

Informational system for the solution of molecular spectroscopy problems. 4. Transitions in molecules of C_{2v} and C_s symmetry

N.A. Lavrent'ev, A.I. Privezentsev, and A.Z. Fazliev

*V.E. Zuev Institute of Atmospheric Optics,
Siberian Branch of the Russian Academy of Sciences, Tomsk*

Received August 5, 2008

The part of a molecular spectroscopy informational system is described, which is designed on the basis of applied ontology of tasks. The description is presented of informational models of two tasks (direct and inverse) of finding the transitions used for the building of the informational system. Databases related to these tasks have been populated for the water molecule and its isotopomers.

Introduction

Data on transitions between energy levels in an isolated molecule are key in solution of direct and inverse tasks of molecular spectroscopy. The frequency, characterizing the transition, is the most in-demand characteristic in allied subject fields, studying radiation fluxes, on the one hand, and the basis for obtaining fundamental molecular characteristics, i.e., energy levels, on the other hand. Identified energy levels, obtained from the processing of experimental spectra, underlie the validation of assigning quantum numbers and the construction of potential energy surface. In our opinion, the latter is one of important engineering problems of modern molecular spectroscopy. Solutions, obtained in both direct and inverse tasks, require, first of all, formal restriction test,¹ imposed by mathematical models.

The openness of transition data sources and possibility of data validation at any stage of work with them are important. For this purpose, we have designed the software, which allows collecting data on transitions, testing them for restrictions following from mathematical models, storing the data in a relativistic database, providing them on user request as files of different formats, in tabular or graphical form.

The features of the designed information system (IS) enable a user to download a data array in IS and assign it to one of molecules, paper-based or online (URI assignment) publishing, supplying these data with a set of semantic metadata, which are automatically generated when downloading data in the described IS. These metadata allow solving two problems: the semantic search of information resources, presenting in the IS, and the systematization of information resources, downloaded by users in the system using a public knowledgebase.

Physical characteristics of the direct and inverse tasks on finding transition frequencies and the inverse task on retrieving a set of frequencies from

experimental data are given in Section 1, as well as the structure of metadata for molecules of C_{2v} and C_s symmetries. The procedure of data downloading by a user is discussed in Section 2. It is described in more detail in Ref. 2. Applications for search by a limited set of physical characteristics, data delivery to a user, comparison of data from different information sources, and their representation in tabular and graphical forms are described in Section 3. This description is devoted to the W@DIS (Water internet-accessible Distributed Information System) IS,³ oriented on work with data and metadata, related to the water molecule and its isotopomers.

1. Transition data and metadata

Within the chosen information model of molecular spectroscopy,⁴ transitions between molecule levels are determined in two problems of molecular spectroscopy: when solving the direct problem by the effective Hamiltonian or variational method and when solving the inverse problem by fitting the prepared list of transitions, retrieved from processing of measurement results, to a certain computation, which is reference for a given molecule.

The solution of the direct task determines vacuum transition frequencies, related quantum numbers, and Einstein coefficients. Solution of the inverse task yields vacuum transition frequencies, corrections to them caused by the measurement accuracy, and quantum numbers characterizing the transition. Assignment of quantum numbers to the transition frequencies, calculated from measurement results, is an estimating process, but not a formalized one; therefore, the results of such estimation are often contradicting. An example of such divergence of results is given below. Facts of the divergence reveal themselves when calculating rms deviations of data for couples of information sources. The Einstein coefficients usually are not found when solving the inverse task. This is caused by significant errors (up

to hundreds of percents) in experimental determination of intensities.

The value of total angular momentum can vary only by unity in traditional molecular spectroscopy, because only single-photon processes are considered in the absorption and emission of light by molecules:

$$J'' \rightarrow (J' \text{ or } J''). \quad (1)$$

For molecules of C_{2v} and C_s symmetries, the restriction

$$k_a + k_c = J \text{ or } J + 1 \quad (2)$$

is to fulfill for each energy level.

In transition from the lower to upper level, the restriction on rotational quantum numbers k_c

$$|k'_c - k''_c| = 2n - 1, \quad n = 1, 2, 3, \dots \quad (3)$$

is to fulfill.

For molecules of C_{2v} symmetry (H_2O , D_2O , $H_2^{17}O$, $H_2^{18}O$, $D_2^{17}O$, $D_2^{18}O$), the restriction on vibrational and rotational quantum numbers is to fulfill:

$$v''_3 + v'_3 + k''_a + k'_a = 2n - 1, \quad (4)$$

connected with a prohibition of transitions between energy levels of *para*- and *ortho*-water.

At present, the W@DIS IS database is the most populated with the data on transitions in the water molecule among all known IS. The data were collected by the IUPAC group,⁵ providing us with the data for IS populating. Data on transitions in the water molecule and its isotopomers in the W@DIS IS are collected from more than 250 sources.

The W@DIS IS calculation data were taken from Refs. 6–9. Frequency ranges, covered by them, are shown in Table.

Molecule	Frequency range, cm^{-1}	Reference
H_2O	0–30000	6, 7
$H_2^{17}O$	0.452854–25400.16924	7
$H_2^{18}O$	0.189153–25399.985226	7
HDO	0.02752–19966.75785	8
D_2O	0.306822–15980.26341	9

Solutions of inverse tasks include spectral fragments, which have been completely observed in Ref. 5.

Along with values of physical parameters, the W@DIS IS includes semantic metadata, forming a knowledgebase, which contains sets of statements on individuals, describing data sources, and class taxonomy.

The general approach to construction of metadata for description of the above task solutions is partly described in Ref. 10. When choosing metadata to describe problems of finding vacuum frequencies of transition in a molecule, extreme values of some physical parameters and the number of

errors made by the authors¹⁰ with respect to different restrictions, constrained by the used mathematical model of a molecule, were used as quantitative metadata.

Five levels of description were applied for data sources, used in downloading data on transitions:

1. *Description of data source attributes*: the title of the data source, the publication reference, the comment, the task type and the name of the substance, energy levels of which relate to the data source.

2. *Description of data intensionals presented in the data source*: vacuum transition frequencies, Einstein coefficients, and uncertainties in determining these parameters, quantum numbers of transitions (in notation of Normal Modes and Barber and Tennyson BT2 notation).

3. *Metadata on the task concerned with the data source* are determined in ontologies of spectroscopy problems and contain the information on the input data and the method of the task solution.

4. *Quantitative metadata on parameters describing transitions*: minimum and maximum of vacuum frequency in the set of data, incoming in the information source, the number of transitions, the presence of errors, the presence of Einstein coefficients in the information source, the number of transitions with unique quantum numbers, the number of transitions with partial quantum numbers.

5. *Quantitative metadata* characterizing errors in quantum numbers, used in transition identification (see restrictions defined by Eqs. (1)–(4)).

Figure 1 shows the screenshot containing all types of metadata provided for the user, working with data sources on transitions in the W@DIS IS. The presented metadata include hyperlinks to tabular forms of corresponding metadata. For example, the symbol set [T] points to jump to the tabular form.

Branch on such link in the line “A number of transitions with incorrect quantum numbers ($k_a + k_c \neq J$ or $J + 1$)» results in the jump to a page containing a line with an incorrect set of quantum numbers in the table. If the number of incorrect assignments is not large, then they are enumerated in the comment to the information source.

2. Downloading of transition data

To collect the data, a downloading system with web-interface has been worked out. The downloading of transitions consists of four steps.

1. Choice of the information source.
2. Preset of the molecule and the number of entities, values of which are contained in the downloaded file.
3. Description of positions of the values of entities chosen for the downloading (Fig. 2).
4. Preset of the data file address on the user side.

To download data in the IS, the user should build the information source, enclosing the reference

Annotation (1995_PaHo_D2O of 2009-02-18 17:58:17 by vss) Calculation/Experiment

Substance		Output data	
D ₂ O		Wavenumbers	
Method		Unit	cm ⁻¹
UNDEFINED		Wavenumber _{min}	110.586802
Reference		Wavenumber _{max}	413.794939
R. Paso and V.-M. Horneman, High-resolution rotational absorption spectra of H ₂ ¹⁶ O, HD ¹⁶ O, and D ₂ ¹⁶ O between 110 and 500 cm ⁻¹ . // Journal of Optical Society of America, B, 1995, v. 12, no. 10, p. 1813-1838.		Number of transitions	215 [T]
Pure rotational spectra of 16O isotopic species of H ₂ O, HDO, and D ₂ O were measured with a high-resolution Fourier-transform spectrometer in order to obtain better calibration line positions for high-resolution spectroscopy. The number of assigned lines for HDO was ,380 and for H ₂ O it was ~310 between 170 and 500 cm ⁻¹ . For D ₂ O, ~220 lines in the 110-420-cm ⁻¹ region were assigned. The calibration was performed with the highly accurate ν ₂ band of carbonyl sulfide.		Vacuum wavenumber	true
Incorrect assignment 0 0 0 13 3 12 0 0 0 13 2 11 161.037344		Uncertainty	false
Root-mean square deviations by quantum numbers		Einstein coefficient	
Type:	Number of information source [3]	Unit	s ⁻¹
NormalModes		Availability	false
		Uncertainty	false
		Transition quantum numbers	
		Quantum numbers type	NormalModes
		J _{min}	3
		J _{max}	19
		Number of transitions with unique quantum numbers	215 [T]
		Number of transitions with nonunique quantum numbers	0 [T]
		Number of unassigned transitions	0 [T]
		Number of allowed transitions for water (k _a +k _c = J or J+1)	214 [T]
		Number of forbidden transitions for water (k _a +k _c ≠ J or J+1)	1 [T]
		Number of allowed transitions (J' --> J'' or J''+ -1)	215 [T]
		Number of forbidden transitions (J' -/-> J'' or J''+ -1)	0 [T]
		Number of allowed transitions for water (k _c ' - k _c '' = 2n+1, n=1,2,3,..)	215 [T]
		Number of forbidden transitions for water (k _c ' - k _c '' = 2n)	0 [T]
		Number of allowed transitions for water (C _{2v}) (ν ₃ ' + k _a ' + ν ₃ '' + k _a '' = 2n+1)	215 [T]
		Number of forbidden transitions for water (C _{2v}) (ν ₃ ' + k _a ' + ν ₃ '' + k _a '' = 2n)	0 [T]
		Number of transitions rejected by experts	0 [T]
		Number of bands	
		ν ₁ ^{up} ν ₂ ^{up} ν ₃ ^{up} ν ₁ ^{low} ν ₂ ^{low} ν ₃ ^{low}	1 [T]

Fig. 1. Metadata of the information source, containing D₂O transitions. Of note among the metadata are transitions with incorrect quantum numbers.

to the publication. At the next step, it is necessary to indicate the molecule, to which the data array relates. The system allows recognition of the following user errors in user's file composing:

- mismatch of data types used in the file;
- incorrect structure of physical entities, described by the user in preparation for downloading;

– mismatch with the permissible value range of the physical parameters, and so on.

Figure 3 shows the database scheme presenting relations between tables, where solutions of the tasks T2 and T5 are stored. A simplification concerns the list of used quantum number notations and relations with the bibliography database.

Line number	Physical Quantity	First Position of the Column	Last Position of the Column
1	Transition frequency (Vacuum frequency)	1	13
2	Transition quantum numbers (Normal Modes)	14	91
3	Einstein coefficient	92	100
4	Uncertainty of vacuum wavenumber	101	110
5	Uncertainty of Einstein coefficient	111	121
6	Transition quantum numbers (BT2)	122	140

Ok

Current parameters

Information source	2008_FuCsTe_H2O
Substance	H ₂ O
Data schema	Transition frequency (Vacuum frequency) (1,13) Transition quantum numbers (Normal Modes) (14,91) Einstein coefficient (92,100) Uncertainty of vacuum wavenumber (101,110) Uncertainty of Einstein coefficient (111,121) Transition quantum numbers (BT2) (122,140)

Fig. 2. Description of columns of the file with output data of the task T5.

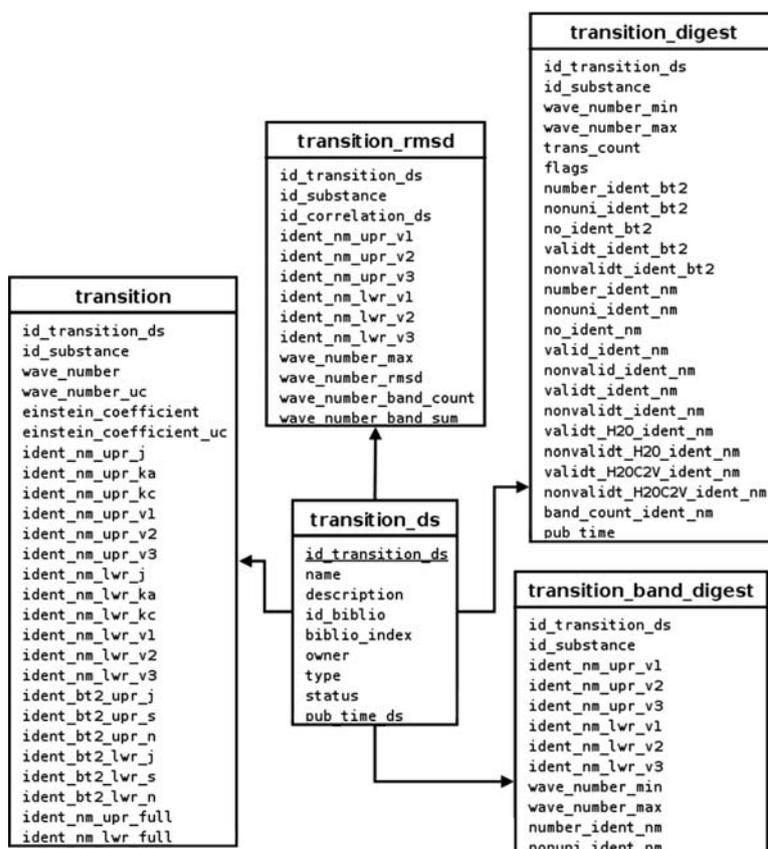


Fig. 3. The structure of database storing the data on transitions of molecules of C_{2v} and C_s symmetries.

3. Search, representation, delivery, and comparison of vacuum frequencies in the W@DIS IS

Vacuum transition frequencies, connected with the information sources, concerned the direct and inverse tasks, can be compared both with frequency value and quantum numbers in the Normal Modes or BT2 notations. Such comparison is important in the analysis of experimental data, the number of which is one-order-of-magnitude less than the calculation data.

The user can export data on vacuum transition frequencies of molecules when looking up or comparing information sources, using their tabular form. It is possible to export the whole data array or its part, preliminary chosen in the table. The data are delivered via Internet in the text format or the format used in Microsoft Excel.

Before the comparison, the user should find the required information sources. There are two search mechanisms in the W@DIS IS:

– attribute search by values of such attributes as the molecule name, the total angular momentum range, quantum number notations, etc.;

– semantic search of information sources in the knowledgebase, generated in the form of explicit specification of molecular spectroscopy, using the OWL language.¹¹

In the attribute search, the user enumerates values of attributes of physical entities or ranges of their variation, using the static interface, fixed by developers of the IS. After the necessary number of information sources is chosen, transition frequencies in tabular or graphical form can be compared in pairs or data from a certain information source can be scanned. For a couple of information sources, rms deviations can be calculated both in lines and throughout the data array. Figure 4 shows the example of such comparison. It is seen that the results can differ significantly.

To organize the search, allowing unambiguous finding of information sources, answering the user's enquiry, it is necessary to create a logic-based specification language of molecular spectroscopy, on the one hand, and to provide the IS user with facilities for constructing the enquiry for the information-source search. Some details of the search realization are given in Ref. 12.

Information source left							Information source right							
2005_MiMeStTa_D2O							2006_ZoOvShPo_D2O							
S.N. Mikhailenko, G.Ch. Mellau, E.N. Starikova, S.A. Tashkun, V.I.G. Tyuterev, Analysis of the first triad of interacting states (020), (100), and (001) of D ₂ ¹⁶ O from hot emission spectra. // Journal of Molecular Spectroscopy, 2005, v. 233, p. 32-59.							N.F.Zobov, R.I.Ovsannikov, S.V.Shirin, O.L.Polyansky, J.Tennyson, A.Jankac and P.F. Bernath, Infrared emission spectrum of hot D ₂ O. // Journal of Molecular Spectroscopy, 2006, v. 240, no. 1, p. 112-119.							
Number of lines with the same quantum numbers							w ₁ -w ₂ _{max} (cm ⁻¹)							
1969							72.14407							
<input type="button" value="Show"/> <input type="text" value="5"/> rows starting from <input type="text" value="0"/> In all rows 26910 <input type="button" value=">"/> <input type="button" value=">>"/> <input type="button" value="Control"/>														
Vacuum wavenumber w ₁ (cm ⁻¹)	v ₁ ^{up}	v ₂ ^{up}	v ₃ ^{up}	j ^{up}	K _a ^{up}	K _c ^{up}	v ₁ ^{low}	v ₂ ^{low}	v ₃ ^{low}	j ^{low}	K _a ^{low}	K _c ^{low}	Vacuum wavenumber w ₂ (cm ⁻¹)	w ₁ -w ₂ [z-a]
2251.00707	0	2	0	17	15	3	0	0	0	18	16	2	2178.863	72.1441
3057.00792	0	2	0	16	14	2	0	0	0	15	13	3	3009.1959	47.812
2858.73997	0	2	0	16	14	2	0	0	0	16	13	3	2810.9334	47.8066
390.37693	1	0	0	15	14	2	0	2	0	14	13	1	435.90035	45.5234
2856.8959	0	2	0	15	14	2	0	0	0	15	13	3	2811.384	45.5119
<input type="button" value="Show"/> <input type="text" value="5"/> rows starting from <input type="text" value="0"/> In all rows 26910 <input type="button" value=">"/> <input type="button" value=">>"/> <input type="button" value="Control"/>														
Spectral band (v ₁ ,v ₂ ,v ₃) ^{up} - (v ₁ ,v ₂ ,v ₃) ^{low}							(Vacuum wavenumber) _{rmsd} (cm ⁻¹)							
All bands							4.078							
0 0 1 - 0 0 0							0.75							
0 0 1 - 0 1 0							0.003							
0 0 1 - 0 2 0							0.004							

Fig. 4. Comparison of vacuum frequencies and calculation of rms deviations. This example points out the transition frequency mismatch, which can be connected with incorrect assignment of frequencies to quantum numbers.

Conclusion

The information sources described in this work are associated with problems of determination of the vacuum transition frequencies, their assignment to quantum numbers, and finding the Einstein coefficients. Population of the W@DIS IS [Refs. 3,13] with designed and experimental data, related to the above problems, now allows one to consider the IS as a prototype of IS, oriented to representation of data on spectral parameters of an isolated molecule of an arbitrary symmetry, as well as knowledge about information sources, where these data are published.

Acknowledgements

The authors are grateful to the Russian Foundation for Basic Research for the financial support; professors J. Tennyson and A. Császár for discussion of the transition data structure and metadata types; Dr. N.F. Zobov for placing the calculation results of the transition parameters of water molecule isotopomers in our disposal; and Dr. A.Yu. Akhlestin for realization of the application for the rms deviation representation.

References

1. T. Furtenbacher, A.G. Császár, and J. Tennyson, *J. Mol. Spectrosc.* **245**, No. 2, 115–125 (2007).
2. A.V. Kozodoev, E.M. Kozodoeva, A.I. Privezentsev, and A.Z. Fazliev, in: *Proc. of XIIIth Baikal All-Russian Conf. "Information and Mathematical Technologies in Science and Management,"* P. 2 (ESI SB RAS, Irkutsk, 2008), pp. 156–162.
3. A.Z. Fazliev, A.G. Császár, and J. Tennyson, in: *Proc. of the 10 HITRAN Database Conference* (2008), pp. 38–39.
4. A.D. Bykov, A.Z. Fazliev, N.N. Filippov, A.V. Kozodoev, A.I. Privezentsev, L.N. Sinitsa, M.V. Tonkov, and M.Yu. Tretyakov, *Geophys. Res., Abstracts* **9**, 01906 (2007).
5. J. Tennyson, P.F. Bernath, L.R. Brown, A. Campargue, M.R. Carleer, A.G. Császár, R.R. Gamache, J.T. Hodges, A. Jenouvrier, O.V. Naumenko, O.L. Polyansky, L.S. Rothman, R.A. Toth, A.C. Vandaele, N. Zobov, L. Daumont, A.Z. Fazliev, T. Furtenbacher, S.N. Mikhailenko, S.V. Shirin, and B.A. Voronin, *J. Quant. Spectrosc. and Radiat. Transfer* (in press).
6. R.J. Barber, J. Tennyson, G.J. Harris, and R.N. Tolchenov, *Mon. Notic. Roy. Astron. Soc.* **368**, No. 3, 1087–1094 (2006).
7. N.F. Zobov, R.I. Ovsyannikov, S.V. Shirin, and O.L. Polyansky, *Opt. and Spectrosc.* **102**, No. 3, 348–353 (2007).
8. D.W. Schwenke and H. Partridge, *J. Chem. Phys.* **113**, No. 16, 6592–6597 (2000).
9. S.V. Shirin, N.F. Zobov, and O.L. Polyansky, *J. Quant. Spectrosc. and Radiat. Transfer* **109**, No. 4, 549–558 (2008).
10. A.V. Kozodoev, A.I. Privezentsev, and A.Z. Fazliev, *Atmos. Oceanic Opt.* **20**, No. 9, 736–740 (2007).
11. Patel-Schneider, Hayes, Horrocks, eds., *OWL Web Ontology Language Semantics and Abstract Syntax*, W3C Recommendation 10 Feb 2004, <http://www.w3.org/TR/owl-semantics/>
12. A.I. Privezentsev and A.Z. Fazliev, in: *Proc. of XIIIth Baikal All-Russian Conf. "Information and mathematical technologies in science and management,"* P. 2 (ESI SB RAS, Irkutsk, 2008), pp. 171–176.
13. W@DIS Information System, <http://wadis.saga.iao.ru>