Vibrational Absorption Intensities in Chemical Analysis. 9. The Near-Infrared Spectra of Methyl Branched Alkanes

Malcolm E. Parker[†] **and Derek Steele***

Department of Chemistry, Royal Holloway, University of London, Egham, Surrey TW20 0EX, U.K.

Martin J. C. Smith[‡]

British Petroleum International plc, Sunbury Research Centre, Chertsey Rd., Sunbury-on-Thames,
Middlesex TW16 7LN, U.K.

Received: March 5, 1997; In Final Form: May 28, 1997[⊗]

The spectra of 43 aliphatic hydrocarbons have been measured between 3900 and 6000 cm^{-1}. It is shown that they can be represented as the sum of a limited set of group spectra. The required number of groups is established using principal factor analysis. A theorem is derived by which these group spectra can be determined directly from the eigenspectra (scores) and eigenvectors of the covariance matrix of the spectral data (loadings). Group spectra, integrated over 4 cm^{-1} intervals, are presented for *n*-alkylmethylene and methyl, isopropyl, *tert*-butyl, 2,3-dimethylpropyl, and midchain methyl, ethyl and *m*,*m*-dimethyl. Using these the average error was 2.2% for the 35 compounds in the set for which the chain branchings were separated by more than one methylene unit. This corresponds to an average predictive error of 0.0004 molar absorbance units (mol dm^{-3} cm). For CH_2 chains with isopropyl units (six spectra) the average error is 2.5%. These figures were comparable both with the deviations in the total integrated intensities as a function of the number of CH_2 groups and also with the impurity levels (98−99% pure according to suppliers and to independent glc checks). We therefore consider that the additivity for remote groupings is good at the selected 4 cm^{-1} resolution. Remote groups are defined as groups which are separated by two or more methylenes. Evidence is presented that the spectra of the functional groups are sensitive as to whether there is a $-CH_2-$ or a $-CH_3$ group present in the β position. For this reason the spectra of functional groups are extended to include corrections for β CH_3 groups. A rationalization of the effect of a β methyl or, in general, of a group substituted at the β position is given in terms of spectral perturbations of the CH_2 adjacent to a methyl as compared to a CH_2 group in the middle of a methylene chain.The principal features of the group spectra are interpreted. A major feature common to all terminal functional groups is explained in terms of changes in spectra of a midchain methylene and of a methylene adjacent to a methyl or to another grouping. Evidence is presented for a reassignment of methyl bands in the 5900 cm^{-1} region, based largely on a comparison of intensities in the spectra of *n*-alkylmethyl, isopropyl, and *tert*-butyl groups.

1. Introduction

The correlation and prediction of chemical and physical properties of mixtures by near-infrared (NIR) spectra is of significant practical importance. Advantages over alternative methods in terms of speed, precision, and suitability for on-line monitoring via optic fibers has led to widespread applications in food, agriculture, pharmaceutical, oil, and chemical industries.[1] In practically all such applications, spectra−property correlations are established—almost invariably via linear multivariate methods (e.g., partial least squares)—using a training set consisting of the spectra of a series of mixtures in which the property has been determined by an appropriate reference method. Subsequently the property is predicted from the NIR spectrum without recourse to the reference method. Many, indeed most, applications involve complex mixtures composed of hundreds of compounds. For example, there have been a number of demonstrations of the ability to predict, with high precision, the octane numbers of gasoline samples from their NIR spectra.[2,3]

The existence of strong spectra−property correlations for such complex mixtures implies a relation between the property and the concentration of some basic structural units. For example, a positive correlation of octane number with content of methyl groups (and hence degree of branching) and with aromatic hydrocarbon content has been noted.[2] Furthermore, the fact that linear correlations are obtained implies that the near-infrared spectra are linearly additive in some structural units.

Given linear additivity, the determination of functional group spectra becomes a meaningful objective. In the present work, near-infrared group spectra are derived for the basic structural units present in aliphatic hydrocarbons from an analysis of the spectra of a series of pure compounds. If principal factor analysis (PFA), or principal component analysis, (PCA) is applied to the spectra of an homologous series—or, in general, of molecules of related structure—the number of factors necessary to reconstruct the spectra to a given level of precision can be evaluated. We recognize that significant variations in the spectra arising from a functional group occur as a result of the sensitivity of Fermi resonance interactions to the skeletal conformation and, to a smaller extent, as a result of environmental perturbations.[4] It follows that the precision required is more subjective than defined by instrumental signal-to-noise ratios. The quality of fit varies according to the groups being studied—as will be discussed below. A crude estimate of the

---

* Corresponding author. E-mail: D.Steele@rhbnc.ac.uk.

[†] Present address: Department of Chemistry, University of Surrey, Guildford, Surrey, U.K.

[‡] Current address: Unicam Chromatography, P.O. Box 205, York St., Cambridge, CB1 2SS, U.K.

[⊗] Abstract published in *Advance ACS Abstracts,* November 1, 1997.

S1089-5639(97)00834-7 CCC: $14.00

TABLE 1: Compounds for Which the Near-Infrared Spectra Were Run and the Percentage Errors in Prediction Using Group Spectra

	compound	integ int	% error	group spectra	figure
1	*n*-hexane	8.82	3.5	2A + 4B	1A
2	*n*-heptane	10.35	2.6	2A + 5B	
3	*n*-octane	11.46	2.7	2A + 6B	
4	*n*-nonane	12.41	1.3	2A + 7B	
5	*n*-decane	13.56	2.1	2A + 8B	
6	*n*-dodecane	15.87	1.1	2A + 10B	
7	*n*-hexadecane	20.60	0.9	2A + 14B	
8	2-methylpentane	8.94	6.2	A + 2B + C	
9	2-methylhexane	10.11	2.1	A + 3B + C	
10	2-methylheptane	11.37	1.2	A + 4B + C	1B
11	2-methyloctane	12.59	1.4	A + 5B + C	
12	2-methylhexadecane	21.90	1.9	A + 13B + C	
13	2-methylheptadecane	23.41	1.7	A + 14B + C	
14	2-methyloctadecane	24.42	1.8	A + 15B + C	
15	2,5-dimethylehxane	11.53	3.8	2B + C	
16	2,6-dimethylheptane	12.58	2.1	3B + C	
17	2,7-dimethyloctane	13.46	4.2	4B + C	
18	3-methylheptane	11.25	1.6	2A + 4B + F + J	
19	4-methylheptane	11.38	1.6	2A + 4B + F	1D
20	3-methyloctane	12.46	1.5	2A + 5B + F	
21	4-methyloctane	14.92	0.1	2A + 5B + F	
22	2,4-dimethylpentane	11.66	5.1	2C + B	
23	2,4-dimethylhexane	11.08	5.7	A + 2B + C+ F + I + J	
24	2,4-dimethylheptane	12.83	5.8	A + 3B + C + F + I	
25	3,5-dimethylhetpane	12.39	6.9	2A + 3B + 2F + I + 2J	
26	2,3-dimethylpentane	10.01	5.2	A + B + H	
27	2,3-dimethylhexane	11.49	(0)	A + 2B + H	
28	2,3-dimethylheptane	12.33	4.1	A + 3B + H	
29	3,3-dimethylpentane	10.02	1.7	2A + 2B + H + 2K	
30	3,3-dimethylhexane	11.27	0.9	2A + 3B + H + K	
31	4,4-dimethylheptane	12.63	0.0	2A + 4B + G	
32	3-ethylhexane	11.50	2.2	2A + 3B + E	
33	3-ethylheptane	12.51	2.4	2A + 4B + E	
34	4-ethylheptane	12.41	2.5	2A + 4B + E	
35	2,5-dimethylhetpane	12.59	2.8	A + 3B + C + F + J	
36	pristane	24.31	5.4	2C + 9B + 2F	
37	squalane	37.69	6.8	2C + 16B + 4F	
38	2,2-dimethylhexane	10.97	2.5	A + 3B + D	1C
39	2,2-dimethylbutane	8.57	12.4	A + B + D	
40	2,2-dimethylpentane	9.80	3.8	A + 2B + D	
41	2,2-dimethyloctane	13.28	0.0	A + 5B +D	
42	2,2,5-trimethylhexane	12.10	5.0	2B + C + D	
43a	3-methyltridecane	18.54	3.3	2A + 10B + F	
43b	3-methyltridecane	18.54	1.6	2A + 10B + F + J	
A	*n*-methyl	2.191			2A
B	*n*-methylene	1.156			2B
C	isopropyl	4.54			2C
D	*t*-butyl	5.308			2D
E	br-ethyl	3.512			2G
F	br-methyl	2.337			2E
G	*n*,*n*-dimethyl, $-C(CH_3)_2$	3.626			2I
H	2,3-dimethylpropyl, $(CH_3)_2HCCCH_3H$	−6.991			2F
I	β-brMe/brMe	−0.148			
J	β-Me/brMe	0.040			3
K	β subst correction to br-*m*,*m*-dimethyl	−1.326			

dimensionality is given by the Malinowski's indicator function (IND).[5] The problem that is then faced is how to relate the factors to structural entities. As the spectral features in the near-infrared are predominantly related to hydrogen motions, one option is to relate to various types of hydrogen deformations. Thus for *n*-hydrocarbons it would be necessary to include methyl C−H stretch, methylene CH stretch, two types of HCH angle deformation, and two types of HCC as a minimum. This would seriously overdefine the problem. For *n*-alkanes, as discussed below, only two factors are necessary. This suggests that the entities might be considered as CH_3 and CH_2. It will transpire that, though we have progressed with this hypothesis, there is a significant oversimplification involved which has serious effects when chain branching is considered.

When 2-methylalkanes are added to the data set, one extra factor is introduced. This might have been taken as the methine hydrogen, but this would have required that the methyl vibrations were essentially unaltered when two methyls are attached to the same carbon. It is well-known however that, for example, the symmetric HCH umbrella vibration is split into two bands with approximately 10 cm^{-1} separation. This is a direct consequence of kinematic coupling through the common carbon atom, though Bellamy has quoted evidence in favor of a through space interaction as responsible for the splitting.[6] It was obviously better to treat the isopropyl unit as the "*new factor*". Following along lines such as the above, an attempt was made in the first instance to interpret near-IR spectra in terms of the units: *n*-alkane CH_3, defined as CH_3 attached to $(CH_2)_n$, $n \geq 2$; *n*-alkane CH_2, defined as CH_2 attached only to CH_2 or CH_3; isopropyl; *m*,*m*-dimethyl ($-C(CH_3)_2-$); branched methyl ($-CHCH_3-$); branched ethyl ($-CH(CH_2CH_3)-$); and *tert*-butyl.

In section 3 we shall present a theorem allowing the direct extraction of group spectra from the results of factor analysis of the data set.

In the Discussion, section 4.1, a synopsis of earlier relevant work on band assignments in the near-infrared will be followed by evidence for a reassignment of the fundamental antisymmetric ν_{C-H} of a methylene adjacent to a methyl. This will be seen to be of considerable importance when seeking to understand the complexity of spectra of groups containing methyls. Following this, the determination of group spectra of *n*-alkane type methyl with those of terminal isopropyl and *tert*-butyls will give strong evidence for a spectral reassignment of the highest wavenumber features of the C−H first overtone.

Along the way attention will be drawn to certain problems for which we have as yet no answer.

2. Experimental Procedure

Spectra were recorded at a resolution of 1 cm^{-1} on a Perkin-Elmer 1760X spectrometer with cell lengths of 0.5, 2.0, and 10.0 cm as appropriate. Cells were of the Infrasil variety with very low absorbances throughout the near-IR. The resulting spectra were integrated over 4 cm^{-1} intervals. This procedure is employed to mitigate perturbations of the band centers and shapes of vibrational bands by effects such as Fermi resonance and intermolecular interactions. This follows the logic and pattern employed in earlier papers in this series, but here we employ a smaller interval than the 25 cm^{-1} previously used.[4]

Samples were generally identified as 98 to 99% pure. This was checked by gas chromatography (GC), and in most cases the nature of the principal impurities determined using FTIR/GC. Some samples were discarded on the basis of containing carbonyl or amino impurities which would be likely to generate significant bands in the near-ir or because they were of less than 97% purity. The spectra were converted into molar absorbance spectra employing tabulated densities to evaluate concentrations. The compounds studied are listed in Table 1, in which are given the total integrated intensities over the spectral range 6000−3888 cm^{-1} and the figure numbers corresponding to the spectra. It is always a problem to identify with any precision where the true background lies, even when

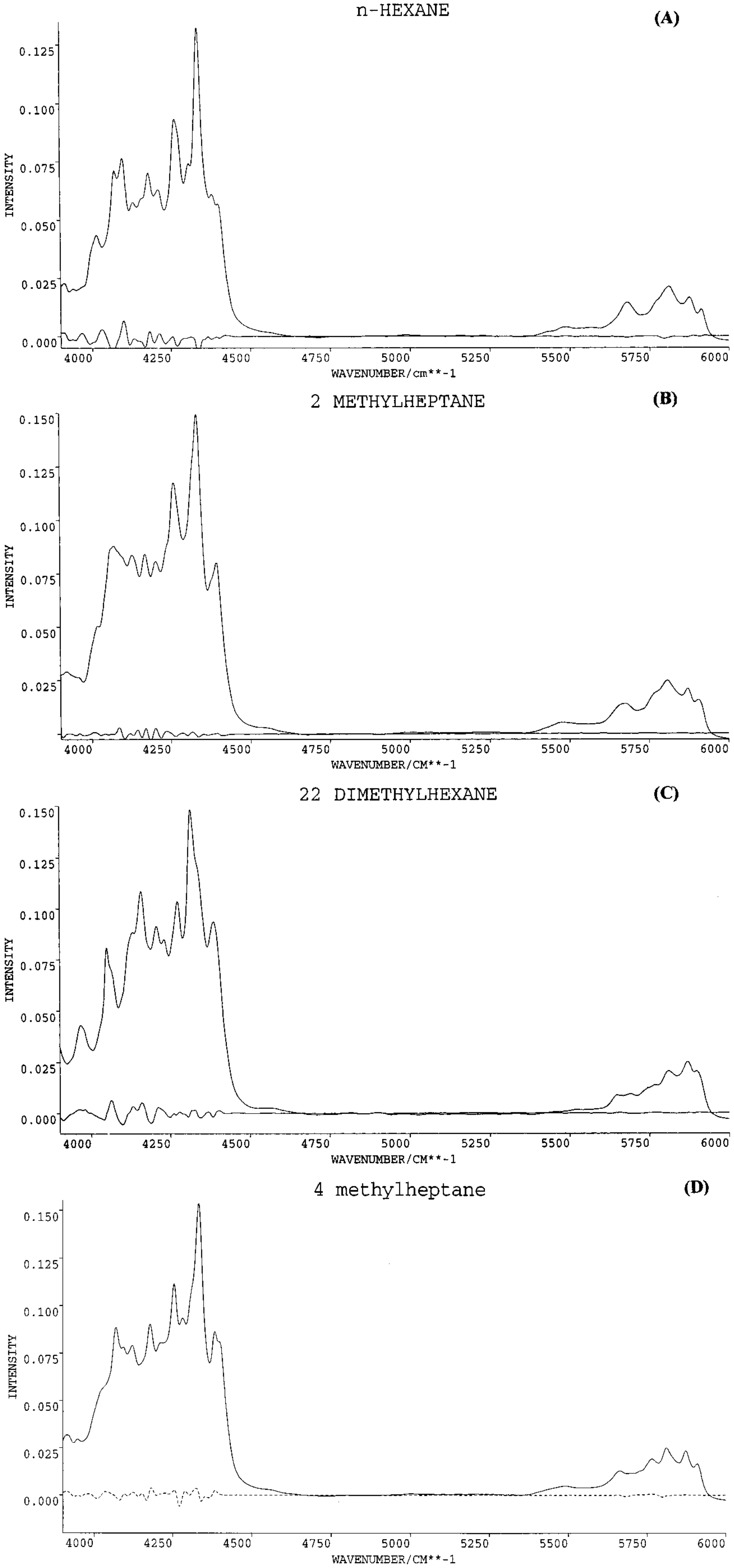


Figure 1. The observed NIR spectra between 3900 and 6000 cm^{-1} (solid lines) and the errors in prediction using the group spectra shown in Figure 2 (broken lines) of (A) *n*-hexane, (B) 2-methylheptane, (C) 2,2-dimethylhexane, (D) 4-methylheptane.

this is a smooth function of the abscissa. We opted to position it along the flat spectral region 4900 to 4700 cm^{-1}. This looks eminently reasonable until it is observed that this procedure consistently leads to negative absorption at 6000 cm^{-1}, just

TABLE 2: Designation of CH Stretching Frequencies and Their Approximate Fundamental Wavenumbers (Values Taken from Ref 10)

symbol	description	approx wavenumber/cm^{-1}
d^-	methylene antisym stretch	2915
d^+	methylene sym stretch	2846
d^+_{FR}	metylene sym stretch in Fermi resonance with combination of deformations	~2890
d^+_{ω}	methylene sym stretch for terminal CH_2	2853
d^-_{ω}	methylene antisym stretch for terminal CH_2	~2936† [a]
r^-_a	methyl CH sym stretch, in-plane bond dominant	2962
r^+	methyl CH sym stretch out-of-plane bonds dominant	2870
r^-_b	methyl CH antisym stretch	2952
r_{FR}	methyl CH sym stretch, Fermi resonance with 2δ	2931

[a] † denotes a new assignment. Note that for the methyl modes, designated r, only C_s symmetry is assumed.

above the first overtones of the C−H stretching bands (see Figure 1). As no artificial factor could be found to explain this negative region it has to be assumed that there is a general background of absorption in the 4900−4700 cm^{-1} region. Indeed the absorption is reasonably proportional to the number of methylene units. Despite this observation it was decided to retain the original assumption on the basis that it was a simple matter to achieve consistency in choice of the background. The principal effect is to underestimate significantly the total absorbance arising from the methylene entity.

All data were processed with software written by the authors. The only unusual program was one written to extract group spectra from eigenspectra of the factor covariance matrix.[7] This is discussed in more detail in the next section.

As the complexity of the carbon skeleton increases, so does the need for factors beyond the number afforded by the above structures. In part it will be seen that this is due to neglect of the influence of a given group (e.g., n-alkane−CH_3) on neighboring units (e.g., n-alkane CH_2). Interaction through two carbon atoms (β-interactions) is the next level of complexity introduced. Due to the limited data set currently available, only two β interactions of significance can be determined. They are the interaction of two methyls (midchain type) separated by a methylene (β-brMe/brMe) and of a midchain methyl branch β to a terminal methyl group (β-Me/brMe).

3. Theory of a Method To Extract Group Spectra from Eigenspectra of the Data Covariance Matrix

We shall define the eigenspectra, **R**, (or scores in traditional nomenclature) by

$$\underset{(n\text{freq}\times m)}{\mathbf{R}} = \underset{(n\text{freq}\times n)}{\mathbf{D}}\ \underset{(n\times m)}{\mathbf{Q}_m} \qquad (1)$$

where **Q** is the normalized eigenvector (or loadings) matrix of the data covariance matrix **Z** and $\mathbf{Q}_m$ is that matrix generated by retaining only the m most important eigenvectors. Since **Q** is orthonormal,

$$\mathbf{D} \approx \mathbf{D}'_m = \mathbf{R}\mathbf{Q}^{t}_{m} \qquad (2)$$

we see that **D** may be reformed from the m eigenspectra, the reconstituted data matrix being $\mathbf{D}'_m$. Of course these eigenspectra are not the spectra of the components but are related by a linear transformation. For the purposes of evaluation of group spectra from an homologous series of molecules, a procedure based on constraining the eigenvector components to relate to known concentrations has been developed[9] and used to generate the spectra of methylene and methyl groups of n-hydrocarbons in the range 3600−1300 cm^{-1}. An alternative and more direct method is presented here.

Let the matrix of the group spectra be designated **Σ** and the matrix of group concentrations be **P** (defined here as a matrix of row vectors). Then, by the assumption of the additivity of the group spectra

$$\underset{(n\text{freq}\times n)}{\mathbf{D}} = \mathbf{R}\mathbf{Q}^{t} \cong \underset{(n\text{freq}\times m)}{\mathbf{\Sigma}}\ \underset{(m\times n)}{\mathbf{P}^{t}} \qquad (3)$$

As in standard factor analysis procedures, property matrices can be related to eigenvectors using statistical procedures. We define **B** as the property (concentration) correlation matrix. c is the number of species whose concentrations are being defined.

$$\underset{(n\times m)}{\mathbf{Q}} = \underset{(n\times c)}{\mathbf{P}}\ \underset{(c\times m)}{\mathbf{B}} \qquad (4)$$

Substitution of the transpose of (4) into (3) yields

$$\mathbf{R}\mathbf{B}^{t}\,\mathbf{P}^{t} = \mathbf{\Sigma}\mathbf{P}^{t} \qquad (5)$$

and hence

$$\mathbf{\Sigma} = \mathbf{R}\mathbf{B}^{t}$$

B is unique only for $c = m$. For $c < m$, it can be determined from **Q** and **P** by multiple linear regression. Thus

$$\mathbf{B} = (\mathbf{P}^{t}\mathbf{P})^{-1}\,\mathbf{P}^{t}\mathbf{Q} \qquad (6)$$

However it is the case of $c = m$ that is of concern here. Equation (4) still applies. Since principal factor analysis informs us that the original data set can be generated from m linearly additive spectra, it remains to determine the appropriate concentration matrix, **P**. In the present context the spectra from which **D** is constituted are of pure molecular compounds. The concentrations may therefore be defined in terms of atomic groupings. It will be demonstrated later how the appropriateness of the choice of groupings can be checked. When PFA was applied to n-alkane spectra ranging from n-hexane to hexadecane only two factors were required. It is natural for a chemist to deduce that these factors arise from the additivity of the methylene and methyl group spectra. This, as we shall see, is an oversimplified picture.

4. Discussion

In the following we have adopted the notation of ref 10 for group vibrations, though in later discussions it will be convenient to change to accepted local mode designations when discussing a particular group. This is summarized in Table 2.

Within the range 6000−3900 cm^{-1} there are two major groups of absorbances. The first overtones and combination region of the CH stretches occur within the region 6000−4000 cm^{-1}. Not surprisingly, the similarity of spectra of different hydrocarbons is strong, with the principal variations arising from the CH_3:CH_2 ratio. Nevertheless subtle differences due to structural variations in the skeleton do occur. Some insight into these variations can be obtained from a consideration of the literature on *n*-alkanes. Ricard-Lespade, Longhi, and Abbate[11] examined the 6000−3700 cm^{-1} absorption of a series of *n*-alkanes (C_4, C_5, C_7, and C_{36}) and of a solid polymethene. Attention was focused on the methylene contribution, and a model was developed using up to fourth-order in the potential expansion and a constrained dipole expansion. Fermi resonance was taken directly into consideration. The principle conclusions were that the absorption at 5664 cm^{-1} is due to the pure overtones of the symmetric ($2d^+$) and of the antisymmetric ($2d^-$) CH_2 stretches, shifted down by Fermi resonance of $2d^+$ with the ternary combination of d^+ with skeletal and deformation modes. One major problem with this analysis is the neglect of the differences in spectral wavenumbers for the midchain methylenes and those adjacent to the methyls. Strictly d^+ and d^- refer to the in-phase motions of the chain $d^+(\pi)$ and $d^-(\pi)$. The Raman active bands occur at substantially lower wavenumbers [see, for example, ref 12). Another disconcerting feature of the analysis of Ricard-Lespade et al. is that the intrinsic intensity of the ternary combination was required to be 0.72 of that of the overtones, whereas experience has shown that intensity generally drops by 1 order of magnitude for each extra quantum involved.

An alternative analysis has been presented[13,14] based on a local mode model using harmonically coupled anharmonic oscillators. This model was applied to propane,[15] acetone, and dimethyl ether.[16] Both frequencies and intensities of the principal features are explained. Of special note is that the importance of the combination features, such as $|1,0\rangle|1\rangle$, of the methyl become increasingly important for lower quanta transitions relative to the overtones. In the notation $|n_a,n_b\rangle\,|n_c\rangle$ the n_a and n_b refer to the two equivalent oscillators out of the local skeletal plane and n_c refers to the CH oscillator in the plane. The combination transitions are of low probability for the higher quanta manifolds, but for the two-quanta manifold they dominate.

Before discussing our results it is useful to summarize the results established for the fundamental region. It has been shown that the CH in the plane of the adjacent C−C of an *n*-alkane is vibrationally distinct from the other two. These are associated with vibrational bands at 2950 and 2936 cm^{-1}, respectively, in the case of *n*-propane-d_7.[17] Ab initio calculated bond lengths of the C−H bonds in a variety of alkane structures are shown to correlate to a high precision (±0.0001 Å) with the observed wavenumbers of the isolated C−H stretches as observed in monohydro perdeuterated alkanes.[18] The ab initio calculated bond lengths also correlate extremely well with the higher overtone local mode wavenumbers.[19] Such studies show a sensitivity of the C−H stretching wavenumbers to conformation. Remarkably the observed wavenumbers cluster in groups separated by about 13 cm^{-1} (0.0010 Å). The C−H of the methylene adjacent to methyl (ω-CH_2) have vibrational wavenumbers which are raised 13 cm^{-1} for each adjacent methyl, assuming trans configurations of the skeleton. Thus for *n*-propane, *t*-*n*-butane, and *tt*-*n*-pentane the isolated vibrations occur at 2920, 2907, and 2906/2892 cm^{-1}, respectively (in effect at the mean of the d^+ and d^- wavenumbers).

Further evidence of the conformational dependence of CH stretching wavenumbers comes from cyclohexane.[20] The equatorial $\tilde{\nu}_{CH}$ has been identified 31 cm^{-1} higher than that of the axial. Calculations show that less than 1 cm^{-1} is due to kinematic effects. This is a specific case of a widely recognized pattern that a CH trans to a carbon is stronger than a CH bond trans to a hydrogen.[21] Such effects can lead to changes from CH_2 symmetric and antisymmetric mode to effectively local mode behavior. Thus for 1,3,5-trioxane it has been shown that bands at 3030 and 2852 cm^{-1} are due 100% to equatorial and axial CH stretching, respectively.[22]

In the methyl symmetric deformation (umbrella mode) the carbon atom vibrates along the C−C line with substantial amplitude, balancing the momentum of the three hydrogens. This results in strong coupling of the umbrella mode with C−C stretching and with the corresponding vibration of other methyls attached to the same carbon. The well-known splitting of the band in isopropyl and tertiary butyl vibrations[23] arises in this way.

4.1. Assignment of the C−H Stretching Absorptions of a Methylene Adjacent to a Methyl. The existence of distinct ω-CH_2 absorbances was first recognized by Hill and Levin.[24] They located the symmetric CH_2 stretch of the ω-methylene in the Raman and infrared spectra at 2855 cm^{-1} in *n*-hexadecane. This wavenumber should be compared with 2846 cm^{-1} for the equivalent motion of the other CH_2 groups.

The corresponding antisymmetric frequency was tentatively located[24] near 2900 cm^{-1} (2917 cm^{-1} for midchain). This, however, appears to be too low. The infrared spectra of crystalline *n*-alkanes have been measured at $T < 10$ K.[10] The symmetric CH_2 mode, d^+_ω of the terminal CH_2, was clearly seen at 2853 cm^{-1}. Nonterminal −CHD− gave absorption at 2886 cm^{-1},which is within 5 cm^{-1} of the average of the symmetric and antisymmetric wavenumbers (2846 and 2915, average 2881 cm^{-1}). At least part of this small discrepancy is due to Fermi resonance with the combination band, $d^+_{FR,}$ at 2890 cm^{-1}. Visually this band appears to be about 5% of the absorption at 2853 cm^{-1}. Assuming that all the intensity comes from the fundamental leads to a perturbation of 5% of the 44 cm^{-1} separation, $d^+_{FR}-d^+$. This in turn results in 3 cm^{-1} perturbation of the mean of the CH_2 modes. Assuming that the terminal (ω) CHD at 2897 cm^{-1} is again at the mean of the symmetric ω-CH_2 stretch, d^+_ω, (at 2853 cm^{-1}) and the antisymmetric stretch, d^-_ω, indicates that the latter is at 2941 cm^{-1}. This band is masked by the methyl bands. The terminal methyl thus produces 7 and 26 cm^{-1} shifts on the adjacent symmetric and antisymmetric methylene bands. The average shift of 16.5 cm^{-1} can be compared with the shift observed between midchain and ω-CHD of d_{11} *n*-pentane of 14 cm^{-1}. This suggests that the value for the antisymmetric mode, as estimated above, may be too high by about 5 cm^{-1}. Of course Fermi resonances can influence these wavenumbers as this is not taken into account in deriving the antisymmetric wavenumber from the d_{11} pentane isolated $\tilde{\nu}^{iso}{}_{CH}$.

Data on the first overtone spectrum of $CD_3(CH_2)_nCD_3$[11] may also be interpreted to show that the absorption due to CH_2 adjacent to the methyl is at a significantly higher wavenumber than that for midchain methylenes. There is a pair of CH_2 bands for each *n*-alkane. These two bands have been assigned to Fermi resonance of the CH overtone with $\tilde{\nu}_{CH} + 2\delta$ where δ is the HCH bending wavenumber. If we take the intensity ratios of the two bands as recorded by Ricard- Lespade et al. for d_6 *n*-alkanes and weight the wavenumbers according to these intensities (assumes that the tertiary band has a negligible intrinsic intensity in comparison with the binary band) then the CH_2 overtone wavenumbers, corrected for this perturbation, are found to be the following: butane, 5783.5 cm^{-1}; pentane, 5740.5 cm^{-1}; heptane, 5729.8 cm^{-1}; hexatriacontane, 5729.9 cm^{-1}.

The complexity of the bands is actually much greater than the two principal bands, and these wavenumbers must be given at least a 3 cm^{-1} uncertainty. However this again shows a strong perturbation of the CH_2 adjacent to the methyl (26.8 cm^{-1}). This figure is to be compared with twice the difference in the d CH stretch fundamentals of about 2 × 14 cm^{-1}. For *n*-pentane and *n*-heptane there is no apparent resolution into separate overtone bands due to the α- and β-methylenes.

4.2. *n*-Alkanes. The factor method showed that the near-IR spectra of the seven *n*-alkanes were well-reproduced by two factors. Correlating these with the number of methylenes and methyls led to group spectra[4] that reproduced the molecular spectra to an average error of 1.9% (see Figure 1A for representative spectral fit). It has been noted[24] that the ω-methylenes, adjacent to the methyls, absorb at different band centers from the main chain. However, in the absence of species deuterated at either the ω-CH_2 or the methyl, there is no way that the contributions of the ω-CH_2 and the midchain CH_2 can be separated on the basis of the longer chain hydrocarbons. It follows that the CH_3 spectrum includes a contribution equal to the difference between the spectra of the ω-CH_2 and the midchain CH_2. This could have resulted in negative group absorbances at wavenumbers which were more characteristic of the midchain CH_2 than of the ω-CH_2. In practice, all the group absorbances of CH_2 and of CH_3 are positive, though negative values have been noted for interacting methyl groups (see discussion of spectra of branched methyl groups below) and for the midchain *m,m*-dimethyl group, $-C(CH_3)_2-$. As can be seen in Figure 2A, the methyl spectrum shows a strong sharp dip with a minimum at 5780 cm^{-1}, close to the very strong CH_2 overtone that peaks at 5784 cm^{-1}. There is also a clear shoulder at 5680 cm^{-1} near the maximum for the CH_2 at 5674 cm^{-1}. It is believed that this is a negative-going peak on the side of the CH_3 band. As we do not yet have data on species such as CD_3 which would allow separation of the ω-CH_2 component, it makes sense to report the spectrum of the terminal ethyl unit by adding the CH_3 and CH_2 spectra (Figure 2H). A 5780 cm^{-1} minimum is now almost absent. The principal features, which will be discussed below, occur at 5912, 5876, 5812, 5682, and 5478 cm^{-1}.

It is of value to recognize the origin of the bands, so that the effects of structural changes can be estimated. Let ω and $\tilde{\nu}$ be the harmonic and experimental wavenumbers, respectively. In a local mode study of high-overtone CH stretching modes of various hydrocarbons, Wong and Moore[25] showed that the anharmonicity is given by

$$\omega x/\mathrm{cm}^{-1} = 123.9 - 0.0207\tilde{\nu}^{\mathrm{iso}}{}_{\mathrm{CH}} \qquad (7)$$

where $\tilde{\nu}^{\mathrm{iso}}{}_{\mathrm{CH}}$ is the CH stretching wavenumber isolated from CH coupling with other CH bonds by otherwise complete deuteration.[26] Applying this to the normal CH_2 mode wavenumbers, d^+ and d^-, taking $\tilde{\nu}^{\mathrm{iso}}{}_{\mathrm{CH}}$ as 2886 cm^{-1}, and then using the relation

$$\tilde{\nu}_{0\rightarrow 2} = 2\tilde{\nu}_{0\rightarrow 1} - 2\omega x$$

the overtones of $2d^+$ and $2d^-$ are predicted at 5701 and 5564 cm^{-1}. The agreement with the observed bands at 5784 and 5674 cm^{-1} is unacceptable. An alternative analysis can be based on the local mode model.[13] In this case the expected bands are due to the degenerate {0,2} transitions of the CH stretching modes and the {1,1} mode. The separation should be $2\omega x$, that is, about 126 cm^{-1}, with the lower overtone band at $2\tilde{\nu}^{\mathrm{iso}}{}_{\mathrm{CH}} - 2\omega x$ (~5646 cm^{-1}). Note that for the $|0,0\rangle \rightarrow |1,1\rangle$ transition the observed frequency should be simply twice that of the fundamental. Allowing for an incomplete transfer from normal to local mode behavior and for the Fermi resonances with such as the ternary combination of the CH stretch with 2 quanta of the HCH scissoring motion (fundamental at about 1465 cm^{-1}), the accord with the observed separation of 110 cm^{-1} and of the lower band center near 5674 cm^{-1} indicates that the first overtone spectra are much closer to the local mode model. Kjaergaard et al.[15] solved the vibrational Hamiltonian for propane allowing for intra- and intermanifold coupling using an harmonic coupling restriction. The intermolecular coupling parameter γ_{m}, related in the case of methylene to twice the separation of the symmetric and antisymmetric fundamental frequencies, has a substantial effect on the ground state to $|02\rangle_+$ and $|11\rangle$ state frequencies. These frequencies are given by

$$\tilde{\nu} = 2\omega_{\mathrm{m}} - 5\omega_{\mathrm{m}}x_{\mathrm{m}} \pm [(\omega_{\mathrm{m}}x_{\mathrm{m}})^2 + 4(\gamma'_{\mathrm{m}})^2]^{1/2}$$

where the + sign refers to the $|11\rangle$ transition. Taking Bellamy's values for the average CH_2 symmetric and antisymmetric wavenumbers of 2926 and 2853 cm^{-1}, the γ'_{m} value is 36.5 cm^{-1}. Such a large coupling moves the predicted wavenumbers to 5940 and 5756 cm^{-1}. This is clearly not in accord with observation, even allowing for the fact that the γ'_{m} may have been estimated too large as indicated by comparison with the value of 23.4 cm^{-1} for propane. It would seem that on a polymethylene chain there is a more rapid movement to local mode behavior. Certainly in propane, theory and experiment are in very good accord.

Finally, Ricard-Lespade et al. also identified a very weak background absorption in the region 5400−5600 cm^{-1} which they identified as the ternary methylene combination ($d + \omega + \delta$). This correlates with the third feature apparent in the methylene overtone spectrum in the region 5508−5572 cm^{-1} and is thus assigned to the $|10\rangle$ ($\omega + \delta$) transition.

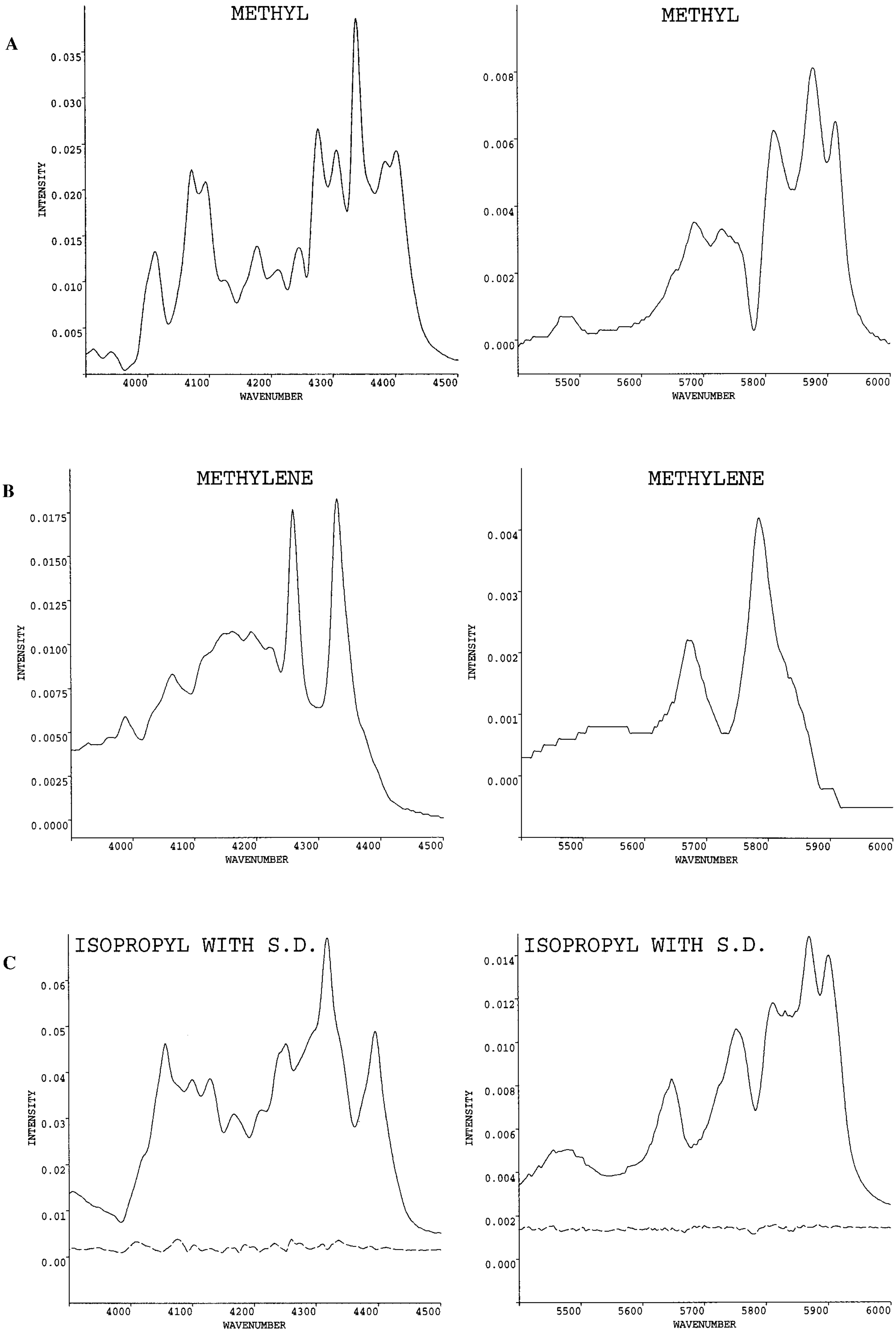
The methylene combination region is somewhat more complex. The spectra are much less well predicted in terms of a limited set of factors indicating the more important role of nonlocalized modes. The most important intense features are common to all spectra however, being two narrow bandwidth peaks at 4332 and 4260 cm^{-1}. Ricard-Lespade et al.[11] calculated that the methylene ($d^- + \delta$) and ($d^+ + \delta$) combination modes absorb at 4330 and 4250 cm^{-1}, respectively, the symmetric combination resonating with a (3δ) overtone observed as a shoulder at 4340 cm^{-1}. These combination peaks are thus assigned to the $|10\rangle_-\delta$ and $|10\rangle_+\delta$ transitions, although once again the Fermi peak is unresolved in the methylene component spectrum. The remaining features below 4250 cm^{-1} are thought to be binary combinations of methylene stretches with wagging and twisting deformations.

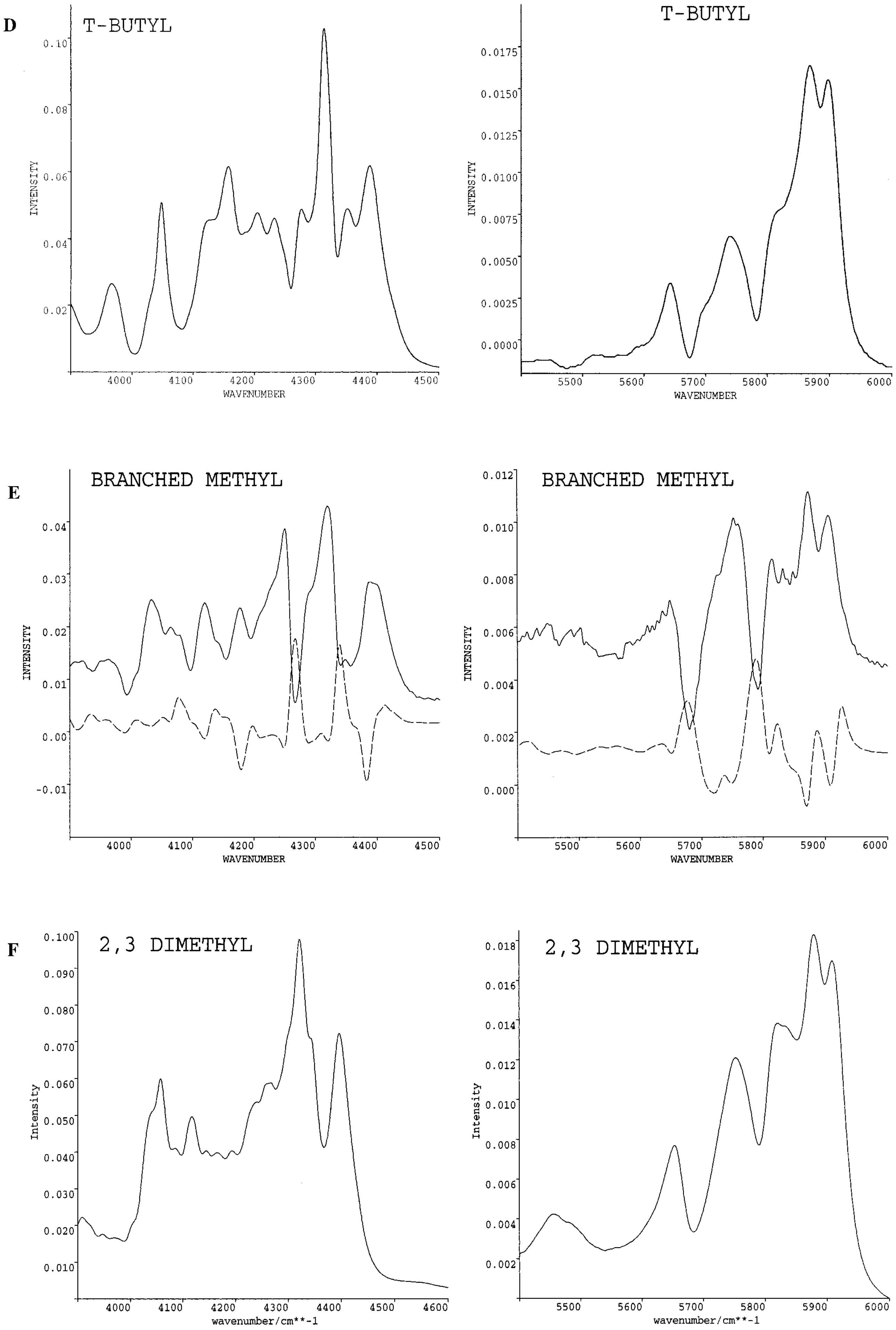
Moving attention to the CH_3 modes and using an ωx of 63 cm^{-1} the overtones of the modes r_a^-, r_b^-, r^+, and $r^+{}_{FR}$ leads to predicted wavenumbers of 5798, 5778, 5614, and 5736 cm^{-1}, respectively, in this pure normal mode picture. This is, again, in poor accord with experiment. In the local mode view we require the wavenumbers appropriate to the CH bond oscillators in and out of the skeletal plane. Assuming (a) that the methyl is perfectly tetrahedral, and (b) that CH/CH interaction force constants are the same irrespective of the orientation of the CH bonds, then the normal mode wavenumbers can be related to the force field by

$$\lambda_1 = 4\pi^2\tilde{\nu}_1^2 = \left(\mu_{\mathrm{H}} + \frac{\mu_{\mathrm{C}}}{3}\right)\left(\frac{f^{\mathrm{i}} + 2f^{\mathrm{o}}}{3} + 2f_{\mathrm{CH/CH}}\right) \qquad (8)$$

$$\lambda_2 = 4\pi^2\tilde{\nu}_2^2 = \left(\mu_{\mathrm{H}} + \frac{4\mu_{\mathrm{C}}}{3}\right)\left(\frac{2f^{\mathrm{i}} + f^{\mathrm{o}}}{3} + f_{\mathrm{CH/CH}}\right)$$

$$\lambda_3 = 4\pi^2\tilde{\nu}_3^2 = \left(\mu_{\mathrm{H}} + \frac{4\mu_{\mathrm{C}}}{3}\right)(f^{\mathrm{o}} - f_{\mathrm{CH/CH}})$$


A
METHYL
INTENSITY
0.035
0.030
0.025
0.020
0.015
0.010
0.005
4000
4100
4200
4300
4400
4500
WAVENUMBER
METHYL
0.008
0.006
0.004
0.002
0.000
5500
5600
5700
5800
5900
6000
WAVENUMBER
B
METHYLENE
INTENSITY
0.0175
0.0150
0.0125
0.0100
0.0075
0.0050
0.0025
0.0000
4000
4100
4200
4300
4400
4500
WAVENUMBER
METHYLENE
INTENSITY
0.004
0.003
0.002
0.001
0.000
5500
5600
5700
5800
5900
6000
WAVENUMBER
C
ISOPROPYL WITH S.D.
INTENSITY
0.06
0.05
0.04
0.03
0.02
0.01
0.00
4000
4100
4200
4300
4400
4500
WAVENUMBER
ISOPROPYL WITH S.D.
INTENSITY
0.014
0.012
0.010
0.008
0.006
0.004
0.002
0.000
5500
5600
5700
5800
5900
6000
WAVENUMBER


D
T-BUTYL
INTENSITY
0.10
0.08
0.06
0.04
0.02
4000
4100
4200
4300
4400
4500
WAVENUMBER
T-BUTYL
INTENSITY
0.0175
0.0150
0.0125
0.0100
0.0075
0.0050
0.0025
0.0000
5500
5600
5700
5800
5900
6000
WAVENUMBER
E
BRANCHED METHYL
INTENSITY
0.04
0.03
0.02
0.01
0.00
-0.01
4000
4100
4200
4300
4400
4500
WAVENUMBER
BRANCHED METHYL
INTENSITY
0.012
0.010
0.008
0.006
0.004
0.002
0.000
5500
5600
5700
5800
5900
6000
WAVENUMBER
F
2,3 DIMETHYL
Intensity
0.100
0.090
0.080
0.070
0.060
0.050
0.040
0.030
0.020
0.010
4000
4100
4200
4300
4400
4500
4600
wavenumber/cm**-1
2,3 DIMETHYL
Intensity
0.018
0.016
0.014
0.012
0.010
0.008
0.006
0.004
0.002
5500
5600
5700
5800
5900
6000
wavenumber/cm**-1

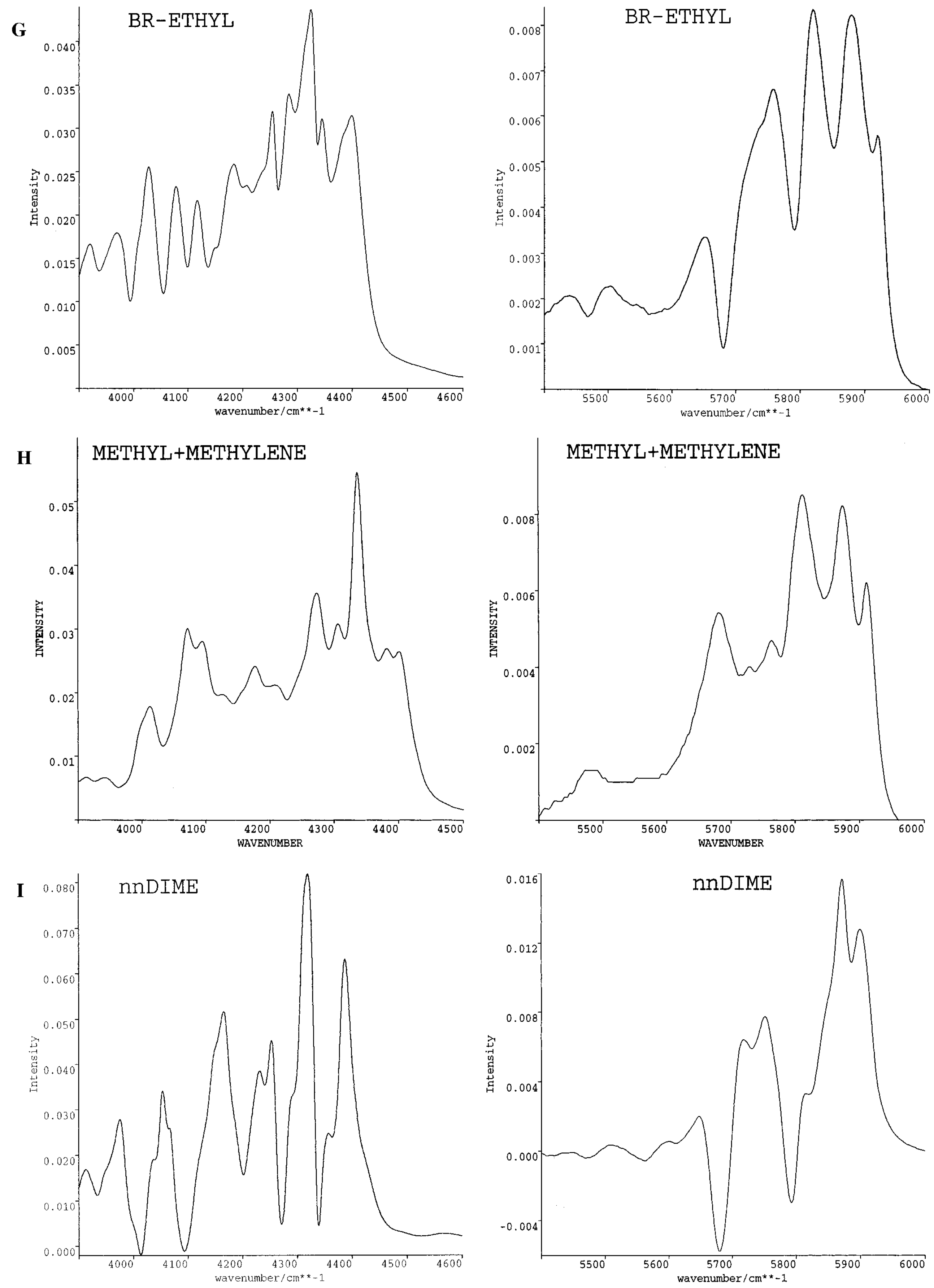


Figure 2. The NIR spectra of some hydrocarbon groups between 3900 and 4500 cm^{-1} and between 5400 and 6000 cm^{-1}. Absorption between these regions is negligible: (A) methyl attached to *n*-alkyl chain, CH_3-; (B) methylene in long alkyl chain; (C) isopropyl, $(CH_3)_2CH-$ (broken line, standard deviation); (D) *tert*-butyl; (E) methyl branch (CH_2 either side), $-CHCH_3$; (F) 2,3-dimethylpropyl, $(CH_3)_2CHCH(CH_3)-$; (G) branched ethyl, $-CH(C_2H_5)-$; (H) terminal ethyl (sum of A and B); (I) *m*,*m*-dimethyl, $-C(CH_3)_2-$.

The superscripts i and o designate the CH bonds in and out of the skeletal plane. Substituting the experimental fundamental vibrational wavenumbers[10] of 2962 and 2870 cm^{-1} leads to values of the force constants $f^{\,\mathrm{i}}$, $f^{\,\mathrm{o}}$, and $f_{\mathrm{CH/CH}}$ of 4.733, 4.680, and 0.029 mdynes/Å, respectively. In turn these lead to $\tilde{\nu}^{\mathrm{i}} = 2939$ cm^{-1} and $\tilde{\nu}^{\mathrm{o}} = 2923$ cm^{-1} for isolated CH bonds ($-$CHD$_2$). This separation is in very good accord with that observed using partial deuteration studies[18] where $\tilde{\nu}^{\mathrm{i}}$ and $\tilde{\nu}^{\mathrm{o}}$ were found to be 2950 and 2936 cm^{-1}. The higher wavenumbers are due to the use of vapor data. Thus the separation of the in-plane and out-of-plane transition wavenumbers for the isolated CH bonds increases from 10 cm^{-1} in the coupled CH$_3$ to 16 cm^{-1} in CHD$_2$, and the predicted combination bands for the transitions from the ground to $|02\rangle|0\rangle$ and $|00\rangle|2\rangle$ states will lie at about 5752 and 5736 cm^{-1}, respectively. The state nomenclature $|\mathrm{v}_1\mathrm{v}_2\rangle|\mathrm{v}_3\rangle$ is such that v_1 and v_2 refer to the spatially equivalent bonds, in this case those out-of-plane. The transition to the $|11\rangle|0\rangle$ state will be $2\omega x_{\mathrm{e}}$ higher than the degenerate transition, that is at 5846 cm^{-1}, and the $|01\rangle|1\rangle$ wavenumber should be 5862 cm^{-1}. As seen in Figure 2a the dominant features of the methyl group spectrum are three peaks at 5912, 5876, and 5814 cm^{-1} and a sharp dip at 5780 cm^{-1}. This dip corresponds to the midchain methylene band. As the spectra analyzed did not include any structure which permitted the separation of midchain and ω-methylene contributions, the difference in the spectra of the two types of methylenes is added to the methyl. The spectrum of *n*-butane would have been an asset here, but its volatility renders quantitative measurements difficult with the available equipment. In the ethyl spectrum (sum of the CH$_2$ and CH$_3$, Figure 2H) the negative-going feature has totally disappeared. As previously discussed, the CH stretch of an ω-CHD has a $\tilde{\nu}_{\mathrm{CH}}$ which is about 11 cm^{-1} above that of a midchain methylene. This should lead to the $|02\rangle$ transitions of the methylene midchain and ω-groups being 22 cm^{-1} apart. The 5820 cm^{-1} feature of the ethyl spectrum is actually about 35 cm^{-1} higher. There is some indirect evidence that we may have underestimated the ω-effect. Kjaergaard et al. studied the $\Delta\mathrm{v} = 2 - 6$ overtones of the CH stretches of *n*-propane. They assigned the 5851 cm^{-1} transition as arising from both transitions to the methylene $|11\rangle$ level and the methyl $|110\rangle$ level. This methylene is adjacent to two methyl groups and therefore one might expect the perturbation of the methylene to be twice that of a methylene adjacent to one methyl. On the basis of the value 5782 cm^{-1} for a midchain methylene, this would suggest that the ω-methylene should be near 5816 cm^{-1}, in excellent accord with our observations. With the satisfactory explanation of the lower wavenumber features it would appear that the two higher bands must be correlated with the predicted 5862 and 5846 cm^{-1} transitions. However the discrepancies appear rather high. Furthermore, as will be discussed below, there is additional evidence requiring a reassignment of these features.

Tosi and Pinto[27] identified a band in *n*-alkanes, the absorptivity of which varied with methyl content, but the position of which varied according to the expression

$$\tilde{\nu}/\mathrm{cm}^{-1} = 5853 + 49F_{\mathrm{CH}_3}$$

F_{CH_3} represents the mole fraction of methyl. This yields a predicted band center of 5860 cm^{-1} for hexadecane, rising to 5878 cm^{-1} for *n*-butane. The correlation of this band with the experimental value in this work of 5876 cm^{-1} can be confidently made.

Halonen and Child[28] developed the model of three Morse oscillators, harmonically coupled. Solving their secular equations for the A$_1$ and for the E states of the two-manifold, as given in their appendix, yields the following transition wavenumbers:

$$\tilde{\nu}(\mathrm{A}_1) = 2\omega' - 5\omega x + \lambda \pm [\omega^2x^2 + 9\lambda^2 + 2\omega x\lambda]^{1/2}$$

$$\tilde{\nu}(\mathrm{E}) = 2\omega' - 5\omega x - \lambda/2 \pm [\omega^2x^2 + {}^9/_4\lambda^2 - 7\omega x\lambda]^{1/2}$$

Since $3\lambda = \langle 010{:}\mathrm{A}_1| - \langle 100{:}\mathrm{E}|$ we can, as for the methylene case, determine λ from accepted values of the fundamental transition wavenumbers. Accepting again the values quoted for hydrocarbons in Bellamy, λ is about -30 cm^{-1}. For the A$_1$ overtone and combination the terms in λ largely cancel. For the E pair the terms reinforce. The predicted separation increases to about 180 cm^{-1} from 120 cm^{-1} in the absence of the coupling term. Once again the introduction of the coupling term appears to be detrimental to agreement with experiment.

4.3. 2-Methylalkanes. Factor analysis of the seven *n*-alkanes and seven 2-methyl alkane spectra showed them to be well-reproduced with four factors. 2-Methylhexane showed a relatively large Mahalinobis number, and indeed removal of this spectrum from the set reduced the number of factors to three. The isopropyl spectrum was generated including and excluding 2-methylhexane. The spectra were very similar. In fact for ease of statistical understanding, the isopropyl spectrum was deduced from the spectra of 2-methylhexane, 2-methylheptane, 2-methyloctane, 2,6-dimethylheptane, 2,7-dimethyloctane, 2,5-dimethylhexane, and 2-methylhexane by subtraction of the CH$_2$ and CH$_3$ (*n*-alkane) contributions. The average error was 2.5%. In Figure 2c an estimated uncertainty of the isopropyl spectrum is shown (Σ) which was derived by taking the standard deviation from all the residual isopropyl spectra determined as described above. The uncertainty is clearly very low, and the 2.5% figure certainly includes some zero background errors. The principal peaks are now at 5900 ($\epsilon \approx 0.012$ mol^{-1} dm^3 cm^{-1}), 5872($\epsilon \approx 0.013$), 5812 ($\epsilon \approx 0.012$), 5756 ($\epsilon \approx 0.0087$), and 5648 cm^{-1} ($\epsilon \approx 0.003$). The two highest wavenumber bands are depressed 6 to 12 cm^{-1} from their *n*-alkane counterparts, but are clearly correlated. The 5814 cm^{-1} band is much less well-defined, but still present, indicating that the isopropyl is affecting the adjacent methylene in a manner similar to methyl affecting methylene. The methine CH is difficult to detect in the fundamental region, but it is believed[12] to be close to 2890 cm^{-1}. The $\tilde{\nu}_{\mathrm{CH}}$ mean of CH$_2$ and CH$_3$ is 2880 and 2930 cm^{-1}, respectively, and so 2890 cm^{-1} is an intermediate value. There is no obvious band which can be assigned to the tertiary CH $|2\rangle$ transition. The ab initio calculations show bond lengths for the secondary and tertiary CH bond lengths in 2-methylbutane of 1.090 and 1.0885 Å. Such a bond length difference suggests that the $\tilde{\nu}^{\mathrm{iso}}$ for the tertiary CH should be about 15 cm^{-1} lower than that for methylene.

The local mode overtone bands show significant differences from the equivalent primary methyl transitions. Two distinct features are noted at 5756 and 5648 cm^{-1} in the isopropyl component spectrum. These highly resolved features replace the much less distinct 5732 and 5682 cm^{-1} primary methyl features. Once again these are assigned to the symmetrized overtones, the one with the higher wavenumber being due to transitions to the $|20\rangle_{-}|0\rangle$ level and the lower to the $|20\rangle_{+}|0\rangle$ level (see section 4.5). As with the *n*-alkylmethyl, this high degree of splitting of the symmetrized states is indicative of the degree of kinetic and potential coupling present, and it is very noticeable that this is significantly greater in isopropyl.

4.4. *tert*-Butyl. The spectra of four 2,2-dimethylalkanes are available. Extraction of the spectrum of the methylene chain contributions from each spectrum shows that the 2,2-dimethylhexane and 2,2-dimethyloctane yield similar results (±2%). The discrepancy increases with shorter chain lengths, being 9% for the 2,2-dimethylpentane (Figure 1C) and 13% for the butane. The bulk of the discrepancy for 2,2-dimethylbutane is clearly due to the ω-methylene effect. The pentane error may well indicate a significant β-methylene interaction with the *tert*-butyl. Possible procedures for dealing with short chain interactions will be described, but at present it has to be said that more work is needed here. The *tert*-butyl spectrum is shown in Figure 2D. The high wavenumber methyl bands are very similar to those in isopropyl and branched methyl, but stronger roughly in proportion to the number of methyls present. Although not surprising at first, further consideration of the relevant molecular structures reveals that this degree of similarity should not be so based on the assignments previously advanced for propane[15] (section 4.5).

The shifted methylene structure is still present with the dip at ~5784 cm^{-1}. The ω-CH$_2$ band on the high-wavenumber side is much reduced, or perhaps removed and replaced by a very broad higher wavenumber feature. Indeed we might well expect the ω-CH$_2$ band itself to be significantly different from that in *n*-alkanes. If replacement of an α-CH by a C−C produced a monotonic shift, then the new ω-feature which should appear, in addition to the constant negative dip at the midchain CH$_2$ frequency, would be a positive band on the low-frequency side of the negative band in isopropyl and *tert*-butyl. In confirmation of this expectation Tosi and Pinto identified a feature in the *tert*-butyl compounds, the absorptivity that increased with increasing methylene content. In 2,2-dimethylbutane this feature was observed at 5818 cm^{-1}. It was observed to decrease in wavenumber with increasing chain length, but with approximately constant absorptivity, thus confirming the 5820 cm^{-1} band as due to the ω-CH$_2$.

The final feature which is most noticeable on the *tert*-butyl + methylene composite spectrum is found at 5700 cm^{-1}, appearing as a shoulder on the low-energy side of the 5740 cm^{-1} *tert*-butyl band. Tosi and Pinto[27] observed a methylene feature at 5692 cm^{-1} in 2,2-dimethylhexane and at 5685 cm^{-1} in 2,2-dimethylheptane. This may well indicate that the $|20\rangle_\pm$ methylene feature decreases in wavenumber as the chain length increases, as does the $|11\rangle$ feature. If this is the case, then this shoulder may be due to the ω-methylene effect. Indeed a similar feature may be observed on the high-energy side of the isopropyl 5756 cm^{-1} band, although this is a lot less pronounced.

4.5. The Assignment of the Highest Wavenumber Features of the First Overtones of the Groups Containing Methyl Groups. Kjaergaard et al. treated the in-plane and out-of-plane CH bonds of gaseous propane by extending the two harmonically coupled identical Morse oscillators model, with which they described methylene, with the addition of a third harmonically coupled nonidentical Morse oscillator. The states are represented by $|v_1,v_2\rangle|v_3\rangle$. The calculated transitions were compared with gas-phase spectra. They observed an absorption at 5915 cm^{-1} that they attributed to a transition to the $|10\rangle_\pm|1\rangle$ state. This is in agreement with the 5912 cm^{-1} component we observe in the methyl spectrum. The second feature at 5851 cm^{-1} they assigned to the $|11\rangle|0\rangle$ transition from the ground state. This clearly correlates with the 5876 cm^{-1} feature. Using ab initio derived dipole derivatives, the oscillator strengths were calculated for transitions from the ground state to the various local mode states. The observed oscillator strengths for the $|11\rangle|0\rangle$ and $|10\rangle_\pm|1\rangle$ transitions are 6.7×10^{-8} and $11. \times 10^{-8}$, respectively. Thus they predict the lower wavenumber transition as having 0.6 times the intensity of the higher wavenumber transition. This is not surprising inasmuch as the lower energy transition is singly degenerate whereas the higher energy transition is doubly degenerate neglecting normal mode splitting. It would appear from inspection of our spectra that there is a reversal of the intensities from above. More convincing evidence for a reassignment comes from a comparison with the spectra of the isopropyl and *tert*-butyl groupings.

For the isopropyl group, the corresponding bands occur at 5900 and 5872 cm^{-1}. For the *n*-alkylmethyl, v$_1$ and v$_2$ represent excitation of the CH oscillators trans to a CH bond, whereas for an isopropyl the pairs of equivalent CH bonds on the methyl are trans to a CC bond. Thus for the isopropyl the higher wavenumber feature is predicted to be due to the $|11\rangle|0\rangle$ transition on the grounds that the CH frequency of a CH bond trans to a CC is higher than when trans to CH. Based simply on a consideration of the degeneracies of the states, the higher wavenumber feature should carry most of the intensity in the *n*-alkyl spectrum, which it does not. Conversely, in the isopropyl spectrum it should be the lower wavenumber feature which is the stronger. What is also noticeable is the reduction in the intensity ratio of the stronger to the weaker bands in going from *n*-alkylmethyl to isopropyl. In an ab initio study to be published separately, the wavenumbers of butane and 2-methyl butane were calculated using a 6-31G** basis. The wavenumber $\tilde{\nu}(\mathrm{CH^{CC}})$ of the methyl is constant in going from *n*-alkyl to isopropyl. However, the $\tilde{\nu}(\mathrm{CH^{CH}})$ for the CH bond trans to CH drops by 10 cm^{-1}. Thus there is a predicted 10 cm^{-1} drop in this $|10\rangle_\pm|1\rangle$ combination feature in going from *n*-alkyl to isopropyl. This is indeed close to the observed decrease in the high-wavenumber feature.

The most convincing evidence for a reassignment, however, comes from the *tert*-butyl grouping. Here there is only one type of CH bond, trans to CC. There might still be a case for distinguishing between those bonds trans to C−CH$_2$ and those trans to C−CH$_3$. However, in the ab initio studies there would appear to be no correlation between these distinctions and the CH bond lengths. Therefore the discussion will proceed using the generalized notation $|v_1v_2v_3\rangle$. The *tert*-butyl has approximately C_{3v} symmetry as opposed to the other two cases which have C_s symmetry. We can compare then with the analysis of CH$_3$D by Halonen and Child.[28] As discussed earlier, the two highest energy features in the propane $\Delta v = 2$ manifold occur at 5915 and at 5851 cm^{-1}. In the case of *tert*-butyl two similar features are observed at 5896 and 5868 cm^{-1}. These features cannot arise from CH bond inequivalence arising from different trans groups. However, under the local mode scheme, the $\Delta v = 2$ combination forms two distinct symmetrized local mode levels, one symmetric ($|110;A_1\rangle$) and one doubly degenerate ($|110;E\rangle$). In the case of CH$_3$D these levels are observed at 5980 and 6022 cm^{-1}. This degree of separation is indicative of a great deal of normal mode character. In the case of *tert*-butyl, the band separation is about $^2/_3$ of this, 28 cm^{-1}. In accordance with these observations, the 5896 and 5868 cm^{-1} *tert*-butyl features shall be assigned to transitions to the $|110;E\rangle$ and the $|110;A_1\rangle$ levels, the degree of splitting being due to the incomplete transition to local mode behavior in the two-quanta manifold. By implication then the two bonds in the *n*-alkylmethyl and isopropyl spectra arise from a splitting of the two levels $|10\rangle_\pm|1\rangle$. If this is correct there remains the problem of locating the $|11\rangle|0\rangle$ transition.

In the case of an *n*-alkylmethyl the $|11\rangle|0\rangle$ energy level involves the simultaneous single-quanta excitations of two CH bonds trans to CH. In the case of isopropyl this transition involves the simultaneous single-quanta excitations of two CH$^{\mathrm{CC}}$ bonds. Thus the transition to the *n*-alkyl $|11\rangle|0\rangle$ energy level is predicted to be about 18 cm^{-1} lower than the equivalent isopropyl transition. As the $|10\rangle_{\pm}|1\rangle$ is split according to this interpretation and the doublet being rather broad, it would seem reasonable to assume that the $|11\rangle|0\rangle$ transition may be buried under this feature. Moreover, in the case of the *n*-alkylmethyl, the transition to the $|11\rangle|0\rangle$ energy level will be on the low-energy side of the origin of the $|10\rangle_{\pm}|1\rangle$ transition. Conversely, in the case of isopropyl, the transition will be on the high-energy side of the $|10\rangle_{\pm}|1\rangle$ band. This could possibly explain the decrease in the intensity ratio of the two features in going from *n*-alkyl to isopropyl, the intensity contribution from the $|11\rangle|0\rangle$ transition being shifted to a higher energy in the case of the latter, hence increasing the intensity of the high-energy feature.

4.6. Branched Methyl (br-Me). This is defined as a methyl branch which is not part of an isopropyl nor attached to an isopropyl. The initial "calibration" set used consisted of 3- and 4-methyloctane and 3- and 4-methylheptane. As with all group spectra it was apparent that a distinction existed between the group spectrum for situations where a β-methyl existed as opposed to a β-methylene. The effect here is quite large, as can be seen from the differences shown in Figure 3. The spectra, whose differences are shown, are for 3- and 4-methyloctanes from which the spectral absorptions of two methyl and five methylene groups have been subtracted. The resulting trace therefore corresponds to the $-CHCH_3-$ spectrum in the environments $-(CH_2)_2-CHCH_3-(CH_2)_2-$ and $-CH_3CH_2-CHCH_3-(CH_2)_2-$. Although the calibration set is limited to only four spectra to produce the two groups, the fit is acceptably good and errors are less than 1.5% of total intensity (see, for example, Figure 1d). After this initial work, an additional five branched methyl spectra were added to the set. Although the number of factors as indicated by the indicator function was increased from three to five, the effect on the errors of the original calibration set was minimal. With four factors the prediction errors were barely increased from those with five factors, and so the group spectra for branched methyl with only β-methylenes and with one β-branching were taken as originally determined. The difference between these two spectra (Figure 3) is called a "β-correction spectrum" (β-Me/brMe), as its subtraction from the spectrum of the $-CH(CH_3)-$ grouping β to a CH_2−methyl will yield a "normal" mid chain spectrum.

The strong β-methyl spectral corrections can be interpreted in terms of the ω-CH_2 spectrum of an *n*-alkyl chain. In the CH stretch first-overtone region there are large positive absorbances in the spectrum of the β-Me/brMe at 5782 and 5676 cm^{-1} coinciding in shape and in position with negative peaks of the reference branched methyl spectrum. As previously noted these features are due to the noncoincidence of the midchain and ω-methylenes. The implicit inclusion of this contribution in the *n*-alkyl spectrum results in its being included twice when two groups are separated by only one CH_2. The existence of two ω-methylene contributions in the β-dimethyl group spectrum results in genuine negative features as opposed to the negative-going absorbances noted in previous cases. Thus, apart from any accumulative perturbation on the methylene in an XCH_2Y situation, the β-interaction terms include the methylene perturbation due to an adjacent methyl. Clearly the effect of a β-Me/brMe interaction could be used as an estimate of the difference between the spectra of the ω-CH_2 and the midchain group. To test this idea further, the "true", or corrected, CH_3 (alkane) group spectrum was evaluated by adding the β-Me/brMe correction spectrum to CH_3. The results are shown in Figure 4. Comparison shows that, at least in the CH stretch overtone region, there is a significant simplification with the removal of all strong features between 5830 and 5720 cm^{-1}. It would be possible to redefine all the relevant group spectra so that an ω-CH_2 was extracted. It is considered, however, that it would be better to wait until spectra are available, for example, for $CH_3CD_2(CH_2)_nCH_3$ from which the ω-CH_2 spectrum can be deduced directly, assuming negligible CD_2 contribution, by subtraction from the $CH_3(CH_2)_{n+1}CH_3$ spectrum.

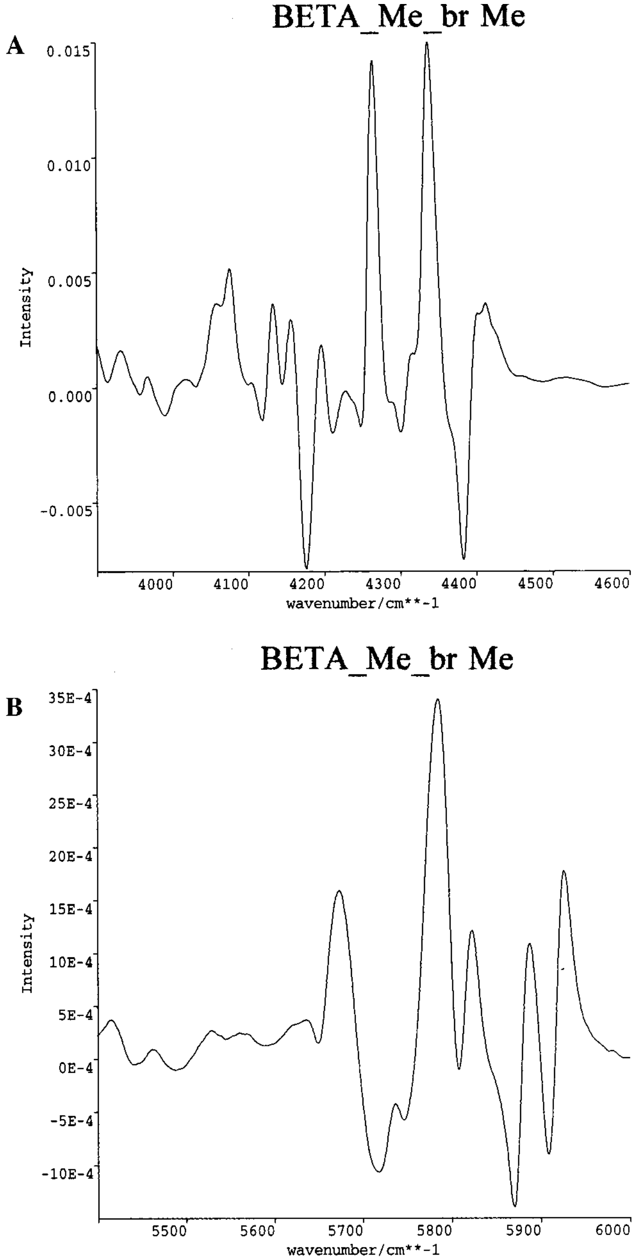


Figure 3. The β-methyl/br-methyl interaction spectrum as derived from the spectra of 3- and 4-methyloctane by subtraction of the contributions of the *n*-alkyl type methyl and methylene units.

In 3,5-dimethylheptane we have an interaction between two branched methyls (β-brMe/br/Me). The β-Me/brMe correction worsens the fit. This is not surprising since the β-Me/brMe term is correcting for a methylene being replaced by a hydrogen, but here we need to correct for a methyl replacing a hydrogen. More examples of β-dimethyl-substituted alkanes are needed before any constructive discussion can be made.

4.7. 2,3-Dimethyl. In view of the strong α-effects expected—and due to the implicit inclusion in all group spectra of the ω-methylene, less the main chain methylene contributions—it was considered advisable to include, where possible, adjacent groups as a single group. The present database is too small to include other species than 2,3-dimethyl.

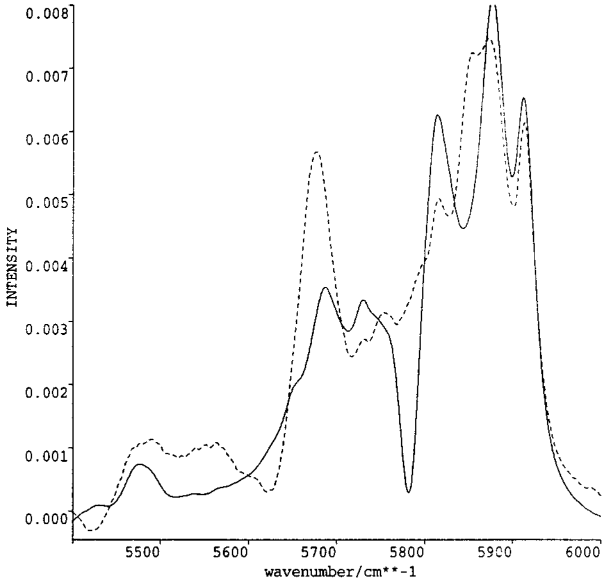


Figure 4. A comparison of the spectrum of the CH_3 group spectrum (solid line) with that resulting from the addition of the β-Me/brMe spectrum to that of the methyl group (dashed line).

Even here it is obvious that it is inappropriate to include 2,3-dimethylbutane and 2,3-dimethylpentane. This leaves only 2,3-dimethylhexane and 2,3-dimethylheptane. It was opted to take the extracted spectrum of the 2,3-dimethyl group from the hexane and predict the heptane spectrum. The error was 4% and is indeed reasonably satisfactory. For 2,3-dimethylpentane the small residual spectrum contains some similar features to the β-Me/brMe interaction spectrum. This is as expected since once again the ω-methylene effect has been included twice.

4.8. *m,m*-Dimethyl. An excellent example showing the additivity of group perturbations is shown by 4,4-dimethylheptane, 3,3-dimethylpentane, and 3,3-dimethylhexane. The spectrum of the latter is the mean of the other two (Figure 5). It is to be noted that the three curves always cross at the same wavenumbers. The presence of such isobestic points is further corroboration of the degree of additivity. At the same time, the correction to the spectrum of the $-(CH_2-CH_2)-C(CH_3)_2-(CH_2-CH_2)-$ group is substantial. The dominant features are the negative peaks 5685 and 5795 cm^{-1}. These are obviously the inverse peaks of the nonterminal methylenes as in branched methyl. Indeed the comparison is sufficiently marked that further discussion at present is unwarranted.

4.9. Branched Ethyl. Three compounds of this sort are present in the data set, these being 3-ethylhexane, 3-ethylheptane, and 4-ethylheptane. Factor analysis unambiguously determined the presence of three components which were taken as CH_3, CH_2, and $-CH(CH_2CH_3)-$. The degree of reproducibility was excellent.

5.0. Spectral Predictions

Only a limited number of spectral predictions have been made, since most of the recorded spectra were used to generate the group spectra or the spectral corrections. Nevertheless some useful indications of reliability and of limitations have resulted. The following are worthy of comment.

(1) 2,5-Dimethylheptane: The fit is very good with an error of 0.35 in 12.59. Nevertheless there are indications that a small γ-effect exists by comparison of the residual with that in pristane and squalane (see below).

(2) Squalane (2,6,10,15,19,23-hexamethyltetracosane) and pristane (2,6,10,14-tetramethylpentadecane): The predictive errors for these compounds are about 7% and 5%, respectively. Both structures consist of two isopropyl units separated by a chain of methylene and branched methyl units. There are four and two, respectively, of γ-separated branched methyls. Comparison of the residuals in Figure 6 shows a remarkable similarity with an intensity ratio close to 2:1. This is a strong indication that these residuals are due to a γ-methyl$-$methyl interaction. The principal features also seem to occur in the residual of 2,5-dimethylheptane, but with a small shift to lower wavenumbers. It appears likely that this effect arises from a preferred skeletal conformation. The existence of group spectra implicitly depends on an independence from skeletal conformation. Yet such an independence cannot be totally true. In an all trans *n*-alkane, selection rules demand spectral differences compared with structures having gauche conformations. In addition to the effects of selection rules, significant changes in normal frequencies do occur on conformational changes. An example is available from the ab initio calculations of the conformers of butane given in ref 29. A further example in a more general context is afforded by the evaluation of an intra- and intermolecular force field for *n*-alkane chains.[30] The evaluation of an intramolecular field component, independent

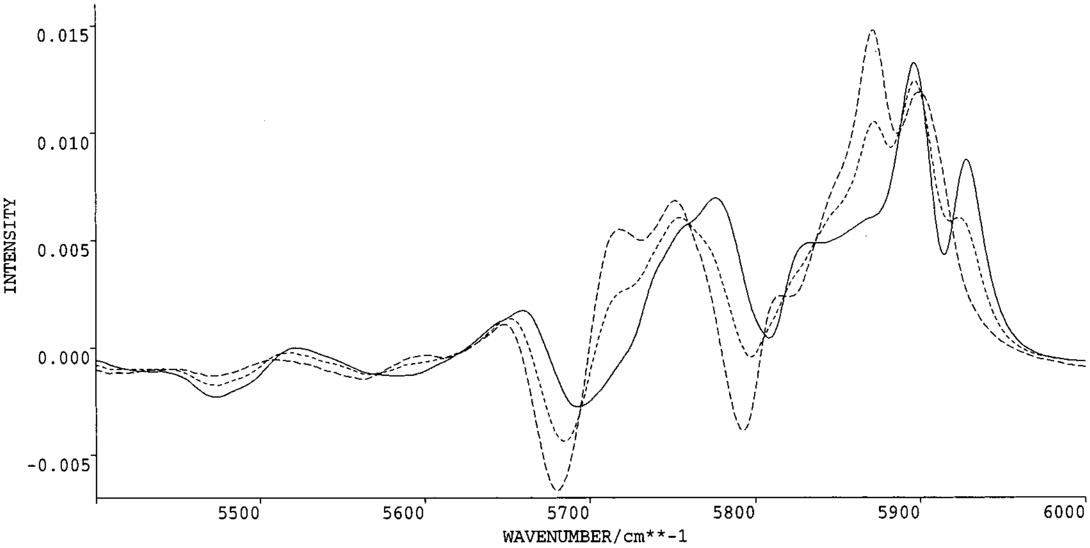


Figure 5. A comparison of the observed spectra of 4,4-dimethylheptane (long dash), 3,3-dimethylhexane (short dash), and 3,3-dimethylpentane (solid line).

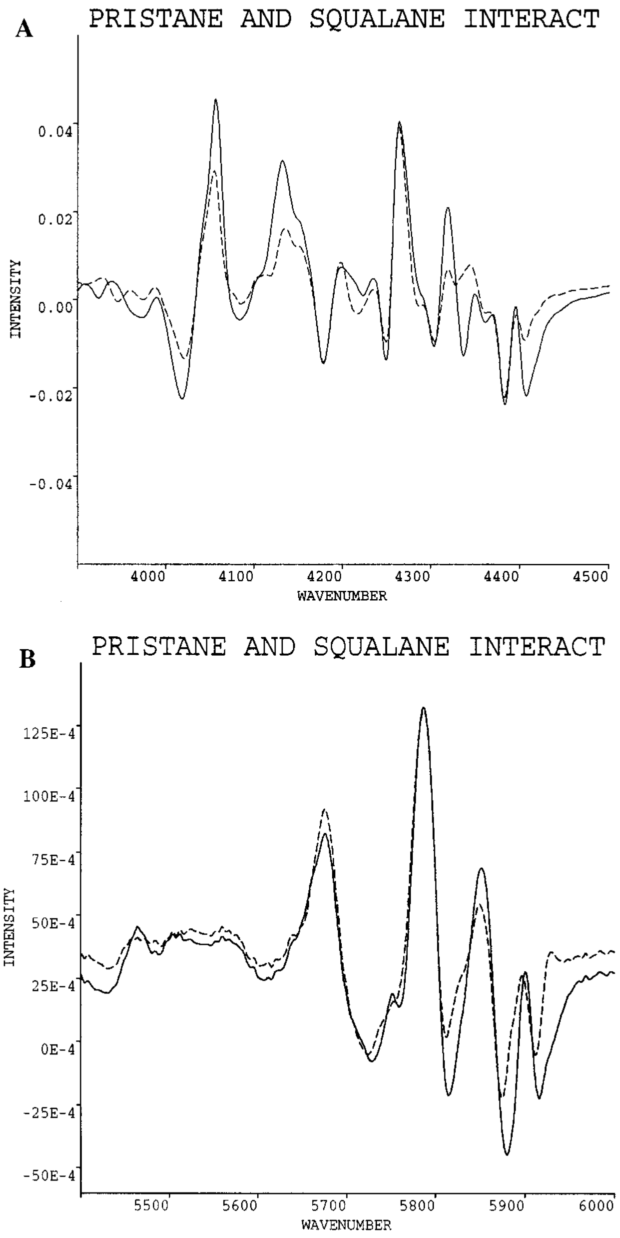


Figure 6. A comparison of the residuals in the prediction of squalane (—) and of pristane (- - -).

of conformation, results in a significant dependence of several interaction force constants on dihedral angles involving the skeletal chain. Various studies have been made of conformational dependences of spectra and of CH stretching frequencies in particular. In the latter category, for example, rules have been deduced to relate the CH stretching frequency to the types of atoms in trans and gauche configurations.[18] There is some evidence that there is an increased proportion of trans structure in liquid *n*-alkanes as compared with the vapor phase, though there is some controversy about the details.[31]

It would appear then that, although our studies have shown a surprising independence of near-IR group spectra to environment, there are some indications of long-range structural dependence which are of conformational origin.

6. Conclusion

It has been demonstrated that some features of the first-overtone spectra of alkanes have been incorrectly assigned. Local mode behavior is dominant, though certain features demonstrate a residual symmetrization. Much more important, however, is the fact that it has been demonstrated that very useful NIR group spectra in the first-overtone/combination band region can be deduced for hydrocarbons. The absorbances are additive to a good approximation. Without this additivity NIR applications would not work. The different perturbative effects of adjacent methylenes and methyls must be allowed for. Studies of such as partially deuterated hydrocarbons are required to clarify these interactions.

References and Notes

(1) Cowe, I. A.; McNicol, J. W. *Appl. Spectrosc.* **1985**, *39*, 257.

(2) Kelly, J. J; Barlow, C. H.; Jinguji, T. M.; Callis, J. B. *Anal. Chem.* **1989**, *91*, 313.

(3) (a) Espinosa, A.; Sanchez, M.; Osta, S.; Boniface, C.; Gil, J.; Martens, A.; Descales, B.; Lambert, D.; Valeur, M. *Oil Gas J.* **1994**, *92*, 49. (b) Asker, N.; Kokot, S. *Appl. Spectrosc.* **1991**, *45*, 1153.

(4) Steele, D. *Spectrochim. Acta* **1988**, *44A*, 1255.

(5) Malinowski, E. R. *Anal. Chem.* **1977**, *49*, 612.

(6) Bellamy L. J. *The Infrared Spectra of Complex Molecules*, 2nd ed.; Chapman and Hall: London, 1980; Vol. 2, p 14.

(7) Emmence, R. A.; Parker, M.; Smith, M. J. C.; Steele, D. *J. Mol. Struct.* **1993**, *292*, 295.

(8) Malinowski, E. R. *Factor Analysis in Chemistry*, 2nd ed.; John Wiley & Sons: New York, 1991.

(9) Emmence, R. A.; Steele D. *Anal. Chem.* **1991**, *63*, 2091.

(10) McPhail, R. A.; Strauss, H. L.; Snyder, R. G.; Elliger, C. A. *J. Phys. Chem.* **1984,** *88*, 334.

(11) Ricard-Lespade, L.; Longhi, G.; Abbate, S. *Chem. Phys.* **1990**, *142*, 245.

(12) Snyder, R. G.; Strauss, H. L.; Elliger, C. A. *J. Chem. Phys.* 1991, *94*, 5844.

(13) Child, M. S.; Lawton, R. T. *Faraday Discuss. Chem. Soc.* **1981**, *71*, 273.

(14) Mortensen, O. S.; Henry B. R.; Mohammadi M. A. *J. Chem. Phys.* **1981**, *75*, 4800.

(15) Kjaergaard, H. G.; Yu, H.; Schattka, B. J.; Henry, B. R.; Tarr, A. W. *J. Chem. Phys.* **1990**, *93*, 6239.

(16) Kjaergaard, H. G.; Henry, B. R.; Tarr, A. W. *J. Chem. Phys.* **1991**, *94*, 5844.

(17) McKean, D. C.; Biedermann, S.; Bürger, H. *Spectrochim. Acta* **1974**, *30A*, 845.

(18) Snyder, R. G.; Aljibury, A. L.; Strauss, H. L.; Casal, H. L.; Gough, K. M.; Murphy, W. F. *J. Chem. Phys.* **1984**, *81*, 5352.

(19) Gough, K. M; Henry, B. R.; Schattka, B. J; Wildman, T. A. *J. Phys. Chem.* **1991**, *95*, 1579.

(20) Kjaergaard, H. G.; Henry, B. R. *J. Chem. Phys.* **1992**, *96*, 4841.

(21) McKean, D. C.; Biedermann, S.; Bürger H. *Spectrochim Acta* **1974**, *30A*, 845.

(22) Dimitrova, Y.; Steele D. *Spectrochim. Acta* **1991**, *47A*, 75.

(23) Bellamy L. J. *′Infrared Spectra of Complex Molecules*, 3rd ed. John Wiley: New York, 1975; Vol. 1.

(24) Hill, I. R.; Levin, I. W. *J. Chem. Phys.* **1979**, *70*, 842.

(25) Wong, J. S.; Moore, C. B. *J. Chem. Phys.* **1982**, *77*, 603.

(26) McKean, D. C. *Chem. Soc. Rev.* **1978**, *7*, 399.

(27) Tosi, C.; Pinto, A. *Spectrochim. Acta* **1972**, *28A*, 585.

(28) Halonen, L.; Child, M. S. *J. Chem. Phys.* **1983**, *79*, 4355.

(29) Snyder, R. G.; Aljibury, A. L.; Strauss, H. L.; Casal, H. L.; Gough, K. M.; Murphy, W. F. *J. Chem. Phys.* **1984**, *81*, 5352.

(30) Palmö, K.; Mirkin, N. G.; Pietiläand, L. O.; Krimm, S. *Macromolecules* **1993**, *26*, 6831.

(31) Snyder, R. G. *J. Chem. Phys.* **1982**, *76*, 3342.