches are all allowed for $^3\Pi - {}^3\Pi$ transitions, Q branches are rarely observed except in cases where the region of the "zero gap" between P and R branches is very uncluttered, as transition intensities drop off rapidly with increasing J. (In transitions where $\Delta\Lambda \neq 0$, however, the Q branch is approximately twice as strong as the P and R branches.)

More quantitatively, the intensity of a vibronic (emission) transition is a product of several factors: the population distribution of the initial state(s), the transition frequency (a factor of $\nu^4_{J'J''}$ originating in the Einstein A coefficient expressed in power units), the intrinsic strength of the electronic transition (the square of the *transition dipole moment* $\mathbf{R}_e$), the overlap between the initial state and final state vibrational wavefunctions (the Franck – Condon factor $q_{v',v''}$), and a rotational line strength term known as a Hönl – London factor ($S^{\Delta J}_{J''}$). The overall expression[45] takes the form

$$P_{J'J''} = \frac{16\pi^3}{3\varepsilon_0 c^3} \frac{N_{J'}}{(2J'+1)} \nu^4_{J'J''} q_{v',v''} |\mathbf{R}_e|^2 S^{\Delta J}_{J''},$$

and the appropriate Hönl – London factors for triplet – triplet transitions were first worked out by Budó[46], and are treated in considerably more detail in the book by Kovács[47]. Where the upper and lower states are in the the limit of Hund's (a) or (b) case and smaller terms are omitted from the molecular Hamiltonian the expressions are quite straightforward, but in the more generally applicable intermediate case they grow considerably more complicated, and are again parameterised in terms of Y, the ratio of the spin – orbit coupling constant and the rotational constant. In the intermediate case intensities of a total of 27 dipole-allowed spectral branches are computed for each J (each, in turn, subject to Λ-doubling); transitions can be labelled according to a system where, for example, $P_1(J)$ corresponds to a $\Delta J = -1$ transition between upper

and lower states both of type F_1 (that is, with $J=N+1$), while a transition between the F_2 level of the upper state and the F_1 level of the lower state with $\Delta J=-1$ (remembering that ΔJ refers to the change in J on going from the *lower* to the *upper* state, regardless of the actual direction of the transition) has $N=J$ in both the upper and lower state and so can be labelled ${}^{Q}P_{21}(J)$. (Transitions of the latter type, where $\Delta N \neq \Delta J$, are known as satellite branches, and are usually weaker than branches where $\Delta N=\Delta J$. They are strictly forbidden in the Hund's (a) case.)

Modern computer programs such as PGOPHER[48] utilise a high-quality model Hamiltonian to fit or simulate spectroscopic data to an almost arbitrary degree of accuracy. Using PGOPHER, Brooke *et al*.[49] reanalysed several merged data sets on the Swan system of C_2 (from the work of Curtis and Sarre[50], Prasad and Bernath[51], Lloyd and Ewart[52], Tanabashi[53], and Bornhauser[54,55]) and obtained the most accurate molecular constants currently available (See simulated transition in Figure 7).

2.4 Other series, perturbations, and the true ground state of C_2

While the Swan bands dominate the emission spectrum of C_2, multiple other vibronic band systems have been identified spectroscopically ranging in energy from the infrared to the ultraviolet. Known bands involving the singlet and triplet manifolds of C_2 are summarised in Figure 8 and Table 2. Very recently Peter Radi and his co-workers at the Paul Scherrer Institute (PSI) in Switzerland identified the first transitions known to occur between two *quintet* states of the molecule[62] *via* a sophisticated four-wave mixing laser double resonance

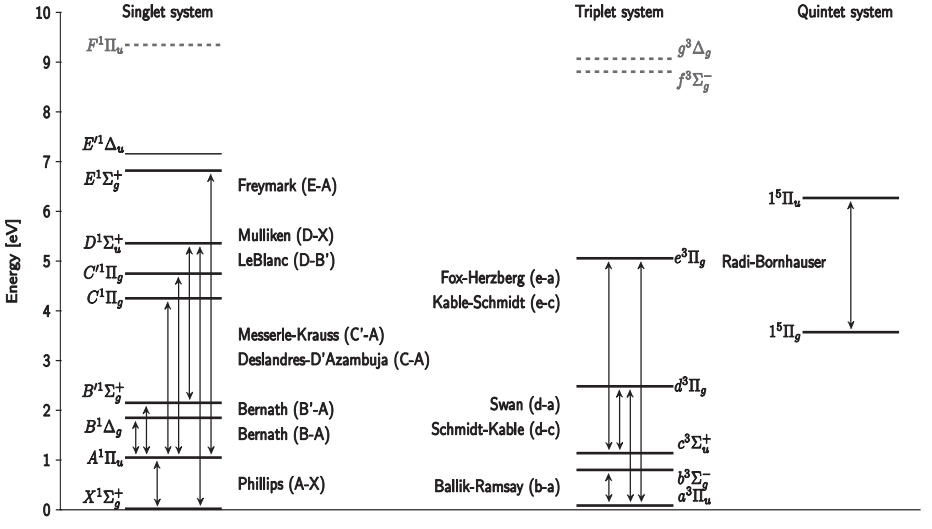


Figure 8 *States in the singlet and triplet manifolds of C_2 together with observed bands. Energies are electronic term energies T_e. States shown in grey (dashed lines) are thought to be Rydberg states.*

Table 2 *Spectroscopic systems of C_2.*

System	States involved	Energy (cm^{-1})	Observation conditions
Swan	$d\,^3\Pi_g - a^3\Pi_u$	19378[49]	Bunsen flame, discharge tube with hydrocarbon vapours, carbon arc
Fox–Herzberg	$e\,^3\Pi_g - a^3\Pi_u$	39807[57]	Discharge through He containing benzene vapour
Ballik–Ramsay	$b\,^3\Sigma_g^- - a\,^3\Pi_u$	5632[60]	Emission from carbon-tube furnace
Kable–Schmidt (e–c)	$e\,^3\Pi_g - c\,^3\Sigma_u^+$	31144	Discharge through acetylene/Ar followed by jet expansion and LIF
Schmidt–Kable (d–c)	$d\,^3\Pi_g - c\,^3\Sigma_u^+$	10715	Discharge through acetylene/Ar followed by jet expansion and LIF
Freymark	$E\,^1\Sigma_g^+ - A\,^1\Pi_u$	46668[61]	Discharge through acetylene
Mulliken	$D\,^1\Sigma_u^+ - X\,^1\Sigma_g^+$	43227[58]	Emission in carbon arc, oxy-acetylene flame, hydrocarbon discharge
LeBlanc	$D\,^1\Sigma_u^+ - B'\Sigma_g^+$	28033[58]	Two-photon dissociation of C_2H_2 in supersonic jet followed by LIF
Messerle–Krauss	$C'\,^1\Pi_g - A^1\Pi_u$	31022[59]	Weak, seen through perturbations of Deslandres-D'Azambuja
Deslandres–D'Azambuja	$C\,^1\Pi_g - A^1\Pi_u$	25969[59]	Condensed discharge through CO, CO_2, or C_2H_2, carbon arc at high T
Phillips	$A\,^1\Pi_u - X\,^1\Sigma_g^+$	8268[60]	High-current discharges, flames, arcs
Bernath B′	$B'\,^1\Sigma_g^+ - A\,^1\Pi_u$	6928[60]	Hydrocarbon discharges
Bernath B	$B\,^1\Delta_g^+ - A\,^1\Pi_u$	3590[60]	Hydrocarbon discharges
Radi–Bornhauser	$1\,^5\Pi_u - 1\,^5\Pi_g$	21791[62]	Discharge, double resonance

Notes on observation conditions are mainly drawn from Pearse and Gaydon[56] and Wallace[57]. Energies are of 0–0 transition band origins, usually computed from global spectral fits. (Note that the 0–0 transition may not be prominent in some band systems due to Franck–Condon factors if the bond lengths in the two electronic states are very different.) The Schmidt (d–c) band origin is estimated from the difference between Brooke's value for the Swan band 0–0 line[49] and Joester's reported value for T_0(c–a),[63] while the Kable (e–c) band origin is similarly estimated from the difference between the origin of the Fox–Herzberg band and the same T_0(c–a) value.

spectroscopy experiment utilising a molecular beam. An initial ("pump") laser pulse excites molecules from the $v=5$ vibrational level of the $a^3\Pi_u$ state (lower state of the Swan band transition) to the $v=6$ level of the $d^3\Pi_g$ state. This state overlaps strongly with the $v=0$ level of the $1^5\Pi_g$ state, giving it mixed triplet-quintet character, and thus acts as a "gateway" state to the quintet manifold. A second ("probe") pulse excites the 0–0 line of the $1^5\Pi_u - 1^5\Pi g$ transition near 21,800 cm^{-1} (459 nm), in the violet range of the visible spectrum. The "signal" photon, with the same energy as the "probe" photon, is detected along a specific direction (determined by the phase relationship between laser beams) perpendicular to the propagation direction of the molecular beam, thus effectively eliminating Doppler broadening.

The history of observation of the vibronic bands of C_2 has run in parallel with that of increasing sophistication in experimental methodology and spectroscopic technique. After the discovery of the Swan bands, the first other bands to be documented were probably those of the Deslandres–D'Azambuja system, a relatively weak system lying mostly in the near-ultraviolet, generated in electrical discharges through carbon monoxide or dioxide; these were first reported by Henri Deslandres and Lucien, D'Azambuja in 1905[64]. This set of bands is now understood to correspond to a series of transitions in the singlet manifold, linking the $A^1\Pi_u$ and $C^1\Pi_g$ states. Other early band systems include those named for Mulliken[65], and Fox and Herzberg[66], both in the ultraviolet, with 0–0 lines near 231 nm and 251 nm, respectively, and a near-infrared band system with a 0–0 transition at 1209 nm identified by John Gardner Phillips[67], then of the Yerkes observatory of the University of Chicago, where he was a graduate student of Gerhard Herzberg; all of these were observed in discharges, and fell within a range of wavelengths that permitted them to be recorded on suitably sensitised photographic plates. Similar spectroscopic methods were used by Freymark in 1951 to detect a set of bands lying somewhat farther in the ultraviolet than the Mulliken bands, with a 0–0 transition at 214 nm[61]. In 1958, one year before their letter to JCP identifying the correct ground state of C_2, Donald Ramsay and his undergraduate research student Edward Ballik had already made a significant contribution to the spectroscopy of C_2 though the identification of a new band system, farther in the infrared than the Phillips system, in the emission spectrum of a carbon furnace[68]; this system has its 0–0 transition at a wavelength of 1,769 nm. In this case, recording of the spectrum necessitated the use of a cooled PbS photoconductive detector.

High-resolution vibronic spectra often show *rotational perturbations* in the fine structure of a band; one line, or a series of lines in a particular narrow range of frequencies, can deviate from the orderly progression predicted by a Dunham-type analysis (Figures 9 and 10). Such perturbations can lead to displacement of lines from their expected positions, alteration in their intensities from theoretical predictions, or changes in their structure, for example *via* splitting of individual lines into doublets[32]. Matching perturbations are typically seen in the *P* and *R*

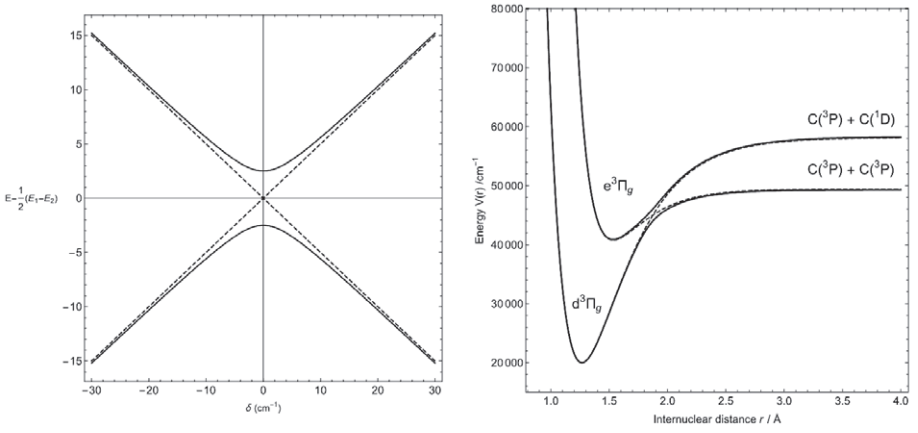


Figure 9 *(Left) Perturbation of two energy levels as a function of the unperturbed levels. The perturbed levels are represented as solid curves, the unperturbed levels as dashed lines (after Figure 134 in Herzberg[32]). (Right) Non-crossing of energy curves for states of the same symmetry. The curves shown as examples are for the* $d^3\Pi_g$ *and* $a^3\Pi_g$ *states of* C_2*, following the approach of Ballik and Ramsay[69].*

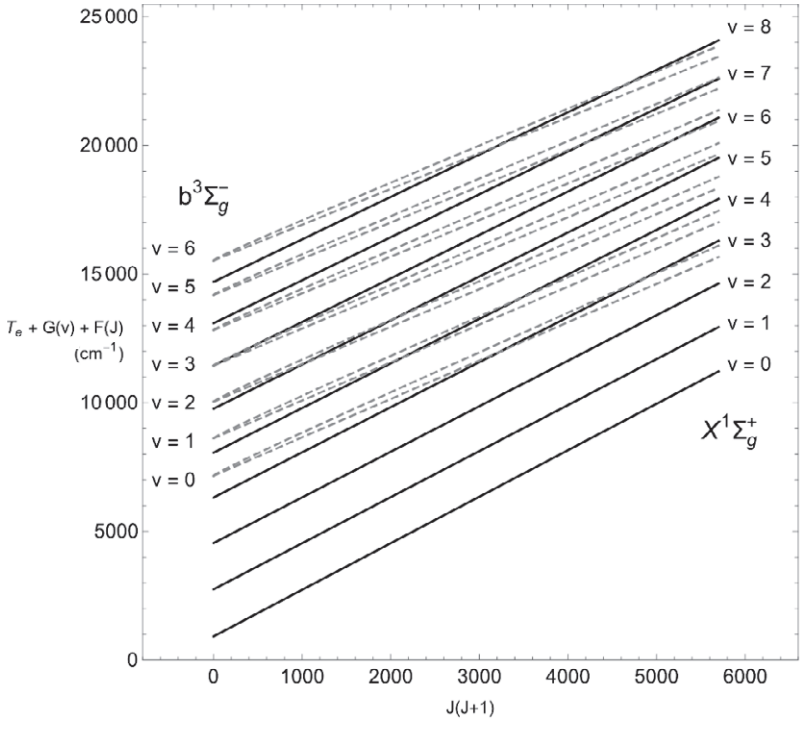


Figure 10 *An approach to predicting rotational perturbations – vibrational-rotational terms of two electronic states are plotted as a function of* J(J + 1)*; perturbations are predicted to occur at intersection points, subject to selection rules. Here, the states are* $X^1\Sigma_g^+$ *(solid black lines) and* $b^3\Sigma_g^-$ *(dashed grey lines). After Ballik and Ramsay[74].*

branches at the same J' or J'' values. This phenomenon occurs when rotational series from two different states of appropriate characteristics intersect. When the states approach in energy the approximation that the two states are non-interacting and independent loses its validity because the wavefunctions of the members of the two series *mix*. Such perturbations can occur if the states have the same total angular momentum J (not to be confused with the rotational quantum number, also called J) and symmetry. (Weaker rules also apply to enforce the same multiplicity (*i.e.* spin), and electronic orbital angular momentum quantum number Λ differing by at most 1, where these are well-defined quantities.) Vibrational perturbations, affecting the progression of an entire band, are also possible. These occur when the potential energy curves for two different electronic states intersect.

It was through such rotational perturbations that Ballik and Ramsay discovered that the true ground state of C_2 was the lowest singlet state, rather than the lowest triplet state[38,69]. Previous opinion that the $a^3\Pi_u$ state was the ground state appeared to be supported by the sheer ubiquity of the Swan bands, as well as the fact that they could apparently be seen fairly easily as absorptions (implying a high population of the lower level), together with specific observations, including long-lived Swan band emissions seen to persist for several hours in C_2 deposited in an inert gas matrix at 4.2 K[70] and cometary Swan band emissions thought to be a resonance fluorescence phenomenon[71] (indicating that the molecules were already in the lower state of the transition when then absorbed the solar photon). Ballik and Ramsay suggest that the cometary emissions can either be accounted for through a very long lifetime of the $a^3\Pi_u$ state or through a continual repopulation of this state *via* an optical pumping mechanism from the $X\,^1\Sigma_g^+$ state; later two-photon excitation measurements by Bondybey indicate a lifetime of about 65 μs for the $v=0$ vibrational state of $a^3\Pi_u$ in inert gas solid matrices, suggesting that the latter explanation may be more likely, though the gas phase radiative lifetime may be considerably longer than this[72]. Additionally, some bands previously seen in absorption and attributed to the C_2 Swan system were later correctly identified by Dolphus Milligan and Marilyn Jacox as transitions originating in the Herzberg – Lagerqvist system of the C_2^- anion[73].

The perturbations seen by Ballik and Ramsay affect the rotational fine structure of the $b\,^3\sum_g^- - a\,^3\prod_u$ bands that now bear their name, with corresponding perturbations equal in magnitude and opposite in direction affecting transitions in several bands of the $A\,^1\prod_u - X\,^1\sum_u^+$ Phillips system at the same energy. These indicate that the sequences of rotational sublevels of the $b\,^3\sum_g^-$ and $X\,^1\sum_g^+$ states intersect (Figure 10); from the extracted spectroscopic constants, Ballik and Ramsay were able to compute the energies of the $v=0$, $J=0$ levels of the $a^3\Pi_u$ and $X\,^1\sum_g^+$ states (with some averaging over fine splittings in the former case) and determine that the $X\,^1\sum_g^+$ is lower in energy by about 610 ± 5 cm^{-1} [69].

Rotational perturbations of a similar kind were found in 1967 by G. Messerle and L. Krauss[59] in the Deslandres–D'Azambuja bands which link the $C^1\Pi_g$ and $A^1\Pi_u$ states in the singlet manifold. In this case the potential energy curve of the $C^1\Pi_g$ state undergoes an avoided crossing with a second state of the same symmetry, now labelled $C'^1\Pi_g$, in such a way that the $v=3,4,5,6,7$ and 8 vibrational levels of the C manifold are near-degenerate with the $v=0,1,2,3,4$ and 5 levels of the C' manifold. The two states dissociate into different separated-atom limits, a process complicated by the avoided crossing. In the absence of interaction, the separated-atom limit of the C state is two carbon atoms in ^1D states, while that of the C' state is two carbon atoms in ^3P states. The avoided crossing effectively reverses this scenario, as demonstrated by Messerle and Krauss in a second paper through rotational dissociation of C_2[75].

A *Rydberg state* of an atom or molecule is a state in which an electron has been promoted to a sufficiently high principal quantum number n that its interaction with the ionic core resembles the interaction between the electron and the nucleus in a hydrogen atom. Thus, binding energies of Rydberg states exhibit a $1/n^2$ dependence, and transitions between such states can be described using the Rydberg formula. Given that in the Bohr model the radius of an atom is proportional to n^2, highly excited Rydberg states can be thousands of times larger than the same atom in its ground state. In experiments using flash discharges through methane, Herzberg and two Swedish colleagues, Lagerqvist and Malmberg, discovered three new absorption transitions of C_2 lying in the vacuum ultraviolet region[76]. This is the name given to the range of wavelengths shorter than about 200 nm, where the photons are strongly absorbed by molecular oxygen (and at wavelengths below 150 nm, by nitrogen) and spectroscopy must be carried out using instrumentation that has been evacuated. (Methane itself is fairly transparent to these wavelengths.) The most energetic of these transitions was assigned as $F\,^1\prod_u - X\,^1\sum_g^+$, originating in the ground state, with a 0–0 band head at 134.1 nm (74,550 cm^{-1}, or 9.25 eV). The other two transitions, with 0–0 band heads at 142.4 nm (70,208 cm^{-1}) and 139.5 nm (71,674 cm^{-1}), were assigned as $f\,^3\sum_g^- - a\,^3\prod_u$ and $g\,^3\Delta_g - a\,^3\prod_u$, and originate in the lowest triplet state; thus, their energies are about 8.79 eV and 8.97 eV above the ground state. Herzberg *et al*. identified the upper states of these transitions as Rydberg states (in this case implying the occupation of molecular orbitals derived from atomic 3s orbitals, rather than 2p), based on their energies, and proposed that they originate in the electron configurations

$$(\sigma_g 2s)^2\,(\sigma_u^* 2s)^2\,(\pi_u 2p)^3\,(\sigma_g 3s)^1$$

for the F state, and

$$(\sigma_g 2s)^2\,(\sigma_u^* 2s)^2\,(\pi_u 2p)^2\,(\sigma_g 2p)^1(\sigma_g 3s)^1$$

for the f and g states. (These states are included in grey in Figure 4, but the transitions themselves are not shown.) If the $\sigma_g 3s$ molecular orbital truly

behaves like a hydrogenic 3s orbital, the *F* state configuration (for example) may reasonably be written as

$$(\sigma_g 2s)^2\ (\sigma_u^* 2s)^2\ (\pi_u 2p)^3\ (3s)^1$$

suggesting that the state may be described as originating in a $\pi_u \rightarrow 3s$ transition. The idea that these states might be Rydberg states was partially confirmed through MRCI calculations by Bruna and Grein in 2001[77]; they found that the singlet F state is a true Rydberg state, while the triplet *f* state has mixed Rydberg and valence character, and the triplet *g* state is predominantly a valence state. Speculation by Herzberg, Lagerqvist, and Malmberg that the singlet *E* state (the upper state of the Freymark bands) might also be a Rydberg state had in the interim been shown to be incorrect through the calculations of Barsuhn[78] and, later, those of Bruna and Wright[79]. According to these calculations, the *E* state is dominated by the configuration

$$(\sigma_g 2s)^2\ (\pi_u 2p)^4\ (\sigma_g 2p)^2$$

a doubly-excited state in which both electrons from the $\sigma_u^* 2s$ molecular orbital have been promoted to the $\sigma_g 2p$ orbital.

While technically difficult to observe on Earth, the far-ultraviolet *F*–*X* band (along with the *D*-*X* Mulliken band, also lying in the ultraviolet, and the infrared *A*-*X* Phillips band) has proved to be a useful signature of C_2 in space. In particular, data obtained using the GHRS and STIS instruments on the Hubble Space Telescope (one active 1990–1997, the other 1997–present) have been useful in identifying and quantifying C_2 present in diffuse molecular clouds lying along sightlines to X Persei and other cloud-occluded stellar light sources[80].

A purported new transition of C_2 near 22,341 cm^{-1} (2.77 eV) identified by van de Burgt and Heaven[81] in 1988 by laser induced fluorescence (LIF) of the photoproducts after photodissociation of toluene in a free-jet expansion and identified as being between the $c\ ^3\sum_u^+$ state and an unknown higher-energy state of Σ symmetry was reassigned on the basis of Bruna and Wright's calculations to a transition of the dicarbon cation C_2^+.

Again in 1988, another new state of C_2 was proposed, this time on the basis of resonance-enhanced multiphoton ionisation (REMPI) spectroscopy on a molecular beam of C_2 formed by two-step C–H bond-breaking photolysis from acetylene[82]. In this type of experiment, the carrier of the spectral features is unambiguously the neutral C_2 molecule, detected upon multiphoton ionisation *via* the mass 24 cation by time-of-flight mass spectrometry. The investigators, Peter Goodwin and Terrill Cool, reported an intermediate state in the ionisation of C_2 from its $A^1\Pi_u$ state to C_2^+ in its $^2\Pi_u$ ground state at an energy about 7.16 eV (57,719 cm^{-1}) above the C_2 ground state; analysis of the rotational fine structure led to the assignment of $^1\Delta_u$ symmetry to this state. The authors designate this

state $1^1\Delta_u$; it lies energetically slightly above the $E\,^1\Sigma_g^+$ state of the Freymark band. (Because of its proximity to the $E\,^1\Sigma_g^+$ state, it has been labelled $E''^1\Delta_u$ here. The existence and position of this state has been confirmed by Deheng Shi and co-workers[83] using CASSCF–MRCI calculations.)

Two further sets of transitions in the infrared were discovered by Douay, Nietmann and Bernath[60] in that same year through reanalysis of previously recorded spectra carried out on the basis of a suggestion made by Douglas McLean of the IBM Research Laboratory, a coauthor on previous work on the first definitive spectroscopic analysis of the congeneric (albeit heteronuclear) molecule SiC. McLean noted that the analogue of the $d^1\Sigma^+ - b^1\Pi$ transition seen in the spectrum of SiC (which has a triplet ground state) had never been observed in the spectrum of C_2. Though the relevant $B'\,^1\Sigma_g^+ - A^1\Pi_u$ and $B^1\Delta_g^+ - A^1\Pi_u$ transitions are actually quite strong, the region of the infrared where the bands occur, ranging from about 2000 cm^{-1} to about 8500 cm^{-1}, also contains bands from the Ballik–Ramsay and Phillips systems of C_2 as well as impurity bands from species such as CN and ArH leading to a very complex emission spectrum. Analysis of this highly congested spectrum required a "bootstrapping" approach in which band positions were at first guessed and a few initial line assignments made; a fit to these lines yielded spectral parameters that could then be used to pick out and assign other, weaker lines belonging to the same band. Eventually the entire band would be characterised and an improved set of parameters obtained.

A further set of transitions, commonly known as the LeBlanc bands, occurs between the $B'\,^1\Sigma_g^+$ and $D'\,^1\Sigma_u^+$ states in the singlet manifold. The standard reference for this set of bands is to the short report by Bao, Urdahl and Jackson[84] of their laser-induced fluorescence study on two-photon dissociation of acetylene in a supersonic molecular jet. The initial photolysis steps are analogous to the method of Goodwin and Cool, using two sequential ultraviolet photons of wavelength 193 nm to photolyse the C–H bonds of acetylene, initially producing C_2H, then C_2. The process generates C_2 molecules distributed over several excited states, including *B′*. Excitation from this state into the *D* state is then accomplished using an excimer-pumped tunable dye laser operating in the near-ultraviolet at around 356 nm. The observed rotational fine structure was then assigned using the molecular constants of Douay *et al.*[60] for the *B′* state and the tabulated data of Huber and Herzberg for the *D* state. This band system, along with the Deslandres–D'Azambuja band system, was later studied in solid neon matrices at 6 K by Wakabayashi, Ong, and Krätschmer[85].

2.5 Recent experimental developments

Within the past decade there has been a flurry of new developments in C_2 spectroscopy. In 2006, Kokkin *et al.*[86] reported the observation of a new band system, identified as $d\,^3\Pi_g - c\,^3\Sigma_u^+$, *via* an indirect method involving laser excitation of C_2 molecules formed in a discharge in the $c^3\Sigma_u^+$ state and

expanded into a vacuum chamber in the form of a supersonic beam, followed by monitoring of the Swan band fluorescence intensity from the $d^3\Pi_g$ state. The chosen excitation, carried out using an excimer-pumped tunable dye laser scanned over a range of wavelengths in the vicinity of 630 nm, was from the $v''=0$ level of the $c^3\sum_u^+$ state to the $v'=3$ level of the $d^3\Pi_g$ state. (The band origin of the transition is in the vicinity of 930 nm, or 10,752 cm^{-1}.) The fluorescence intensity was monitored at a fixed wavelength of 470 nm (in the $\Delta v=+1$ sequence of the Swan bands) using a monochromator. An initial scan over a broader range of laser wavelengths from 615 to 690 nm showed fluorescence from several distinct vibronic transitions of the new band system, including 7–3, 3–0, 4–1, 5–2, and 6–3, as well as multiple Swan band features from the $\Delta v=-3$ sequence. Given the doubt cast by the calculations of Bruna and Wright[79] on the carrier of the transition observed by van den Burgt and Heaven[81], this is the first robust report of any transition of C_2 involving the metastable $c^3\sum_u^+$ state. While the group that discovered these bands have proposed that they be referred to as the "Duck" bands[87] (by analogy with the Swan bands), they are referred to by Bornhauser *et al.*[62] as the Schmidt–Kable bands, and they are denoted here as Schmidt–Kable (*d*–*c*). More detailed investigations of the same band system were published by the same group in the following year[63]; this work took the form of a "two-dimensional" fluorescence experiment in which both excitation and detection wavelengths were varied, and yielded complete spectroscopic constants for the $c^3\sum_u^+$ state.

Members of the same experimental group, in collaboration with theoreticians from the same institution (the University of Sydney), later observed yet another new band system using a similar approach[88]: guided by high-level CASSCF-MRCI calculations (including relativistic corrections using the Douglas–Kroll–Hess approximation) on the $e^3\Pi_g$ state (the upper state of the Fox–Herzberg system, and the lowest state in the triplet manifold lying above the dissociation energy of the C–C bond[89]), they probed the $(v',v'')=(4,3)$ transition of the hitherto-unseen $e\,^3\prod_g - c\,^3\sum_u^+$ band system by laser induced fluorescence spectroscopy, detecting the fluorescence signal through the $e^3\Pi_g - a^3\Pi_u$ (Fox–Herzberg) transition. In the present work, these are designated as the Kable–Schmidt (*e*–*c*) bands.

Parallel to the expansion of the known spectroscopic systems of C_2 through ingenious experimental approaches, new techniques have also been used to explore some of the best-known spectroscopic bands. The degenerate four-wave mixing (DFWM) method has allowed the measurement of the (1,0) and (0,0) Swan bands to very high resolution and with line positions accurate to about 0.003 cm^{-1} using an ordinary oxy-acetylene welding flame as the spectral source[90–92].

In this approach two counterpropagating "pump" beams are crossed at a small angle (~1°) by a "probe" beam (see Figure 11). The "signal" is a beam that counterpropagates along the probe beam direction and is detected by

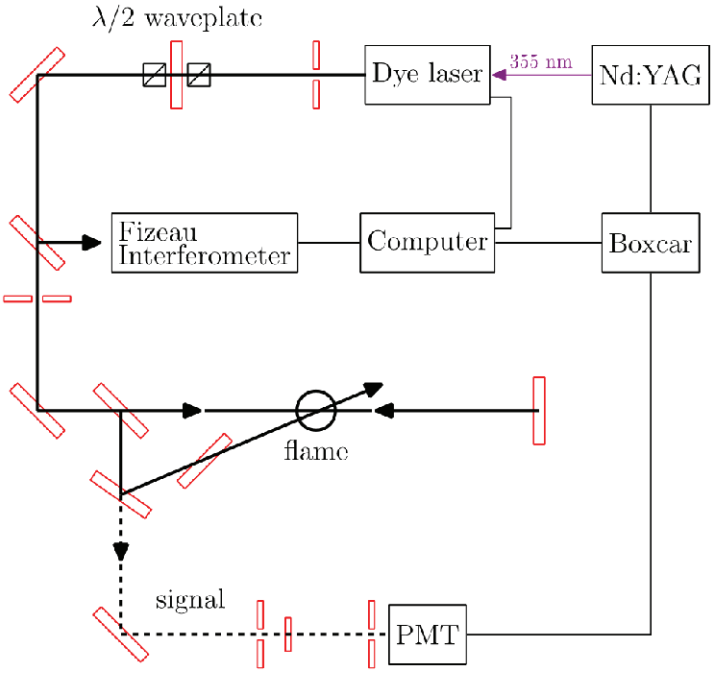


Figure 11 *Degenerate four-wave mixing approach used by Lloyd and Ewart[92].*

a photomultiplier tube. This is known as the backward or "phase-conjugate" geometry, and is the most common geometry used in DFWM. All four beams have the same wavelength (which is the origin of the word "degenerate" in the name of the method), swept across a chosen range using a pulsed dye laser (pumped, in the case of the work of Lloyd and Ewart, by a frequency-tripled Nd:YAG laser). The laser wavelength is carefully monitored using a Fizeau interferometer, a device in which two slightly nonparallel plates produce a series of interference fringes with a spatial relationship dependent on the wavelength. Scattered pump and probe light can be rejected using polarisers. Discrimination of signal from noise in the DFWM experiment is accomplished through the highly directional character of the signal beam, as well as its laser-like phase coherence which allows it to be "spatially filtered" (by passage through a series of apertures) to remove undesirable components. (This makes the method very useful in situations of strong background luminosity, such as flames.) The DFWM phenomenon can be described in terms of the generation of an effective diffraction grating through the interference of two input laser beams; the third beam then scatters from this grating to produce the "phase-conjugate wave". The grating physically consists of a spatially periodic density distribution of molecules in excited states. The DFWM signal is thus enhanced by resonance at the spectroscopic transitions of the medium. The use of the counter-propagating beam geometry allows almost complete elimination of

thermal Doppler broadening from the spectrum obtained by DFWM, leading to improved signal-to-noise ratios over laser-induced fluorescence (LIF) spectra obtained under identical conditions[91].

The most recent work on C_2 spectroscopy has focused on exploration of higher vibrational levels of several known states by supersonic jet/LIF methods[93], as well as exploration of the quintet manifold of C_2 for the first time[54,55,62]. One advantage of the molecular beam generation method (pulsed corona discharge followed by slit-jet supersonic expansion) is that while the molecules are excited to high vibrational states (leading to a vibrational temperature of several thousand kelvin), the rotational temperatures are low (less than 100 K), reducing the number of rotational transitions present in each band and thus simplifying the analysis. The energetically lowest quintet state, labelled $1^5\Pi_g$, lies about 3.7 eV above the $X\,^1\sum_g^+$ ground state. Spectroscopic evidence of this state was first seen by the group of Peter Radi at PSI in 2011 using the two-colour resonant four-wave mixing (TC-RFWM) method[54]. (TC-RFWM differs from DFWM in that the frequency of the probe laser is different from that of the pump lasers used to create the transient "grating". Signal is only extracted when *both* frequencies are in resonance with molecular transitions; moreover, these transitions must share one energy state in common.) The Bornhauser–Radi experiment was a "deperturbation" study of a type previously carried out by the same group in order to obtain improved spectroscopic parameters for the $d^3\Pi_g$ state, the upper state of the Swan band system[55]. (The $v'=4$ vibronic level of the $d^3\Pi_g$ state is strongly perturbed by the $v'=16$ vibronic level of the $b\,^3\sum_g^-$ state, leading to uncertainty in the analysis and some incorrect assignments in the literature. TC-RFWM using a pump beam fixed on a particular rotational line in the (2,3) Swan band at an energy of 18,086.3 cm^{-1} and a probe beam scanning the spectral region where a line in the (4,3) Swan band is perturbed by a line in the (16,3) Ballik–Ramsay band allowed unambiguous assignment of the transition using a suitable model Hamiltonian. The accurate prediction of the unperturbed spectra was made possible by the existence of earlier extremely high-quality Swan band spectral parameters determined by Tanabashi and co-workers using Fourier transform spectroscopy[53].) The TC-RFWM method allows access to "dark states" from which transitions are not usually observed *via* the probing of "window states", rovibrational levels belonging to known spectral bands that have a quantum-mechanically mixed character due to perturbation by levels belonging to the dark state. In this case, deperturbation analysis of the (6,5) and (6,4) bands of the Swan system revealed mixing between rovibronic levels of the $d^3\Pi_g$ ($v=6$) state and levels from both the $b\,^3\sum_g^-$ ($v=19$) state and the hitherto unseen $1^5\Pi_g$ state and allowed several spectral parameters of the previously hidden state to be determined, including the position of its vibronic origin (determined with respect to the $a^3\Pi_u$ ($v=0$) level) T_0 at 29,258.5922 cm^{-1}. Spectroscopic constants for the $1^5\Pi_g$ state were obtained computationally by Schmidt and Bacskay[94]

using the CASSCF – MRCI – DKH approach, and were found to be in excellent agreement with experimental values, moreover strongly suggesting that the state observed by the Radi group was indeed the $v=0$ substate of $1^5\Pi_g$. The most recent measurements of the Radi group[62], already discussed above, take these studies one step farther by utilising the $d^3\Pi_g$ (v = 6) state as a window state to populate the quintet manifold leading to the observation of 57 rovibronic transitions belonging to five sub-bands of the $1^5\Pi_u - 1^5\Pi_g$ system. The electron configurations of the upper and lower quintet states involved in this band system are

$$(\sigma_g 2s)^2\,(\sigma_u^* 2s)^1\,(\pi_u 2p)^2\,(\sigma_g 2p)^2\,(\pi_g^* 2p)^1 \qquad 1\,^5\Pi_u$$

and

$$(\sigma_g 2s)^2\,(\sigma_u^* 2s)^2\,(\pi_u 2p)^2\,(\sigma_g 2p)^1\,(\pi_g^* 2p)^1 \qquad 1\,^5\Pi_g$$

making the Radi – Bornhauser system analogous to the Swan system in the sense that both formally involve the promotion of an electron from a $\sigma_u^* 2s$ molecular orbital to a $\sigma_g 2p$ molecular orbital; correspondingly, the excitation energies in both cases are near 20,000 cm^{-1}, and both transitions involve an increase in rotational constant B of around 7%, indicating a bond that is shorter in the upper state than in the lower state.

3. Theories of bonding in C_2

3.1 Early work – Mulliken and molecular orbital theory

Almost since molecules began to be understood through the lens of quantum theory, calculations of electronic structure have played a role in the understanding of the bound states of C_2. Many of the concepts and much of the terminology related to both the spectroscopic analysis and the computational theory of diatomic molecules are due to Friedrich Hund and Robert Sanderson Mulliken, the latter of whom received the Nobel Prize in Chemistry in 1966. Mulliken's first contribution to the understanding of C_2 came in 1927[95], in a paper applying quantum theory (with transition intensities and selection rules discussed using the Bohr correspondence principle, the ink being barely dry on Schrödinger's theory) to the intensity distribution of the Swan bands, building on work by Hund on the "second positive band" (observed at the positive electrode in a discharge) of nitrogen. Mulliken identifies the bands as having $^3P - ^3P$ character (the modern $^3\Pi - ^3\Pi$ notation would not be introduced by Hund until the following year[96]); the structure of the bands, with weak or absent Q (that is, $\Delta J = 0$) branches, is well explained under the assumption, based on the small triplet splittings, that the molecule belongs to the "Hund's (b) case" (spin vector **S** weakly coupled to the internuclear axis, orbital angular momentum $\Lambda \neq 0$) even for small values of rotational quantum number J. Other spectral patterns are noted, including Λ-doubling and a phenomenon of intensity

alternation attributed by Mulliken to contributions to the overall symmetry from the spins of the nuclei. (Later, Francis A. Jenkins of UC Berkeley would use the intensity distribution of the Swan bands to determine the spins of the ^{13}C and ^{14}C nuclei[97]. The alternately missing components of the Λ-doublets in the Swan bands of $^{12}C_2$ are a consequence of the fact that ^{12}C has nuclear spin zero. The same pattern is apparent in $^{14}C_2$, indicating that its spin is also zero, while $^{13}C_2$ shows Λ-doublets with a 3 : 1 intensity ratio, indicating a nuclear spin of ½.)

Mulliken's early contribution to the theory of diatomic molecules and their spectra came in the form of a series of monumental articles in the *Review of modern physics*. These articles serve as a bridge between the old (Bjerrum–Bohr) and new (Schrödinger) quantum theories, beginning in parts I and II[98] (a single long paper) with an overview of the current state of knowledge with regard to the theoretical understanding of the band spectra of diatomic molecules, describing the various forms exhibited by the fine structure of spectral bands, introducing the quantum numbers used in their analysis, and working through the steps by which the details of the spectrum can be used to reconstruct the form of the molecule's internuclear potential. In part III[99] of *The interpretation of band spectra* (as the series was entitled overall), Mulliken introduces the terminology of molecular orbital (MO) theory and applies it to the energy levels and spectroscopic transitions of diatomic molecules. It is in this paper that he discusses the electronic structure of C_2, making the (guarded) assumption that the $^3\Pi_u$ state is the ground state and predicting the positions of 10 other singlet and triplet electronic states in a range of 7.6 eV above it. The σ and π MO occupancies associated with these states are also given. (The terminology σ and π used in relation to bonds is due to Hund[100], and originates around this time.) Making reference to the valence bond theory of Heitler and London, Mulliken comments that VB theory would predict the $^1\sum_g^+$ state to be "far lower in energy" than $^3\Pi_u$ because the former (considered as formed from two carbon atoms in ^5S states, each with four unpaired electrons) contains four "spin valence bonds" and the latter (considered as formed from two ^3P atoms) only one; MO theory, on the other hand, assigns a valence bond (VB) order of *two* to both states based on the numbers of bonding and antibonding electrons. In the early days of VB and MO theories as competing quantum-mechanical explanations of molecular electronic structure, this would seem to be a win for MO theory – at least, if $^3\Pi_u$ really is the ground state or the two states are very close in energy.

Several years later Mulliken returned to the question of the nature of the bonding in C_2[101], using MO theory to predict energies and bond lengths for several states of C_2 and C_2^+, and obtaining results in reasonable agreement with experiment, including a predicted bond length of 1.24 Å for the $^1\sum_g^+$ state of C_2 and 1.32 Å for the $^3\Pi_u$ state. In the paper he compares these bond lengths with "typical" single, double, and triple carbon–carbon bond lengths found in polyatomic molecules (which he gives as 1.62 Å, 1.40, Å, and 1.18 Å), as

well as the length of the quadruple bond associated with an excited $^1\sum_g^+$ state of C_2 lying about 6 eV above the ground state and formed by promotion of the electron pair in the σ_u^*2s MO into the $\sigma_g 2p$ bonding orbital (about 1.12 Å); this state, which is the upper state of the Freymark band, is unusual in that it has a formal bond order of four but lies energetically above the dissociation energy of C_2 into two carbon atoms in their ground states. (It should be noted that the energies given in Mulliken's papers are obtained using a semi-empirical *ansatz*, and are quite approximate.) Mulliken's calculated bond lengths for polyatomic molecules are – as he notes – longer than the observed values, a fact attributable to the lack of detailed consideration of bonds between carbon and other atoms (such as hydrogen) in the calculations.

3.2 Computational approaches to bonding in C_2

The appearance of powerful (by the standards of the time) electronic computers such as the IBM 650 and the UNIVAC I during the 1950s revolutionised theoretical chemistry, but divided theoretical chemists into two camps: those who continued with the semiempirical methods that gave approximate results but yielded clear physical insight (accepting the vast speed improvement in these calculations afforded by the computer over the hand-operated mechanical calculators that had preceded it), and those who believed the future lay with much larger-scale computations that could only be accomplished with the help of the machines. Though Mulliken of himself was of the "old school" (he would eventually learn to write computer programs at the age of 74), his laboratory at the University of Chicago, the Laboratory of Molecular Structure and Spectra, was the leading incubator for the new type of calculation. The introduction of self-consistent-field theory by Douglas Hartree, Vladimir Fock and Clemens C. J. Roothaan (known in the LMSS as C[2]), and others (such as Francis Boys in the UK) had made the implementation of molecular orbital theory a tractable problem for computers to handle.

Even so, as late as around 1960, theoretical work on C_2 was largely semi-empirical. The first[102] of a series[102–104] of papers by Enrico Clementi and Kenneth S. Pitzer uses the "magic formula" approach of Mulliken[105] (an interesting blend of MO and VB theories that uses the VB "criterion of maximum overlapping" championed by Slater and Pauling to relate experimental molecular dissociation energies to theoretical MO overlap integrals) to obtain σ and π molecular orbital energies as a function of internuclear distance and from these the approximate energies of the first few electronic states of C_2; at this point it is considered to be experimentally "quite well established" (though in fact false) that the $^3\Pi_u$ state is lowest in energy, and the theoretical work is not sufficiently exact to distinguish the $^1\sum_g^+$ and $^3\Pi_u$ states in energy order. The second paper in the series[103] uses the ASMO–LCAO (antisymmetric molecular orbitals by linear combination of atomic orbitals – antisymmetrisation using the Slater determinant method guarantees that the resulting wavefunction obeys the Pauli principle) approach,

followed by construction of overall wavefunctions as linear combinations of a limited number of Slater determinants, to study the lowest-energy excited states of C_2 a little more closely. The solution of the secular equations was carried out on an IBM 701 computer, and the results showed the $^3\Pi_u$ state to be lower than $^1\sum_g^+$ by less than 0.25 eV, in good agreement with the prevailing experimental view of the time. The paper therefore represents a transitional work, where it is clear that more accurate results require a "more refined calculation" incorporating a better treatment of the electronic configurations contributing to the wavefunctions, but where brute force computation is still very much a servant to human judgment and ingenuity. Some earlier (and less successful) computational work on C_2 is also alluded to in this paper, including the doctoral work of (Sister) M. Ignatia Frye. The third paper[104] in this series, by Clementi alone, uses theoretical data to predict transition probabilities between these low-lying states, obtaining values for several known band sequences (Swan, Deslandres–D'Azambuja, Mulliken, Fox–Herzberg, Phillips, and Ballik–Ramsay) as well as three unobserved transitions (now identifiable as the Bernath B and B′ band sequences, and the LeBlanc transition).

Very shortly after the publication of these papers, Shirley Read and Joseph Vanderslice carried out a Rydberg–Klein–Rees analysis of the spectroscopic data set to reconstruct potential curves for the electronic states involved[106], and compared their results to the computational work of Clementi and Pitzer; this appears to be the first theoretical paper acknowledging the newly-discovered singlet character of the C_2 ground state. The agreement between the experimentally-derived potential energy curves and the semiempirical computed ones is judged to be reasonably good, and much better than the previous nonempirical calculations cited by Clementi and Pitzer.

The first comprehensive high-quality *ab initio* (that is, fully nonempirical) calculations on C_2 were published by Fougere (a PhD student at Boston University) and Nesbet (of IBM Research Laboratories, San Jose) in 1966[107]; these used three different basis sets of Slater-type (that is, nodeless and exponential) basis orbitals and two different levels of configuration interaction (CI), and led to theoretical potential curves and spectroscopic constants for a total of 27 states. The minimal basis set used in preliminary calculations, designated MB, allowed one Slater function for each atomic orbital occupied in the ground state of carbon, generating symmetric and antisymmetric linear combinations of these to generate the molecular basis: "production" calculations employed a basis set designated XB (extended basis) doubling the number of exponents used to describe orbitals in the valence space, but retaining the minimal core, while the largest basis set, designated 6XB+D, used two basis functions for each core and valence orbital, and supplemented the valence space description with σ and π functions derived from atomic d-orbitals – these would today be known as "polarisation functions", orbitals of higher angular momentum than those occupied in the atom's ground state, included to confer some additional

flexibility on the basis set. The two levels of CI employed (mysteriously named "non-PAN" and "semi-PAN") included, respectively, states generated only from MOs originating in the 2p orbitals and those originating both in 2s and 2p. The XB/non-PAN combination incorrectly leads to a $^3\Pi_u$ ground state, but adding a single further important excited-state determinant (generated by promotion of both σ_u^*2s electrons to $\sigma_g 2p$) with an unchanged basis set lowers the energy of $^1\sum_g^+$ such that it becomes the ground state. Noting the importance of configurations involving the promotion of 2s electrons, Fougere and Nesbet carried out two further sets of calculations at the XB/semi-PAN and 6XB + D/semi-PAN levels; the last of these yields good agreement with the experimentally-derived potential curves of Read and Vanderslice for most of the electronic states computed. Computing spectroscopic constants for the various experimentally-known electronic states and carrying out a regression analysis comparing these to spectroscopic values, Fougere and Nesbet noted that the deviations between theory and experiment were approximately linear (for example, theoretical values of T_e, the energy of the curve minimum above the ground state minimum, are always underestimates), and suggested that interpolated values could be used to predict the properties of the as-yet undetected states.

The next significant attempt at extending the set of electronic states of C_2 for which high-level calculations had been carried out was work carried out by Kate Kirby, of the Harvard – Smithsonian Center for Astrophysics (and currently the Executive Officer of the American Physical Society), in collaboration with Bowen Liu of IBM Research Laboratories[108]. These calculations used a basis set of similar quality to the best basis used by Fougere and Nesbet, consisting of two Slater functions per atomic orbital up to 3d, and employed a multiconfiguration approach with limited (valence-only) configuration interaction; Kirby and Liu computed the energies of 62 valence states of C_2 at 31 internuclear separations ranging from 1.4 to 20.0 Bohr radii (0.74 to 10.6 Å). This is the full set of states that dissociate into valence states of the atom (that is, states not involving occupation of any orbital above 2p). Many of the states are purely repulsive in character, exhibiting no minimum in the potential energy curve. Despite its comprehensiveness, in some sense this is a poorer set of calculations than those of Fougere and Nesbet – for example, the $^3\Pi_u$ state is computed to be lower in energy than the $^1\sum_g^+$ state by 0.03 eV – and any admixture of Rydberg character that might exhibited by the more energetic valence states is systematically neglected. Nonetheless, this remains a landmark study, identifying 19 previously unknown bound states and multiple repulsive states of possible significance in the molecular dissociation process.

Some subsequent computational studies on C_2 have already been described above in the context of the experimental work, for example the MRD-CI calculations of Bruna and Wright[79] on some of the doubly-excited states, and the ongoing large basis set MRCI work that accompanies the LIF experiments

of the University of Sydney group[88,94]; the methodology used by the latter group (large basis set multireference CI theory based on full valence space CASSCF wavefunctions, with energies further improved using Davidson's extrapolation and relativistic corrections) essentially represent the current state of the art, capable of producing energies accurate to within a few kJ mol^{-1} for most valence states. Apart from the observation and analysis of two new transitions[86,88] and the theoretical study of states in the newly-unearthed quintet manifold[94], recent work from this group has been focussed on the accurate calculation of oscillator strengths (which predict transition intensities) for several known band systems of C_2 [49,109].

Davis Sherrill and his collaborators have used C_2 as a test species for evaluation of high-level computational methods in terms of their ability to describe bond breaking processes[110,111]. With "full configuration interaction" (FCI) – which includes all possible Slater determinants of the appropriate symmetry in the solution space of the Schrödinger equation – representing a benchmark "exact solution", the benchmark potential energy curves for the $X\,^1\Sigma_g^+$, $B^1\Delta g$, and $B'\,^1\Sigma_g^+$ states of C_2 were compared with curves calculated using the high-level approximate theories known as MP2 (second-order Møller–Plesset perturbation theory), CISD (configuration interaction with single and double substitutions), CCSD (coupled-cluster theory with singles and doubles) and CCSD(T) (coupled-cluster theory with singles, doubles, and perturbative triples) [110]. Curiously, the basis set used for these comparisons was the modest and completely standard polarised double-zeta basis known as 6-31G*, which was created by the Pople group in the 1970s[112,113]. Calculations using both RHF (restricted Hartree–Fock, in which the α-spin and β-spin spatial orbitals are constrained to be the same) and UHF (in which they are not) reference functions were considered. The UHF-based methods outperformed the RHF-based methods in the sense that they gave qualitatively correct potential energy curves (while many of the RHF-based curves were not even qualitatively correct), but all except the very best methods studied showed "nonparallelity" errors in excess of 100 kJ mol^{-1}. (In the ideal case, the approximate wavefunction would give a potential energy curve simply shifted from the exact curve by an amount independent of internuclear distance; the nonparallelity error measures the difference between the minimum and maximum deviations of the approximate curve from the exact one.) In a second paper some superior approximate methods were compared with the same benchmark energy curve[111]. The most sophisticated of these was a multireference second-order CI approach based on a CASSCF reference wavefunction; this method generated MRCI wavefunctions containing over a quarter of a million determinants, and represents a "complete limit" for the multireference CISD approach. Another of these methods was a version of the multiconfiguration second-order perturbation theory method CASPT2, while yet another set of approaches considered was a group of methods based on "completely renormalised" coupled-cluster theory. The latter group, including

methods such as CR–CCSD and CD–EOMCCSD (equation-of-motion coupled cluster theory, used for excited states), employed a technique known as method of moments coupled cluster theory (MMCC), devised by Piotr Piecuch and his co-workers[114]; MMCC supplies explicit mathematical expressions for the difference between the CC or EOMCC energy of a given state and its full CI value. These methods outperformed the ones considered in the prior study on C_2 bond breaking and gave results of comparable quality to MRCI. Thus, the CR methods have the advantage of yielding MRCI-level results using a "black box" technique that can be employed by non-specialist users.

Other notable recent high-level studies include the work of De-Heng Shi and co-workers[115,116] on multiple states of several C_2 isotopologues, employing the CASSCF/MRCI approach with large basis sets (up to aug-cc-pV6Z) corrected using the Davidson extrapolation, with relativistic effects taken into account using the second-order Douglas–Kroll–Hess approximation; these methods produce accurate spectroscopic constants for rotationally unexcited ($J=0$) molecules. The C_2 molecule has also recently played its role as a proving-ground for new computational methodology in a study using the full configuration interaction quantum Monte Carlo (FCIQMC) approach[117]. This method employs an approach involving the stochastic dynamics of a finite number of integer "walkers" to determine the amplitudes of individual determinants in the full CI space; the walker populations converge over time to approach the coefficients of these determinants in the FCI expansion. The calculation (using 6.3 million walkers) is intended primarily as a test of the method (on the $^1\sum_g^+$ ground state), and the paper does not present computed spectroscopic constants.

3.3 Recent developments in theory – GVB theory enjoys a revival

In the early days of molecular quantum mechanics, valence bond (VB) and molecular orbital (MO) theories provided alternative descriptions of molecular electronic structure and chemical bonding. The VB method has its origins in the 1927 calculations of Walter Heitler and Fritz London on H_2[118] in which the two-electron wavefunction is constructed from a basis of hydrogenic 1s functions where one electron is initially localised on each nucleus; their "ansätze" (guesses – which also turn out to be the solutions to the problem) take the form of symmetric (+) and antisymmetric (–) linear combinations (presented here without normalisation factor)

$$\psi_{\pm}(1,2)=\phi_{1sA}(1)\,\phi_{1sB}(2)\pm\phi_{1sA}(2)\,\phi_{1sB}(1)$$

such that the symmetric combination exhibits a curve of energy against internuclear distance with a minimum and the antisymmetric does not; to generate a wavefunction that is antisymmetric overall this combination of spatial orbitals is then multiplied by an appropriate linear combination of spin functions

$$\psi_{\pm}(1,2) = [\phi_{1sA}(1)\,\phi_{1sB}(2) \pm \phi_{1sA}(2)\,\phi_{1sB}(1)](\alpha(1)\beta(2) \mp \alpha(2)\beta(1))$$

Thus the notion that that pairing of electrons of opposite spin leads to a bond appears as a central concept in VB theory, and forms a quantum-mechanical version of the electron-pair bond previously envisaged by Gilbert Lewis. However, it would be erroneous to assume that bonding (that is, lowering of the total energy with respect to two noninteracting H atoms in their 1s ground states) is *per se* a spin-related phenomenon. The overall energy as a function of internuclear distance *r* is given by the expression

$$E_{\pm}(r) = 2E_H + \frac{J(r) \pm K(r)}{1 \pm T(r)}$$

where $J(r)$, $K(r)$, and $T(r)$ are known as the Coulomb, exchange, and overlap integrals respectively. The Coulomb and exchange integrals as defined by Heitler and London depend on the parts of the potential energy not already included in the $2E_H$ term; that is, the attractive potential between each electron and the "other" nucleus, the repulsive potential between the electrons, and the repulsive potential between the nuclei. Bonding arises from the interplay between these terms in the *J* and *K* integrals and, in particular, from the *r*-dependence of the overlap term in the exchange integral *K*[119,120]. The singlet character of this two-electron state is simply enforced by the Pauli principle.

The prehistory of MO theory lies in the quantum numbers assigned by Hund and Mulliken to electrons in diatomic molecules as a component of their interpretation of vibronic spectra, and in the relationships thus made apparent between the energy levels in the molecular state and in the separated atom states. This qualitative approach was turned into a formal theory by John Lennard-Jones in 1929 in a paper entitled "The electronic structure of some diatomic molecules", in which he demonstrated the paramagnetic character of O_2 based on MO arguments[121]. In the MO approach, electrons were from the very beginning associated with orbitals extending over the entire molecule, constructed from linear combinations of atomic orbitals (LCAO). The wavefunction could be made antisymmetric with respect to interchange of electron labels as required by the Pauli principle by expressing it as a determinant, a technique introduced by Slater in the same year for use in the theory of many-electron atoms[122]. Thus, for H_2, again considering a basis of hydrogenic 1s atomic functions, a one-electron bonding MO for electron *j* takes the (normalised) form

$$\psi(j) = \frac{1}{\sqrt{2}}(\phi_{1sA}(j) + \phi_{1sB}(j))$$

and the two-electron wavefunction is then expressed as the Slater determinant

$$\psi(1,2)=\frac{1}{\sqrt{2}}\begin{vmatrix}\psi(1)\alpha(1) & \psi(1)\beta(1)\\ \psi(2)\alpha(2) & \psi(2)\beta(2)\end{vmatrix}$$

which can be rewritten in the form

$$\psi(1,2)=\frac{1}{\sqrt{2}}\psi(1)\psi(2)[\alpha(1)\beta(2)-\alpha(2)\beta(1)]$$

The spin part of this wavefunction is clearly the spin singlet function seen above for the Heitler–London case; the space part, however, is different, and can be expanded as

$$\psi(1)\psi(2)=\frac{1}{2}(\phi_{1sA}(1)+\phi_{1sB}(1))(\phi_{1sA}(2)+\phi_{1sB}(2))$$

$$=\frac{1}{2}[(\phi_{1sA}(1)\phi_{1sB}(2)+\phi_{1sA}(2)\phi_{1sB}(1))+(\phi_{1sA}(1)\phi_{1sA}(2)+\phi_{1sB}(1)\phi_{1sB}(2))]$$

where the first term is the Heitler–London VB solution and the second comprises a pair of "ionic" contributions where both electrons are localised on the same nucleus. Thus, MO and VB theory represent different starting points from which more realistic solutions can be approached: more extensive basis sets in both cases, the inclusion of ionic structures in the VB case, and the inclusion of configuration interaction in the MO case, all lead to improved results.

Historically, the VB approach dominated over the MO approach in the early days of molecular quantum mechanics, mainly because of its conceptual relationship with the familiar idea of the Lewis electron pair bond, aided by the "evangelism" of Pauling for his "resonance" concept which chemists found intuitive, but gradually the simplicity of the MO interpretation and its relative ease of implementation in computer code led to a reversal in the fortunes of the two methods. For molecules more complex than the two-electron system H_2, MO theory is much simpler to apply. The starting point for a MO calculation is a single Slater determinant, with the spatial parts of the molecular orbitals represented as LCAO; the coefficients of the atomic orbitals in the molecular orbitals are optimised by minimisation of the expectation value of the total energy computed using these MOs. The details of the computation are greatly simplified by the fact that for a single determinant wavefunction the orbitals can be chosen to be orthonormal (making overlap integrals between different orbitals zero)[123]. This is the basis of Hartree–Fock–Roothaan SCF-MO theory as it is typically implemented. The advantages of orthogonality even extend into CI extensions of MO theory, as the configurations thus generated are themselves also orthogonal. The starting point for a valence bond calculation is a wavefunction of considerably more complicated form, fundamentally multideterminantal in character, that represents a product of $n/2$ electron-pair bond wavefunctions.

Both MO and VB wavefunctions can be improved by CI methods: in the former case this is a matter of including additional Slater determinants corresponding to excited state configurations in which one or more electrons has been promoted to a MO unoccupied in the Hartree–Fock reference configuration (hence *configuration interaction*, or CI), while in the latter case the configurations requiring to be added are the "ionic" configurations missing from the VB reference function. (In the VB reference function, corresponding to "perfect pairing" of electrons, all AOs are singly occupied; ionic configurations then correspond to double occupancy of some subset of AOs.) The transparency of interpretation that gives the VB reference configuration its advantage over MO theory is quite rapidly lost in such CI expansions. Additionally, VB theory originated in an era of minimal basis set descriptions (one orbital per AO occupied in the atomic ground state configuration), and improvement of VB theory treatments by enlargement of the basis set is rather complicated, leading to a large number of additional pairing configurations.

These complexities in the implementation of VB theory, together with a string of conspicuous successes of its rival in describing organic chemical reactivity (for example the explanation of electrocyclic reactions using the Woodward–Hoffman rules[124,125]) and spectroscopic observations (for example the photoelectron spectra of methane and the other group 14 hydrides, which exhibit two clear peaks well-described *via* "ionisation from a_1 and t_2 orbitals analogous to the *s* and *p* orbitals of the corresponding inert gases"[126], and show no experimental evidence for the sp^3 orbital hybridisation invented by Pauling to explain the tetrahedral character of carbon[127]) led to a long period of eclipse for the VB approach as a mainstream method of computational chemistry. Nonetheless, exponents of VB theory continued working to improve the utility of the method, leading to the development of a number of new and efficient implementations such as generalised valence bond (GVB) theory[128], which makes use of an elegant trick invented by Charles Coulson and Inga Fischer[129] that allows implicit incorporation of covalent and ionic configurations into a single, formally covalent, VB structure by means of the adoption of "slightly delocalised" orbitals in the place of localised atomic orbitals.

Formally, a Coulson–Fischer orbital for a bond AB can be written (without normalisation constants)[130]

$$\psi_{CF} = |\varphi_L \overline{\varphi_R}| - |\overline{\varphi_L} \varphi_R|$$

where $\varphi_L = \phi_A + \varepsilon\phi_B$ and $\varphi_R = \phi_A + \varepsilon'\phi_B$ (ϕ_A, ϕ_B atomic orbitals on centres A and B); this expression can then be expanded as

$$\psi_{CF} = (1 + \varepsilon\varepsilon')\left(|\phi_A \overline{\phi_B}| - |\overline{\phi_A} \phi_B|\right) + \varepsilon' |\phi_A \overline{\phi_A}| - \varepsilon |\phi_B \overline{\phi_B}|$$

which clearly has the form of a VB wavefunction. The GVB method uses Coulson–Fischer orbitals delocalised over the entire molecule (called "overlap-

enhanced orbitals" or OEOs), and in its most widely used form employs the approximations of "perfect pairing" (PP) and "strong orthogonality" (SO), leading to the GVB-SOPP method. The PP approximation results in a molecular wavefunction expressed as a product of "geminal" functions, each describing a pair of singlet-coupled electrons (which may correspond to a bond or a lone pair), while the SO constraint requires mutual orthogonality of the geminals. Further computational simplification is obtained if the geminals are rewritten in terms of natural orbitals (one-electron orbitals that diagonalise the one-particle density matrix). In this way, a GVB-PP calculation is in fact an MCSCF calculation restricted to a particular subset of configurations.

A complication shared by all GVB methods is the construction of appropriate spin eigenfunctions. (Indeed, this is a general issue in the construction of CI wavefunctions for non-singlet states that are eigenfunctions of the total spin operators $\widehat{S^2} = \sum \widehat{s_i^2}$ and $\widehat{S_z} = \sum \widehat{s_{z,i}}$.) Methods of construction utilise the mathematical theory of permutation group symmetry and are either *synthetic* (in which the complete set of eigenfunctions of a given S and M is built up from one-electron or two-electron eigenfunctions) or *analytic* (in which a spin eigenfunction is *projected* from a wavefunction with mixed spin character using an appropriate projection operator)[131]. A compendium of these methods is presented in the book by Pauncz[132].

Refinements of the GVB approach include the SCVB (spin-coupled VB) method, proposed almost simultaneously by Gerratt and Lipscomb[133] and by Ladner and Goddard[134]; like GVB–PP, this is a single-configuration approach, but one in which the SO and PP restrictions have been removed and all possible modes of spin-pairing are permitted. Both the form of the orbitals and the relative weights of different spin-pairing modes are optimised variationally. The SCVB approach permits a quantitative understanding of the relative importance of different modes of spin-coupling (in rough correspondence with canonical Lewis structures) through the analysis of the pairing coefficients. (Thus, for example, the structure of methane is dominated by a single GVB-PP configuration[135], while that of benzene exhibits larger contributions from two Kekulé structures and smaller contributions from three Dewar structures[136].) The SCVB method is one of the most accurate and flexible theoretical approaches based on a single configuration, and gives a good description of bond breaking and formation during chemical reactions. A variety of CI extensions of SCVB theory are also possible.

Alternative GVB approaches that make use of localised "hybrid" atomic orbitals (HAOs) include the valence bond self consistent field (VBSCF) method of van Lenthe and Balint-Kurti[137,138] and the breathing-orbital valence bond (BOVB) method of Hiberty and co-workers[139]; VBSCF is a multiconfiguration approach in which the wavefunction is a linear combination of VB functions, and both the orbitals of the individual VB structures and the weighting of these structures in the linear combination are optimised simultaneously, while the

BOVB approach, though fundamentally similar, adds some extra flexibility by optimising the VB structures in a way that maximises their mutual interaction energy. Usually, the orbitals of the system are split into a set containing "active electrons" (which change significantly during a process such as bond breaking) and a set containing "spectator electrons" (which remain doubly occupied throughout the process); the two sets of orbitals are allowed to optimise independently, and change shape continuously as the bond breaking process progresses, leading to the name of the method. The BOVB approach does a good job of capturing both static and dynamic electron correlation, and yields accurate values for activation energies and bond dissociation energies.

The interpretive power of GVB methods was recently used by Dunning and co-workers in analysing the contribution of recoupled pair bonding to hypervalence and the "first row anomaly" (the fact that hypervalence is not exhibited by elements in that row of the periodic table)[140]. The recoupled pair bond, touted by its discoverers as a new kind of chemical bond, is a phenomenon that occurs when an electron in a formally singly-occupied ligand orbital decouples formally paired electrons on a central atom and recouples them as a singlet to form a bond. For example, this type of process is envisaged as the mechanism for bond formation between beryllium, which in its ground state has a valence electron configuration $2s^2$, and monovalent ligands. Similarly, the concepts of recoupled pair bonding can be used to describe the bonding of trivalent or tetravalent carbon without resorting to the traditional sp^n Pauling-type hybrid orbitals. In fact, many years earlier, in 1973, Goddard and his co-workers – including Thom Dunning, then at the beginning of his research career – used the same bonding concepts, albeit not by name, in a GVB-based analysis of the bonding in hydrocarbons and fluorocarbons[141].

3.4 Application of valence bond theory to C_2 – the question of bond order

Recent work on C_2 using modern VB theory began with a 2011 paper (referred to below as Su *et al*.[142]) from a Chinese–Israeli–French collaboration between the group of Wei Wu at Xiamen University and the seasoned GVB exponents Sason Shaik and Philippe Hiberty. To motivate the work, the authors note the "enigmatic" nature of the bonding in C_2, where application of a principle of maximum coupling leads to a triple bond (with one σ and two π and two π bonds as in acetylene), with the odd electrons then coupling antiferromagnetically to give the molecule's singlet ground state, while – as discussed in the *Introduction* – straightforward application of MO theory (and the MO energy level diagram of Figure 1) leads to two "suspended" π bonds (in the sense of lacking the support of a σ bond) and "lone pairs" constructed from the fully occupied bonding/antibonding MO pair $\sigma_g 2s$ and $\sigma_u^* 2s$. Other intriguing features noted by the authors are the short ground state bond length of 124.24 pm, the existence of the low-lying triplet state $a^3\Pi_u$, and the fact that

the bond length in this state is substantially longer, at 131.19 pm, despite the closeness in energy of the two states. (It might be mentioned here that the *upper* state of the Swan band transition, $d^3\Pi_g$, has a bond length of 126.21 pm, and there are other excited states of C_2, specifically $c^3\sum_u^+$ and $D'\,^1\sum_u^+$, that have potential energy minima at internuclear distances even shorter than the ground state bond length.) Su *et al*. reason that there are two possible competing explanations for the breaking of the usual bond length-bond energy relationship that might be distinguished by computation. If the MO picture is correct, the difference between the $X\,^1\sum_g^+$ and $a^3\Pi_u$ states is that a π bond is weakened and a σ bond strengthened in the latter as compared to the former. This is in agreement with the view of Sherrill and Piecuch, who propose that a suspended π bond pair should be shorter than a σ bond-π bond pair[111]. Alternatively, in the other extreme represented by the maximum coupling picture, transition from the $X\,^1\sum_g^+$ to the $a^3\Pi_u$ state corresponds to the transformation of a two-electron π bond into a (longer, but not necessarily weaker) one-electron π bond.

The calculations of Su *et al*.[142] used the VBSCF and VBCIS (CI with single excitations) methods as implemented in the Xiamen valence bond (XMVB) program package[143], with all eight valence electrons incorporated in the VB treatment, and strictly localised atomic orbitals on each C atom. (The level of electronic description employed was the modest 6-31G* basis set.) Of the 1764 canonical VB structures generated by eight electrons in eight orbitals, Su *et al*. used a grouping approach to select structures contributing to a particular picture of chemical bonding in C_2. For example, in order to describe a bonding pattern comprising one σ bond, two π bonds, and two unpaired electrons occupying outward-pointing hybrid orbitals (labelled s_1 and s_2 by the authors), contributing structures include the fully covalent one in which all the appropriate singly-occupied orbitals spin-pair (p_{1z} with p_{2z}, π_{1x} with π_{2x}, *etc*.), together with a series of ionic structures where one or more of the orbitals is doubly-occupied while its partner is empty, following the spirit of the idea of the two-electron valence bond which can be described as the superposition of a covalent configuration and two ionic configurations. In the scheme of Su *et al*., doubly ionic configurations are also considered, but only if each carbon atom is left individually electroneutral through possession of a doubly occupied AO and an empty AO. This leads to a set of 21 VB configurations that contribute to the triply-bonded structure (see Figure 12). In addition to this structure, others include one containing two "suspended" π bonds (29 VB configurations) and another containing a σ bond and a π bond (14 VB configurations). Several other sets of configurations are required in order to give a correct description of dissociation, including a set with six electrons in σ orbitals and only two in π orbitals, which are known from previous work by Sherrill[110,111] to become important at internuclear distances greater than about 1.7 Å.

With this computational methodology, Su *et al*. examined the relative weighting of several groups of configurations at the equilibrium internuclear

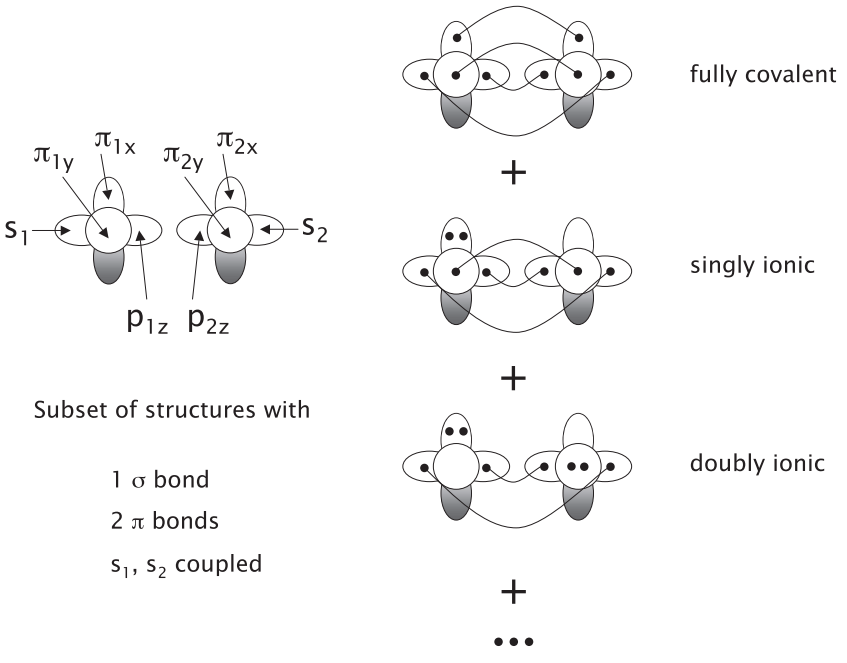


Figure 12 *Some of the valence bond structures featured in the calculations of Su* et al.[142]

distance (1.262 Å for VBCIS) and with the bond stretched to a length of 2.0 Å. At the equilibrium distance, the triply-bonded structure dominates, followed by the $\sigma+\pi$ doubly-bonded structure, with only a small contribution coming from the structure containing a "suspended" π double bond. At 2.0 Å the situation is completely different, with contributions to the electronic structure spread over several groups, including those originating in the configuration $(\sigma_g 2s)^2\,(\sigma_u^* 2s)^2\,(\pi_u 2p)^2\,(\sigma_g 2p)^2$ (with six electrons in σ orbitals and two in π orbitals). Su *et al*. also computed the contributions of the different groups of structures to the energetic stabilisation of C_2, and demonstrated that inclusion of the $\sigma+\pi$ structures has a much larger energy-lowering effect than inclusion of the "suspended" π structures (by 62 kJ mol^{-1} as compared to 9 kJ mol^{-1}). This is reconciled with the usual MO diagram for C_2 (Figure 1) by noting that while the $\sigma_g 2s$ orbital is strongly bonding in character, the $\sigma_u^* 2s$ is only weakly antibonding, leaving some residual σ character in the overall bonding type.

In valence bond theory, any electron-pair bond can be considered as a superposition of a covalent contribution (the Heitler–London contribution) and two ionic contributions (with the two electrons localised on one or the other nucleus), even in the case of a homonuclear diatomic molecule. The relative weighting of these contributions is indicative of the type of bond present. If the Heitler–London contribution dominates, the bond is rather covalent; if one of the ionic contributions is large (in a heteronuclear molecule), the bond is somewhat ionic. In homonuclear molecules, the covalent contribution

is invariably the dominant one. However, in 1992 Shaik and co-workers identified a phenomenon they named "charge-shift bonding"[144,145], in which the Heitler–London configuration is lower in energy than either of the ionic configurations, but contributes only a minority of the overall binding, the remainder coming from the "charge-shift resonance energy" due to the admixture of the ionic configurations into the variationally optimal wavefunction. In the case of homonuclear diatomics, this admixture is symmetrical. Strikingly, this approach gives clear insight into the bonding in F_2, where it is well known that the Hartree–Fock method yields a purely dissociative potential energy curve[146], in obvious contradiction with experiment. From an MO theory perspective, this is tackled by using a multiconfiguration wavefunction. Shaik *et al*.'s VB calculations on F_2 showed that while the Heitler–London structure yields an energy 148 kJ mol^{-1} above that of two ground state F atoms at infinite separation, the perturbative inclusion of ionic-covalent mixing of the equivalent F^+F^- and F^-F^+ structures (which by themselves have an energy over 1000 kJ mol^{-1} higher than the sum of the separated atom energies) produced a "resonance energy stabilisation" of the molecule amounting to 314 kJ mol^{-1}, leading to a dissociation energy of 166 kJ mol^{-1}, close to the experimental value of 160 kJ mol^{-1}.

Su *et al*. examined the characteristics of the individual σ and π contributions to the multiple bond in C_2 in a similar fashion. Considering only the triply-bonded structure (with contributing configurations as shown in Figure 12), they evaluate the quantities

$$RE^{\sigma} = E\left(\Psi^{(1)}_{\sigma-cov}\right) - E\left(\Psi^{(1)}_{full}\right)$$

and

$$RE^{\pi} = E\left(\Psi^{(1)}_{\pi-cov}\right) - E\left(\Psi^{(1)}_{full}\right)$$

where the superscript (1) indicates selection from the group of 21 VB configurations corresponding to the triply-bonded structure (their "group 1"), the quantity $E\left(\Psi^{(1)}_{full}\right)$ refers to the variationally-optimised VBSCF energy for this group of configurations, and the quantities $E\left(\Psi^{(1)}_{\sigma-cov}\right)$ and $E\left(\Psi^{(1)}_{\pi-cov}\right)$ refer to the energy obtained by variational optimisation over only that subset of the 21 configurations in which, respectively, the σ or π bond is Heitler–London paired. Using this approach they determined that the π resonance energy is about three times larger than the σ resonance energy (though both bonds have similar overall strength), indicating that while the σ bond is largely a conventional covalent bond the π bond has substantial charge-shift character. Overall, the bonding is dominated by the triply-bonded structure, but with some contribution from the $\sigma+\pi$ doubly-bonded structure, accounting for the experimental fact that the bond length in C_2 is intermediate between that in ethylene and that in acetylene, but lies closer to the latter.

A follow-up paper in 2012 from a group including some of the same researchers[11], referred to above, proposes that C_2 (along with its isoelectronic species CN^+, BN, and CB^-) possesses a *quadruple* bond, with four separate types of interaction all contributing to the bonding picture. The first three of the four bonds are the familiar σ and π bonds discussed above, while the fourth is the additional energetic stabilisation obtained through the interaction of the outward-pointing sp hybrids with one another. The argument in this paper is based on the calculation of the *in situ* bond energy, D_{in}, which is an analogue of the bond dissociation energy D_e applicable to the individual bonding contributions in a multiple bond. The concept was introduced in a 1995 paper by Hiberty *et al*.[147], and was used by Su *et al*. in the paper previously discussed[142] to estimate the strength of the σ and π bonds in C_2. While D_e can, in principle, be calculated directly from the difference between the energy of the bonded state at the optimum internuclear separation and that of the separated atoms at infinity, no similar straightforward approach exists for the calculation of D_{in}. That is, the separated-atom state plays the role of a nonbonded "reference state" for the calculation of the dissociation energy; the problem becomes one of choosing a suitable reference state for the calculation of D_{in}.

One approach is to use the concept of the "quasiclassical" (QC) state, an idea introduced by Werner Kutzelnigg to clarify the physical origins of the chemical bond in archetypal species such as H_2[148]. In Kutzelnigg's picture, bonding is a three-stage process consisting of: (1) the quasiclassical interaction of the unchanged electronic charge distributions of the separated atoms; (2) the interference of atomic orbitals leading to charge buildup in the binding region and lowering of the kinetic energy; and (3), the deformation of the molecular orbitals to restore the balance between kinetic and potential energies in accordance with the virial theorem. The QC state for H_2, for example, consists of a spin-wave determinant (with α spin localised on one atom and β spin on the other)

$$|a\bar{b}| = \frac{1}{\sqrt{2}}\{a(1)b(2)[\alpha(1)\beta(2)] - a(2)b(1)[\alpha(2)\beta(1)]\}$$

where the numbers label electron coordinates and the letters refer to nuclei. This is not a valid spin eigenfunction, but by antisymmetric or symmetric mixing with $|\bar{a}b|$ yields, respectively, a singlet wavefunction (in fact, the Heitler–London wavefunction) or a triplet wavefunction. The energy of this wavefunction can be represented[147] as

$$E_{QC} = \epsilon_a + \epsilon_b + V_{nn} + J_{ab} + (V_{en})_{ab} + (V_{en})_{ba}$$

where the first two terms are the isolated atom energies, the third is the nuclear repulsion, the fourth is the electron–electron repulsion, and the last two represent the attraction between an electron localised on one atom and the other nucleus. In a truly classical system, the last four terms would sum to zero and the energy curve as a function of internuclear distance would be flat. In a real

molecular system, the behaviour of the energy at small internuclear distances is non-classical, but for $r > r_e$ the QC state does indeed behave approximately classically, and serves to represent a pair of nonbonding atoms at a given internuclear distance.

The QC state concept can be extended fairly easily to larger systems, or systems containing multiple bonds. Practically speaking, the orbitals involved in the bond of interest are replaced by orthogonalised α and β atomic orbitals without changing any other part of the wavefunction, and the energy recalculated. Then the *in situ* bond energy is given by

$$D_{in} = E(\Psi_{QC}) - E(\Psi_{full})$$

where Ψ_{QC} denotes the state where the bond of interest is replaced by spin-alternant orbitals. Su *et al*.[142] note that for C_2 the QC approach is likely to give more accurate energies for the π components of the bond than for the σ component, because the equilibrium internuclear distance in C_2 is much shorter than a typical C–C σ-bond.

Another approach to the computation of D_{in} is to make use of the fact that it is approximately one half of the energy required to promote the electrons in the bond of interest from their singlet-coupled state to a triplet-coupled state[149]. (The two approaches to obtaining a value for D_{in} are illustrated in Figure 13.) Shaik *et al*. note that the sum of the D_{in} values for all bonded electron pairs (which they refer to as the "intrinsic bonding energy") is equivalent to the energy difference between the molecular electronic ground state ($X\,^1\Sigma_g^+$) and the energy of the two separated C atoms in the electronically excited ^5S state (with four unpaired electrons per atom); that is,

$$\sum_{bonds} D_{in} = \Delta E\left(C_2\,(^1\Sigma_g^+) \rightarrow 2C(^3P)\right) + 2\Delta E\left(C(^3P) \rightarrow C(^5S)\right)$$

where the first term is the usual bond dissociation energy (BDE), corresponding to separation of the molecule into two atoms in their ground states.

Using the second approach, Shaik *et al*.[11] calculate D_{in} for the "fourth bond" to be approximately 62 kJ mol^{-1} at the FCI/6-31G* level. (Note that the electron configuration corresponding to the triplet state denoted $c\,^3\Sigma_g^+$ in this work differs from that given by Su *et al*. The configuration given by Shaik and reproduced here in Figure 13 is the leading term in the FCI expansion of the state, with a mixing coefficient of 0.889.) The experimentally-determined energy difference between the $X\,^1\Sigma_g^+$ and $c\,^3\Sigma_g^+$ states (computed from the difference between the Swan and Schmidt–Kable band origins, supplemented by the energy difference between $X\,^1\Sigma_g^+$ and $a^3\Pi_u$) is around 9230 cm^{-1} or 110 kJ mol^{-1}, approximately twice the value of D_{in} obtained in this way. Similar results are obtained using the QC approach.

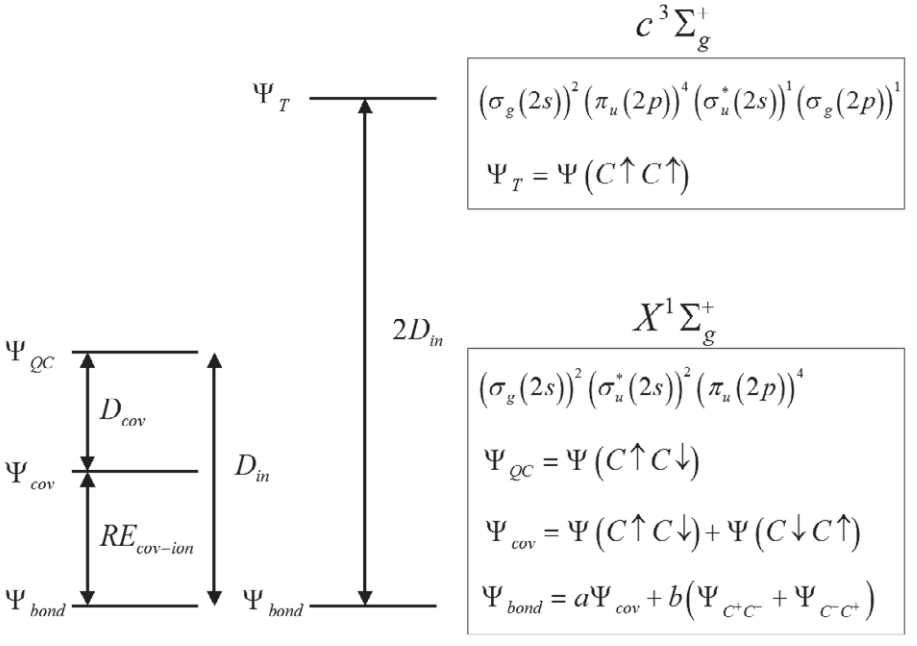


Figure 13 *Two approaches to the calculation of the "*in situ *bond energy",* D_{in}. *On the left,* D_{in} *is given as the difference between the energy of the bound state with wavefunction* Ψ_{bond} *and that of the quasiclassical state (with respect to the particular bond of interest)* Ψ_{QC}; *on the right,* D_{in} *is given as approximately one half of the energy required to uncouple the paired electrons of the bond and form the corresponding triplet state. The triplet state shown here corresponds to uncoupling of the electrons of the "fourth bond". After Shaik* et al.[11]

Similar calculations on the molecule's σ and π bonds lead to *in situ* bond energies of 420 kJ mol^{-1} and 394 kJ mol^{-1} (for *each* π bond), respectively. Thus, Shaik *et al*. conclude that the "fourth bond", originating in the interaction of the outward-pointing hybrids, contributes about 5% of the intrinsic bonding energy (which has a value of around 1270 kJ mol^{-1}, approximately twice the conventionally-defined bond dissociation energy for C_2).

Does this imply that C_2 contains a quadruple bond? (See Figure 14). The answer depends on the definition of bond order one chooses. While the Wiberg bond index for C_2 (one of several definitions of bond order available in computational chemistry) obtained by Shaik *et al*. using a natural bond orbital (NBO) calculation and Kohn–Sham densities is about 3.7, greater than

Figure 14 *Formally quadruply-bonded* C_2.

the bond order in acetylene and N_2 (each of which has a bond order of 3.0 by the same measure), by the rather simpler (perhaps *too* simple) measure of the conventionally-defined bond dissociation energy, or indeed of bond length, reasonable arguments could be made in favour of a bond order slightly less than 3. At any rate, the findings of Shaik *et al*. do imply that there are four distinct contributing types of interaction present in the bonding in the ground state of C_2. (Additionally, they find similar "quadruple" bonding to prevail in the isoelectronic species BC^- and CN^+, but only double bonds to exist in the heavier group 14 diatomics Si_2 and Ge_2.)

3.5 Other insights

Other recent work gives somewhat different perspectives on the bonding in C_2. In a 2013 paper, Ramos-Cordoba *et al*.[150] presented a study of the bonding in several main-group diatomic molecules using the method of *local spin analysis*[151], and concluded that the ground state of C_2 possesses significant diradical character. The essence of the local spin approach is to decompose the expectation value of the $\widehat{S^2}$ operator into atomic and diatomic contributions according to

$$\langle \widehat{S^2} \rangle = \sum_A \langle \widehat{S^2} \rangle_A + \sum_{A,B \neq A} \langle \widehat{S^2} \rangle_{AB}$$

For a prototypical covalent bond in a homonuclear diatomic molecule in a singlet state (such as the ground state of H_2), $\langle \widehat{S^2} \rangle_A$ and $\langle \widehat{S^2} \rangle_{AB}$ must be equal in magnitude and opposite in sign in order to make $\widehat{S^2}$ equal to zero, as required. Useful quantities such as the number of effectively unpaired electrons

$$u_A = 4n(1-n)$$

and the local spin components

$$\langle \widehat{S^2} \rangle = -\langle \widehat{S^2} \rangle_{AB} = \tfrac{3}{2} n\,(1-n)(1+4\delta^2)$$

can be defined in terms of the occupancies (respectively $1-n$ and n) of the bonding and antibonding natural spin orbitals and the overlap integral between these orbitals (defined over the domain of a particular atom), $S_{ij}^A = -S_{ij}^B = \delta$. (The "diagonal" versions of these integrals, S_{ii}^A and S_{ii}^B, have a value of 1/2, as they correspond to density integrals over one half of the molecule.) Thus, for a single-determinant treatment of the molecule, because the antibonding orbital is unoccupied ($n=0$), both u_A and $\langle \widehat{S^2} \rangle_A$ are zero and there is no local spin. This corresponds to two electrons perfectly paired in a canonical covalent bond. When electron correlation is introduced, n takes on a positive value, and the number of effectively unpaired electrons deviates from zero. "Local spin" appears.

Ramos-Cordoba *et al*. considered all period 2 homonuclear diatomics except for the unstable Ne_2, as well as BN, BC^- and CN^+ (isoelectronic with C_2), and Si_2

(a period 3 analogue of C_2), performing a local spin analysis on wavefunctions computed at the CASSCF/cc-pVTZ level of theory. In C_2 both u_A and $\langle \widehat{S}^2 \rangle_A$ are large (respectively 1.10 and 0.81), indicating a significant number of effectively unpaired electrons as well as substantial local spin character (with $\langle \widehat{S}^2 \rangle_A$ even larger than the expected value of 0.75 for a single localised electron); the authors interpret this result as implying that the molecule can be characterised as a diradical. (The unpairing is fairly evenly divided between σ and π electrons.) The diradical character is less pronounced in the heteronuclear isoelectronic molecules and ions, and in Si_2.

The findings of this paper are somewhat substantiated by a 2014 paper by Xu and Dunning[152], which again makes use of a GVB approach. This work focuses on detailed analysis of the GVB wavefunctions as a function of internuclear distance, and concludes that the molecule is best described as possessing a fairly traditional covalent σ bond (so-called "perfect pairing"), but with the remainder of the unpaired electrons better described as antiferromagnetically coupled (with some contribution from perfect pairing). As a result of the complexity of the bonding, the authors propose that it is not possible to assign a specific bond order to C_2.

The GVB wavefunctions for C_2 used by Xu and Dunning are based on optimised GVB orbitals for atomic carbon. These differ somewhat from the pure s and p orbitals usually encountered, with hybrid *lobe orbitals* (see Figure 15) formed from linear combinations of 2s and $2p_z$ orbitals giving a better description of the carbon 2s lone pair than the doubly-occupied 2s orbital of Hartree–Fock theory. (The optimised mixing coefficients are $c_{2s} = 0.93$ and $c_{2p} = 0.36$.) Using this set of spatial wavefunctions, the GVB wavefunction for the ^3P ground state of atomic carbon can be written

$$\Psi_{GVB}\,[C(^3P)] = \hat{a}\phi_{1s}\,\phi_{1s}\,\phi_{2s\_}\,\phi_{2s+}\,\phi_{2px}\,\phi_{2py}\,\alpha\beta\,\frac{1}{\sqrt{2}}\,(\alpha\beta - \beta\alpha)\,\alpha\alpha$$

where $\hat{a}$ denotes an "antisymmetriser" (permutation operator to make the overall wavefunction antisymmetric with respect to interchange of electron labels), and the wavefunction is written as a product of a spatial part (ϕ terms) and a spin part (α and β terms). The above wavefunction can then be understood as describing a state exhibiting paired 1s electrons, singlet-coupled $2s_+$ and $2s_-$ electrons, and triplet-coupled 2p electrons.

For C_2, with eight valence electrons singly occupying eight orbitals, the spins may be coupled in 14 linearly independent ways; thus, the general spin function can be written as a linear combination of these 14 spin coupling patterns

$$\Theta = \sum_{k=1}^{14} c_k \Theta_k$$

and the weighting coefficients $w_k = c_k^2$ then give an indication of the relative importance of particular spin couplings in the overall GVB wavefunction.

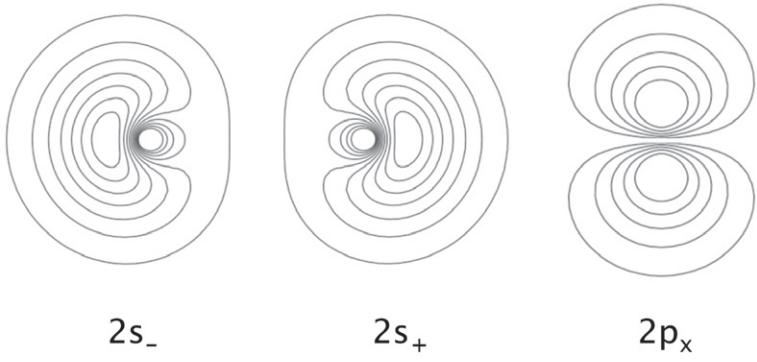


Figure 15 *Valence orbitals for the* ^{3}P *ground state of the carbon atom, including the* $2s_+$ *and* $2s_-$ *lobe orbitals and a representative 2p orbital. After ref. 152.*

The approach of Xu and Dunning is to consider the evolution of the w_k as a function of internuclear distance. An important point to note is that while the overall GVB energy does not depend on the ordering of the basis functions, the mathematical form of the spin function Θ *does* depend on orbital order; the selection of an orbital order that yields a particularly simple form of Θ (dominated by a single term, or only a few terms) leads to insight into the hierarchy of electron–electron interactions in the molecule.

For a consideration of the separated atom limit ($r \rightarrow \infty$), a sensible choice of the ordering of the basis orbitals is ($2s_{-A}$,$2s_{+A}$,$2s_{-B}$,$2s_{+B}$,$2p_{xA}$,$2p_{yA}$,$2p_{xB}$,$2p_{yB}$). This leads to a wavefunction with a very straightforward form, requiring only one of the 14 possible spin coupling patterns. Thus, considering valence orbitals only,

$$\Psi_{GVB}\,[C_2(X^1\Sigma_g^+, \mathrm{r}\rightarrow\infty)]\;\hat{a}\phi_{2s_{-A}}\,\phi_{2s_{+A}}\phi_{2s_{-B}}\phi_{2s_{+B}}\phi_{2p_{xA}}\phi_{2p_{yA}}\phi_{2p_{xB}}\phi_{2p_{yB}}\Theta$$

with

$$\Theta = \frac{1}{\sqrt{2}}(\alpha\beta-\beta\alpha)\;\frac{1}{\sqrt{2}}(\alpha\beta-\beta\alpha)\left\{\frac{1}{\sqrt{12}}[2\alpha\alpha\beta\beta-(\alpha\beta+\beta\alpha)(\alpha\beta+\beta\alpha)+2\beta\beta\alpha\alpha]\right\}$$

This can be interpreted as singlet coupling of the first two pairs of electrons ($2s_{-A}$,$2s_{+A}$) and ($2s_{-B}$,$2s_{+B}$), and triplet coupling of the remaining two electrons on each individual atom (in the $2p_x$ and $2p_y$ orbitals), followed by recoupling of the two separated-atom spin triplets into a singlet. For example, the term $\alpha\alpha\beta\beta$ implies that the p electrons on atom A both have α spin while those on atom B both have β spin. This pattern matches the intuitive picture of two ^3P carbon atoms spin-coupling antiferromagnetically to yield an overall singlet state.

Where the analysis of Xu and Dunning grows more interesting is in the region of the equilibrium internuclear distance (at $r_e = 1.247$ Å). In this region, a different choice of orbital ordering (referred to by the authors as the *molecular orbital ordering*) suggests itself, namely ($2s_{+A}$,$2s_{-B}$,$2s_{-A}$,$2s_{+B}$,$2p_{xA}$,$2p_{xB}$,$2p_{yA}$,$2p_{yB}$). With a "perfect pairing" (PP) spin coupling pattern, that is,

$$\Theta_{PP} = \frac{1}{\sqrt{2}}(\alpha\beta - \beta\alpha)\frac{1}{\sqrt{2}}(\alpha\beta - \beta\alpha)\frac{1}{\sqrt{2}}(\alpha\beta - \beta\alpha)\frac{1}{\sqrt{2}}(\alpha\beta - \beta\alpha),$$

this ordering describes the formation of four electron pairs. These are: the σ-bonding pair ($2s_{+A}$,$2s_{-B}$), the π-bonding pairs ($2p_{xA}$,$2p_{xB}$) and ($2p_{yA}$,$2p_{yB}$), and the "fourth bond" pair formed from the outward pointing hybrids ($2s_{-A}$,$2s_{+B}$). Starting from the separated atom state, as the internuclear distance is shortened the spatial orbitals generated from the first three of these pairs polarise and delocalise onto the other atom following the familiar pattern of bond formation, while the ($2s_{-A}$,$2s_{+B}$) pair remains rather unchanged in form. The weighting coefficients of the 14 spin-coupling patterns also change. At the equilibrium internuclear distance, with this orbital ordering, the dominant spin-coupling pattern is indeed the PP pattern, but its weighting coefficient is only $W_{PP} = 0.67$, much smaller than the value of 0.96 found in N_2. Thus, while a picture of C_2 as a molecule containing four covalent bonds is given some justification by this result, it is clear that this is not the whole story. Such an interpretation is reinforced by the finding that if the GVB calculation is carried out using *only* the PP component of the wavefunction, the total energy at r_e is raised by about 84 kJ mol^{-1} relative to the full GVB energy.

Further insight can be found by means of yet another choice of orbital ordering (which the authors refer to as the *quasi-atomic ordering*); in this ordering, the orbitals participating in the σ bond are listed first, and the remainder are listed by centre. That is, the ordering can be written ($2s_{+A}$,$2s_{-B}$,$2s_{-A}$,$2p_{xA}$,$2p_{yA}$,$2s_{+B}$,$2p_{xB}$,$2p_{yB}$). With the basis orbitals rearranged in this fashion, the leading spin function now dominates strongly, with a weighting coefficient of 0.918. The mathematical form of the dominating spin function is rather complicated, namely

$$\begin{aligned}\Theta = \frac{1}{\sqrt{2}}(\alpha\beta - \beta\alpha)\frac{1}{6}\{&3\alpha\alpha\alpha\beta\beta\beta + 3\beta\beta\beta\alpha\alpha\alpha \\ &+ (\alpha\beta + \beta\alpha)[\beta\alpha\alpha\beta - \beta\beta\alpha\alpha - (\alpha\beta - \beta\alpha)\beta\alpha] + \beta\beta(\alpha\alpha\alpha\beta - \alpha\beta\alpha\alpha) \\ &+ (\beta\beta\alpha\alpha - \alpha\alpha\beta\beta)\beta\alpha - (\alpha\alpha\beta + \alpha\beta\alpha + \beta\alpha\alpha)(\alpha\beta + \beta\alpha)\beta\},\end{aligned}$$

but it can be understood more easily by considering it as a path taken on a Kotani branching diagram[132], which defines the construction of the spin function sequentially as a series of steps where one further spin is added to an existing function. In this context it can be represented $\alpha\beta\alpha\alpha\beta\beta\beta$, indicating that the first two electrons are singlet-coupled while the remaining six are to be understood as quartet-coupled electron triads localised on each atom (($2s_{-A}$,$2p_{xA}$,$2p_{yA}$) and ($2s_{+B}$,$2p_{xB}$,$2p_{yB}$)) which are then recoupled *antiferromagnetically* to yield an overall singlet state. (See Figure 16). Additionally, applying the same trick as before, if the GVB wavefunction is computed using *only* this spin function, a total energy higher than the fully optimised energy by about 45 kJ mol^{-1} is obtained, indicating that this spin configuration, while again not the whole story of bonding in C_2, represents a bigger contribution to it than does the PP configuration.

The authors suggest that the incompleteness of perfect pairing as a model of the bonding in C_2 (as compared, for example, to N_2) is a result of Pauli exchange repulsion. For electrons in closed shells, the quantum-mechanical exchange interaction effectively acts as a repulsive force, as the Pauli principle prevents all the electrons from occupying the same region of space; this is the origin of nonbonded "steric repulsion" effects seen widely across chemical systems. In the particular case of C_2, this effect applies particularly to the "inner" ($2s_{+A}$,$2s_{-B}$) and "outer" ($2s_{-A}$,$2s_{+B}$) σ-electron pairs, which share similar spatial regions.

The conclusions arrived at by Xu and Dunning are challenged in a very recent paper by the Hiberty–Shaik collaboration[154] on the basis of several sets of GVB calculations. One of their approaches used the BOVB (breathing-orbital valence bond) method, applied to only a single bond at a time, to improve the GVB description of that particular bond and assess its covalent character (described by the weighting coefficient $w_{cov} = 1 - w_{ionic}$ of the Heitler–London component of the GVB wavefunction) and bond energy. With a 6-311G(d) basis set the w_{cov} (renormalised value) and D of the "fourth bond" were found to be 0.849 (quite large) and 84 kJ mol^{-1} (about double the VBSCF value), as compared to 0.648 and 400 kJ mol^{-1} for the π bond, indicating that the fourth bond is "quite strong" (about the same strength as the single bond in Li_2) and "quite covalent" (more so than the π bond).

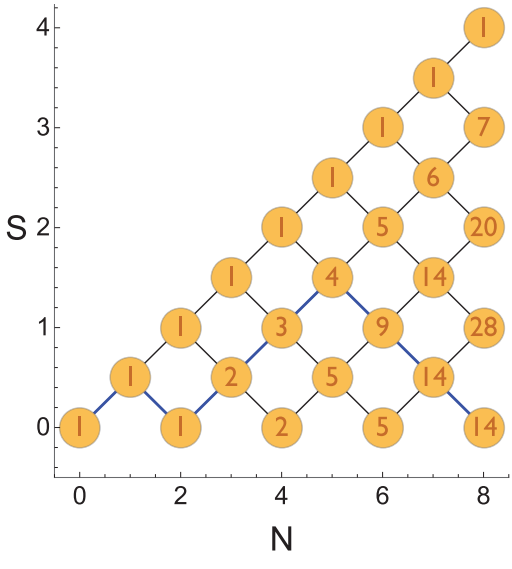


Figure 16 *Branching diagram for a system of eight electrons. First introduced by Van Vleck and Sherman in 1935[153], and discussed in more detail by Pauncz[132], such diagrams are a visualisation tool useful in the construction of many-electron spin eigenfunctions. The x-axis shows the number of electrons, the y-axis shows the total spin, and the vertices are labelled with the number of linearly independent spin eigenfunctions having that total spin, which is equal to the number of paths from the origin to that point. The path shown in blue represents the antiferromagnetically-coupled eigenfunction discussed in the text.*

A second set of calculations strove to assess the degree to which the results obtained by Xu and Dunning were unusual by carrying out an identical analysis on the bonding in the less controversial molecule acetylene, HC≡CH, usually accepted as an example of a canonical triple bond. Using the Kotani scheme for construction of spin eigenfunctions, the GVB-PP contribution to the bonding in acetylene was found to be about 77%, somewhat larger than the 67% found in C_2, but still quite far from 100%. Repeating the calculations on both C_2 and HC≡CH using the alternative Rumer-type spin eigenfunctions led to PP contributions of 88% and 93% respectively, indicating both that the assessment of PP weighting depends somewhat on the spin coupling scheme employed and that C_2 appears not to be as much of a special case as suggested by the calculations of Xu and Dunning. Further calculations on related species (such as BC^- and CN^+) may elucidate the matter.

4. Conclusions

Despite its ephemeral character, the dicarbon molecule continues to act as a touchstone both for new developments in experimental spectroscopy and for fundamental theories of chemical bonding. The first identification of C_2 as the carrier of the Swan bands so ubiquitous in hydrocarbon flame spectra roughly coincided with the beginnings of quantum theory, conferring upon the molecule a recurring role in the interplay between the development of theories of electronic structure, the quantum-theoretical prediction of the features of vibronic spectra, and the measurement of spectra over wider frequency ranges and with ever better resolution.

From the spectroscopic standpoint, the biggest surprise that C_2 had up its sleeve was the fact that its most prominent spectral feature – the Swan bands – was due to a transition between two states neither of which is the ground state. Just as with the hydrogen atom spectrum, where the first features discovered were the Balmer transitions to $n=2$, this is partially an accidental result of the fact that the human eye is attuned to a particular photon energy range. However, the early assumption that what we now call the $a^3\Pi_u$ state, and not the $X\,^1\sum_g^+$ state, was the electronic ground state, and the later overturning of that assumption, served as an important reminder that subtle effects can be at play in the electronic structure of simple species and as an incentive in the development of more accurate computational methods. Indeed, the fact that the ground singlet and lowest triplet states are so close to one another in energy (within 0.1 eV) is quite unusual, and indicates that something interesting is going on in the bonding.

Over the years, a dozen or more valence states and several Rydberg states were discovered in the singlet and triplet manifolds, with transitions to and from these states lying in regions from the infrared (Bernath, Phillips, Ballik–Ramsay) to the far ultraviolet (Mulliken, Freymark, Fox–Herzberg). Several of these transitions are of astrophysical significance, as indicators of

carbon content in the diffuse molecular clouds lying between Earth and various distant stellar light sources. Within the past decade, new band systems involving the $c\ ^3\sum_u^+$ state have been studied in detail for the first time, and exploration of an entirely new spin manifold, the quintet, has begun, with the first observation of a quintet-quintet transition having been made only in the past year. Thus, the sesquicentenarian field of C_2 spectroscopy remains in good health.

Perhaps surprisingly for a period 2 diatomic molecule, the nature of the bonding in C_2 has also continued to excite interest (and quite heated debate). Superficially, the central question might be framed: does dicarbon "really" possess a quadruple bond? While no clear answer to this question may be possible, it seems that persuasive arguments can be made for the presence of four distinct types of orbital interaction all of which contribute significantly to the bonding in the molecule. Additionally, the debate on this topic makes it clear that modern GVB theory has much to contribute to qualitative and quantitative elucidation of questions related to chemical bonding. The debate on C_2 also benefits the chemical community by bringing back into light fundamental topics such as the definition of bond order, which may not be as clear-cut as is widely assumed.

References

1. Swan, W. (1857) *Trans. R. Soc. Edinb.*, **21**, 411–429.
2. Gaydon, A.C. (1974) *The spectroscopy of flames*, 2nd edn. Chapman and Hall, London.
3. Ferguson, R.E. (1955) *J. Chem. Phys.,* **23**, 2085–2089.
4. Marques, C.S.T., Benvenutti, H. and Bertran, C.A. (2006) *J. Braz. Chem. Soc.,* **17**, 302–315.
5. Van Orden, A. and Saykally, R.J. (1998) *Chem. Rev.*, **98**, 2313–2357.
6. Dufour, P. (2012) In: Hoard, D.W. (ed.) *White dwarf atmospheres and circumstellar environments,* Chap. 3, p. 53. John Wiley and Sons.
7. Lange, H., Huczko, A. and Byszewski, P. (1996) *Spectrosc. Lett.,* **29**, 1215–1228.
8. Sorkhabi, O., Blunt, V.M., Lin, H., A'Hearn, M.F., Weaver, H.A., Arpigny, C. and Jackson, W.M. (1997) *Planet. Space Sci.,* **45**, 721–730.
9. Huber, P.K. and Herzberg, G. (1979) *Molecular spectra and molecular structure IV. Constants of diatomic molecules*, Van Nostrand-Reinhold. (Data available from the NIST Chemistry WebBook, *webbook.nist.gov*.)
10. Darwent, B. de B. (1970) *Bond dissociation energies of simple molecules*, *Nat. Stand. Ref. Data Ser.*, Nat. Bur. Stand. (USA).
11. Shaik, S., Danovich, D., Wu, W., Su, P., Rzepa, Henry, S. and Hiberty, P.C. (2012) *Nature Chem.*, **4**, 195–200.
12. Wollaston, W.H. (1802) *Phil. Trans. Roy. Roc. Lond.*, **92**, 365–380.
13. Jensen, W.B. (1986) In: Stock, J.T. and Orna, M.V. (eds) *The history and preservation of chemical instrumentation*, pp. 123–149. Springer, New York.
14. Fraunhofer, J. (1815) *Denkschriften der Königlichen Akademie der Wissenschaften zu München*, **5**, 193–226.
15. Attfield, J. (1862) *Phil. Trans. Roy. Soc. Lond.,* **152**, 221–224.
16. Ångström, A.J. and Thalén, R. (1975) *Nova Acta Reg. Soc. Sc. Upsal.*, **9**.
17. Smith, C.P. (1875) *Phil. Mag.,* **49**, 24.
18. Brand, J.C.D. (1995) *Lines of light: the sources of dispersive spectroscopy, 1800–1930*. Overseas Publishers Association (Gordon and Breach) Amsterdam.
19. Johnson, R.C. (1927) *Proc. Roy. Soc. Lond. A,* **226**, 157–230.
20. Watts, W.M. (1875) *Phil. Mag.,* **49**, 104–106.

21. Bernath, P.F. (2005) *Spectra of atoms and molecules*, 2nd edn. Oxford University Press.
22. Ogilvie, J.F. (2014) *Resonance,* **19**, 834–839.
23. Assmus, A. (1992) *Hist. Stud. Phys. Biol., Sci*., **22**, 209–231.
24. Fortrat, R. (1924*) J. Phys. Radium,* **5**, 33–50.
25. Sommerfeld, A. (1931) *Atomic structure and spectral lines*, 2nd English edn. E.P. Dutton, Aberdeen.
26. Ehrenfest, P. (1913) *Verh. der Deutsche. Physikalische Gesellschaft,* **15**, 451–457.
27. Kragh, H. (2012) *Arch. Hist. Exact. Sci.,* **66**, 199–240.
28. Bjerrum, N. (1949) *Selected Papers*, E. Munksgaard, Copenhagen. (The paper "On the infrared absorption of gases" originally appeared in *Festschrift W. Nernst* (1912) Halle, pp. 90–98.)
29. Kemble, E. (1920) *Phys. Rev.*, **15**, 95–109.
30. Schwarzschild, K. (1916) *Sitzungsberichte d. Preuss. Akad. Wiss.*, 548–568.
31. Shea, J.D. (1927) *Phys. Rev.,* **30**, 825–843.
32. Herzberg, G. (1950) *Spectra of diatomic molecules*, 2nd edn. Van Nostrand Reinhold, New York.
33. Lefebvre-Brion, H. and Field, R.W. (2004) *The spectra and dynamics of diatomic molecules*, Elsevier, Amsterdam.
34. Brown, J.M., Carrington, A. (2003) *Rotational spectroscopy of diatomic molecules*. Cambridge University Press.
35. Bernath, P. (2002) In: Wilson, S. (ed.) *Handbook of molecular physics and quantum chemistry*, Vol. 3, Chap. 16, Wiley, New York.
36. Harris, D.C. and Bertolucci, M.D. (1978) *Symmetry and spectroscopy*. Oxford University Press.
37. Atkins, P.W. and Friedman (1997) *Molecular quantum mechanics*, 3rd edn. Oxford University Press.
38. Ballik, E.A. and Ramsay, D.A. (1959) *J. Chem. Phys.,* **31**, 1128.
39. Morse, P.M. (1929) *Phys. Rev.,* **34**, 57.
40. Dunham, J.L. (1932) *Phys. Rev.,* **41**, 721–731.
41. Hund, F. (1933) *Handbuch der Physik*, **24**, 561.
42. Budó, A. (1936) *Z. f. Phys*., **98**, 437.
43. Phillips, J.G. and Davis, S.P. (1968) *The Swan system of the C_2 molecule; the spectrum of the HgH molecule*. University of California Press, Berkeley and Los Angeles.
44. Brown, J.M. and Merer, A.J. (1972) *J. Mol. Spectrosc.,* **74**, 488–494.
45. Whiting, E.E., Schadee, A., Tatum, J.B., Hougen, J.T. and Nicholls, R.W. (1980) *J. Mol. Spectrosc.,* **80**, 249.
46. Budó, A. (1937) *Z. f. Phys.,* **105**, 579–587.
47. Kovács, I. (1969) *Rotational structure in the spectra of diatomic molecules*. Akadémiai Kiadó, Budapest, English transl. by L. Nemes.
48. PGOPHER version 8.0, Western C.M. (2014) University of Bristol Research Data Repository, doi:10.5523/bris.huflggvpcuc1zvliqed497r2
49. Brooke, J.S.A., Bernath, P.F., Schmidt, T.W. and Bacskay, G.B. (2013) *J. Quant. Spectrosc., Radiat. Transfer,* **124**, 11–20.
50. Curtis, M.C. and Sarre, P.J. (1985) *J. Mol. Spectrosc.,* **113**, 399–409.
51. Prasad, C.V.V. and Bernath, P.F. (1994) *Astrophys. J.,* **426**, 812–821.
52. Lloyd, G.M. and Ewart, P.E. (1999) *J. Chem. Phys.,* **110**, 385–392.
53. Tanabashi, A., Hirao, T., Amano, T. and Bernath, P.F. (2007) *Astrophys. J.,* **169**, 472–484.
54. Bornhauser, P., Sych, Y., Knopp, G., Gerber, T. and Radi, P.P. (2011) *J. Chem. Phys.,* **134**, 044302.
55. Bornhauser, P., Knopp, G., Gerber, T. and Radi, P.P. (2010) *J. Mol. Spectrosc.,* **262**, 69–74.
56. Pearse, R.W.B., Gaydon, A.G. (1976) *The identification of molecular spectra*, Springer, Berlin.
57. Wallace, L. (1962) *Astrophys. J.,* **7**, 165.
58. Sorkhabi, O., Xu., D.D., Blunt, V.M., Lin, H., Price, R., Wrobel, J.D. and Jackson, W.M. (1998) *J. Mol. Spectrosc.,* **188**, 200.
59. Messerle, G. and Krauss, L. (1967) *Z. Naturforsch.,* **22 a**, 2015–2023.
60. Douay, M., Nietmann, R. and Bernath, P.F. (1988) *J. Mol. Spectrosc.,* **131**, 261–271.

61. Freymark, H. (1951) *Ann. Phys. (Leipzig),* **8**, 221.
62. Bornhauser, P., Marquardt, R., Gourlauen, C., Knopp, G., Beck, M., Gerber, T., van Bokhoven, J.A., Radi, P.P. (2015) *J. Chem. Phys.,* **142**, 094313.
63. Joester, J.A., Nakajima, M., Kokkin, D.L., Nauta, K., Kable, S.H. and Schmidt, T.W. (2007) *J. Chem. Phys.,* **127**, 214303.
64. Deslandres, H. and D'Azambuja (L.) (1905) *C. R. Acad. Sci.,* **140**, 917–920.
65. Mulliken, R.S. (1930) *Z. Electrochem.,* **36**, 603.
66. Fox, J.G. and Herzberg, G. (1937) *J. Chem. Phys.,* **52**, 638.
67. Phillips, J.G. (1948) *Astrophys. J.,* **107**, 389.
68. Ballik, E.A. and Ramsay, D.A. (1958) *J. Chem. Phys.,* **29**, 1418.
69. Ballik, E.A. and Ramsay, D.A. (1963) *Astrophys. J.,* **137**, 61–83.
70. McCarty, M. and Robinson, G.W. (1959*) J. Chim. Phys.,* **56**, 723.
71. Stawikowski, A. and Swings, P. (1960) *Ann. d'ap.*, **23**, 585.
72. Bondybey, V.E. (1976) *J. Chem. Phys.,* **65**, 2296–2304.
73. Milligan, D.E. and Jacox, M.E. (1969) *J. Chem. Phys.,* **51**, 1952–1955.
74. Ballik, E.A. and Ramsay, D.A. (1963) *Astrophys. J.,* **137**, 84–101.
75. Messerle, G. and Krauss, L. (1967) *Z. Naturforsch.,* **22 a**, 2023–2026.
76. Herzberg, G., Lagerqvist, A. and Malmberg, C. (1969) *Can. J. Phys.,* **47**, 2735–2743.
77. Bruna, P.J. and Grein, F. (2001) *Can J. Phys.,* **79**, 653–671.
78. Barsuhn, J. (1972) *Z. Naturforsch.,* **27**, 1031.
79. Bruna, P.J. and Wright, J.S. (1992) *J. Phys. Chem.,* **96**, 1630.
80. Hupe, R., Sheffer, Y. and Federman, S.R. (2012) *Astrophys. J.,* **761**, article id 38, 7pp.
81. van den Burgt, L.J. and Heaven, M.C. (1988*) J. Chem. Phys.,* **87**, 4235.
82. Goodwin, P.M. and Cool, T.A. (1988) *J. Chem. Phys.,* **88**, 4548.
83. Shi, D., Zhang, X., Sun, J. and Zhu, Z. (2011) *Mol. Phys.,* **109**, 1453–1465.
84. Bao, Y., Urdahl, R.S. and Jackson, W.M. (1990*) J. Chem. Phys.,* **94**, 808–809.
85. Wakabayashi, T., Ong, A.-L. and Krätschmer, W. (2002) *J. Chem. Phys.,* **116**, 5996–6001.
86. Kokkin, D., Reilly, N.J., Morris, C.W., Nakajima, M., Nauta, K., Kable, S. and Schmidt, T.W. (2006) *J. Chem. Phys.,* **125**, 231101.
87. Schmidt, T.W. (2010) In: Grunenberg, J. (ed.) *Astronomical molecular spectroscopy*, Chap. 13 of *Computational spectroscopy: methods, experiments and applications*, Wiley-VCH, Weinheim.
88. Nakajima, M., Joester, J.A., Page, N.I., Reilly, N.J., Bacskay, G.B., Schmidt, T.W. and Kable, S.H. (2009) *J. Chem. Phys.,* **131**, 044301.
89. Martin, M. (1992) *J. Photochem. Photobiol. A*,**66**, 263.
90. Kaminski, C.F., Hughes, I.G. and Ewart, P. (1996) *Appl. Phys. B*, **62**, 39–44.
91. Kaminski, C.F., Hughes, I.G. and Ewart, P. (1997) *J. Chem. Phys.,* **106**, 5324–5332.
92. Lloyd, G.M. and Ewart, P. (1999) *J. Chem. Phys.,* **110**, 385–392.
93. Nakajima, M. and Endo, Y. (2013) *J. Chem. Phys.,* **139**, 244310.
94. Schmidt, T.W. and Bacskay, G.B. (2011) *J. Chem. Phys.,* **134**, 224311.
95. Mulliken, R.S. (1927) *Phys. Rev.,* **29**, 637–649.
96. Hund, F. (1928) *Z. Physik,* **51**, 759.
97. Jenkins, F.A. (1948) *Phys. Rev.,* **74**, 355.
98. Mulliken, R.S. (1930) *Rev. Mod. Phys.,* **2**, 60–115.
99. Mulliken, R.S. (1931) *Rev. Mod. Phys.,* **4**, 1–86.
100. Hund, F. (1931) *Z. Physik,* **73**, 1.
101. Mulliken, R.S. (1939) *Phys. Rev.,* **56**, 778–781.
102. Pitzer, K.S. and Clementi, E. (1959) *J. Am. Chem. Soc.,* **81**, 4477–4485.
103. Clementi, E. and Pitzer, K.S. (1960) *J. Chem. Phys.,* **32**, 656–662.
104. Clementi, E. (1960) *Astrophys. J.,* **132**, 898–904.
105. Mulliken, R.S. (1952) *J. Phys. Chem.,* **56**, 295–311.
106. Read, S.M. and Vanderslice, J.T. (1962) *J. Chem. Phys.,* **36**, 2366–2369.
107. Fougere, P.F. and Nesbet, R.K. (1966) *J. Chem. Phys.,* **44**, 285–298.
108. Kirby, K. and Liu, B. (1979) *J. Chem. Phys.,* **70**, 893–900.
109. Schmidt, T.W. and Bacskay, G.B. (2007) *J. Chem. Phys.,* **127**, 234310.

110. Abrams, M.L. and Sherrill, C.D. (2004) *J. Chem. Phys.,* **121**, 9211–9219.
111. Sherrill, C.D. and Piecuch, P. (2005) *J. Chem. Phys.,* **122**, 124104.
112. Hehre, W.J., Ditchfield, R. and Pople, J.A. (1972) *J. Chem. Phys.,* **56**, 2257.
113. Hariharan, P.C. and Pople, J.A. (1973) *Theor. Chim. Acta,* **28**, 213.
114. Kowalski, K. and Piecuch, P. (2000) *J. Chem. Phys.,* **113**, 18.
115. Shi, D., Zhang, X., Sun, J. and Zhu, Z. (2011) *Mol. Phys.,* **109**, 1453–1465.
116. Zhang, X., Shi, D., Sun, J. and Zhu, Z. (2011) *Chin. Phys. B,* **20**, 043105.
117. Booth, G.H., Cleland, D., Thom, A.J.W. and Alavi, A. (2011) *J. Chem. Phys.,* **135**, 084104.
118. Heitler, W. and London, F. (1927) *Z. Phys.,* **44**, 455.
119. Gallup, G.A. (2002) In: Cooper, D.L. (ed.), *Valence bond theory,* Chap. 1, p.1. Elsevier, New York.
120. Magnasco, V. and Costa, C. (2005) *Chem. Phys., Lett.*, **403**, 303–307.
121. Lennard-Jones, J. (1929) *Trans. Faraday Soc*., **25**, 668.
122. Slater, J.C. (1929) *Phys. Rev.,* **34**, 1293.
123. Karadakov, P.B. (1998) *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.*, **94**, 3–48.
124. Hoffmann, R. and Woodward, R.B. (1968) *Acc. Chem. Res*., **1**, 17–22.
125. Woodward, R.B. and Hoffmann, R. (1971) *The Conservation of Orbital Symmetry*, Verlag Chemie GmbH, Weinheim.
126. Potts, A.W. and Price, W.C. (1972) *Proc. Roy. Soc. Lond. A* **326**, 165–179.
127. Pauling, L. (1931) *J. Amer. Chem. Soc.,* **53**, 1367–1400.
128. Goddard, W.A., III (1967) *Phys. Rev.,* **157**, 81.
129. Coulson, C. and Fischer, I. (1949) *Phil. Mag.,* **40**, 386.
130. Hiberty, P.C. and Shaik, S. (2006) *J. Comput. Chem.,* **28**, 137–151.
131. McWeeny, R. (1992) *Methods of molecular quantum mechanics*, 2nd edn. Academic Press, San Diego.
132. Pauncz, R. (1979) *Spin Eigenfunctions: construction and use*, 1st edn. Plenum Press, New York.
133. Gerratt, J. and Lipscomb, W.N. (1968) *Proc. Natl. Acad. Sci. USA*, **59**, 332.
134. Ladner, R.C. and Goddard, W.A., III (1969) *J. Chem. Phys.,* **51**, 1073.
135. Pennotti, F., Cooper, D.L., Gerratt, J. and Raimondi, M. (1988) *J. Mol. Struct. (Theochem),* **169**, 421.
136. Cooper, D.L., Gerratt, J. and Raimondi, M. (1986) *Nature,* **323**, 699.
137. van Lenthe, J.H. and Balint-Kurti, G.G. (1980) *Chem. Phys., Lett.,* **76**, 138.
138. van Lenthe, J.H. and Balint-Kurti, G.G. (1983) *Chem. Phys., Lett.,* **78**, 5699.
139. Hiberty, P.C., Flament, J.P. and Noizet, E. (1992) *Chem. Phys., Lett.,* **189**, 259.
140. Dunning, T.H., Jr., Woon, D.E., Leiding, J. and Chen, L. (2013) *Acc. Chem. Res.,* **46**, 359–368.
141. Goddard, W.A., III, Dunning, T.H., Jr., Hunt, W.J. and Hay, P.J. (1973) *Acc. Chem. Res*., **6**, 368–376.
142. Su, P., Wu, J., Gu, J., Wu, W., Shaik, S. and Hiberty, P.C. (2011) *J. Chem. Theory, Comput.,* **7**, 121–130.
143. Song, L., Mo, Y., Zhang, Q. and Wu, W. (2005) *J. Comput. Chem*., **26**, 514.
144. Shaik, S., Maitre, P., Sini, G. and Hiberty, P.C. (1992) *J. Am. Chem. Soc.,* **114**, 7861–7866.
145. Shaik, S., Danovich, D., Wu, W. and Hiberty, P.C. (2009) *Nature Chem.,* **1**, 443–449.
146. Gordon, M.S. and Truhlar, D.G. (1987) *Theor. Chim. Acta,* **71**, 1–5.
147. Hiberty, P.C., Danovich, D., Shurki, A. and Shaik, S. (1995) *J. Am. Chem. Soc.,* **117**, 7760–7768.
148. Kutzelnigg, W. (1990) In: Maksic, Z.B. (ed.) *Theoretical models of chemical bonding*, Vol. 2, p. 1. Springer, Berlin.
149. Shaik, S. (1989) In: Bertan, J. and Ciszmadia, G.I. (eds), *New theoretical concepts for understanding organic reactions*, p. 165. NATO ASI Series 267, Kluwer, Dordrecht.
150. Ramos-Cordoba, E., Salvador, P. and Reiher, M. (2013) *Chem. Eur. J.,* **19**, 15267–15275.
151. Clark, A. and Davidson, E. (2001) *J. Chem. Phys.,* **115**, 7382–7392.
152. Xu, L.T. and Dunning, T.H., Jr. (2014*) J. Chem. Theory Comput.,* **10**, 195–201.
153. Van Vleck, J.H. and Sherman, A. (1935) *Rev. Mod. Phys.,* **7**, 167–228.
154. Danovich, D., Hiberty, P.C., Wu, W., Rzepa, H.S. and Shaik, S. (2014) *Chem. Eur. J.,* **20**, 6220–6232.