

*Российская Академия Наук
Институт геохимии и аналитической химии
им. В.И. Вернадского*

Лев А. Грибов

Упущенные и упускаемые
возможности создания новых
методов и новых приёмов анализа.

Москва, 2006 г.

Экспертные системы для идентификации неизвестных соединений

Начало → математизация качественного анализа по ИК и другим оптическим спектрам

Общий фундаментальный результат –
начало математической химии

Этапы:

система РАСТР (Россия)
система X-PERT (Германия)
система StrucEluc (Канада)

Возможность дальнейшего развития этого направления в России упущена навсегда

Квантовая химия

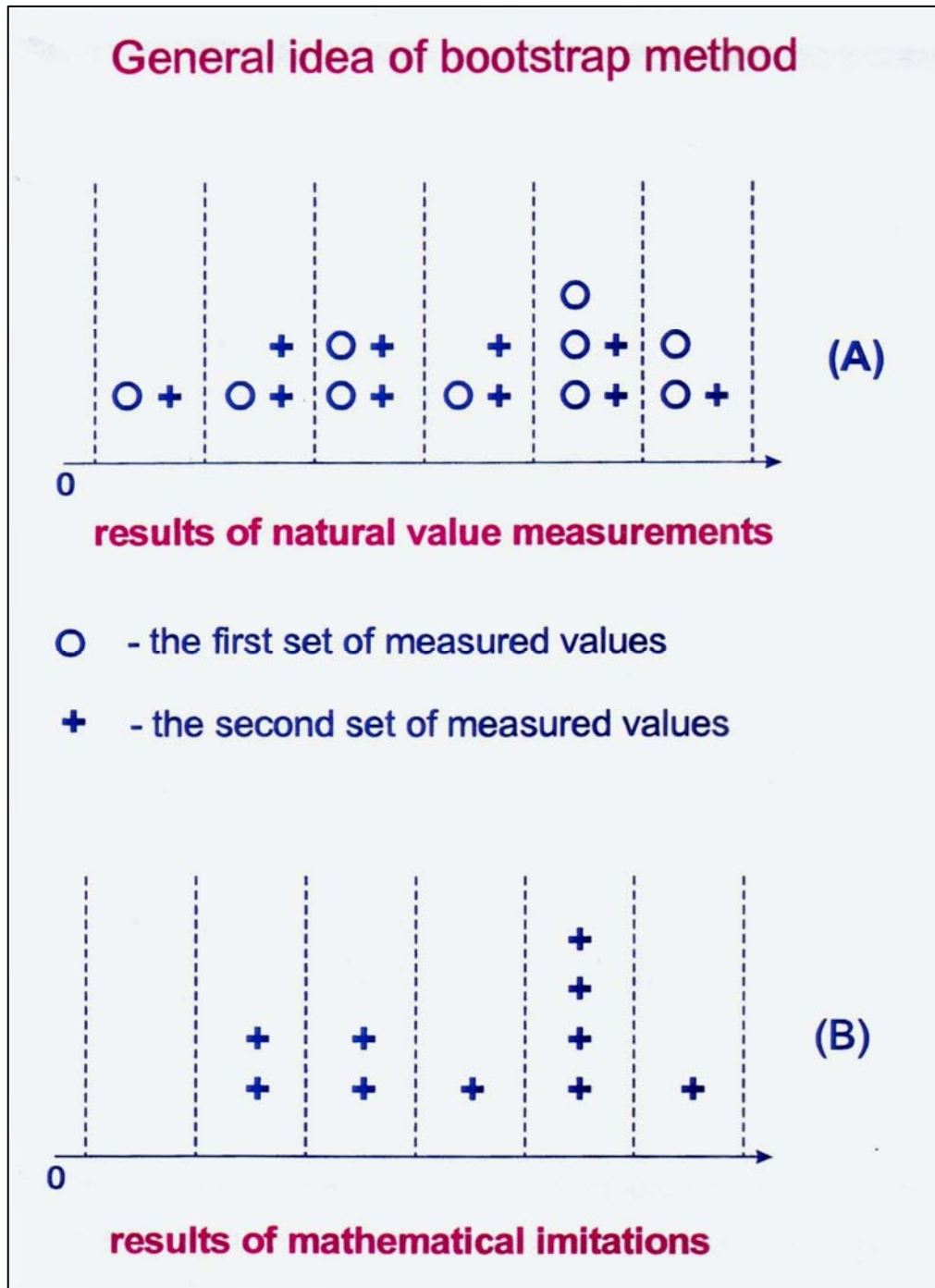
Прогноз действия органических реагентов
Расчёты геометрий молекул, распределения зарядов, хода ион-молекулярных реакций и т.д. Начиналось у нас в одно время с другими странами.

Не создано ни одной отечественной коммерческой программы.

Расширение практического применения полностью зависит от зарубежного продукта.

Обработка результатов наблюдений

Компьютерно-ориентированные статистические методы (1)



Компьютерно-ориентированные статистические методы (2)

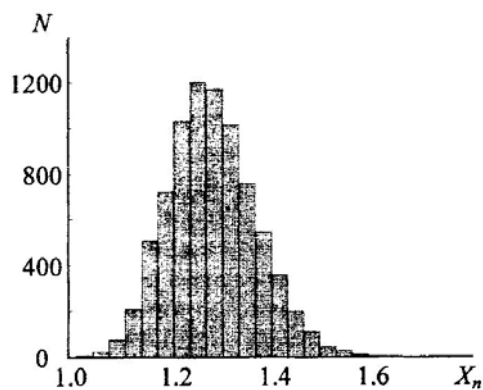


Рис. 1. Полученная по методу бутстрепа гистограмма распределения значений X_n для 30 молекул при 8000 математических испытаний: по вертикали N — число “попаданий” X_n в выбранный интервал значений, а по горизонтали — значения X_n .

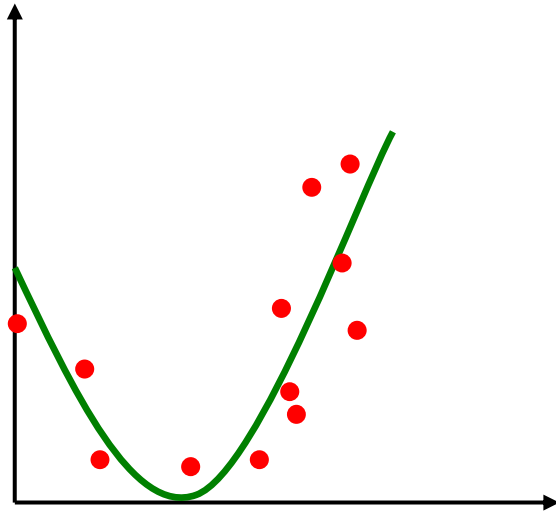
Компьютерно-ориентированные статистические методы (3)

Т а б л и ц а 1. Результаты количественного анализа молекул после обработки по методу бутстрепа

| Соединение | Брутто-формула | X_{\max} | Стандартное отклонение |
|---------------------------------|-----------------|------------|------------------------|
| 1,1-Дихлорпропен | $C_3H_4Cl_2$ | 0.969 | 0.078 |
| 1,2-Дихлорпропен | $C_3H_4Cl_2$ | 1.086 | 0.054 |
| 1,2-Эпокси-3-бутоксипропан | $C_7H_{14}O_2$ | 0.701 | 0.105 |
| 1-2-4-Триметилбензол | C_9H_{12} | 1.342 | 0.075 |
| 1-3-5-Триметилбензол | C_9H_{12} | 0.961 | 0.19 |
| 1-Гексанол | $C_6H_{14}O$ | 1.954 | 0.143 |
| 1-Гексен | C_6H_{12} | 0.894 | 0.117 |
| 1-Хлор-2-пропанол | C_3H_6OCl | 1.726 | 0.33 |
| 1-Хлорпентан | $C_5H_{11}Cl$ | 1.043 | 0.148 |
| 2-Аминотолуол | C_7H_9N | 1.431 | 0.21 |
| 2-Бром-2-хлор-1,1,1-трифторэтан | $C_2HFBrCl$ | 1.014 | 0.048 |
| 2-Бромбутан | C_4H_9Br | 1.302 | 0.196 |
| 2-Хлорпропан | C_3H_7Cl | 1.541 | 0.246 |
| 2-Этоксипропанол | $C_7H_{14}O$ | 0.981 | 0.04 |
| 3-Гидрокси-2-бутанол | $C_4H_8O_2$ | 1.24 | 0.179 |
| 3-Хлор-1-пропанол | C_3H_7OCl | 2.129 | 0.742 |
| Ацетальдегид диацетат | $C_6H_{10}O_4$ | 3.002 | 0.75 |
| Бензальдегид | C_7H_6O | 1.620 | 0.161 |
| Бензил ацетат | $C_8H_{10}O_2$ | 1.129 | 0.158 |
| Бензиловый спирт | C_7H_9O | 1.323 | 0.118 |
| Бромбензол | C_6H_5Br | 1.130 | 0.059 |
| Диметиламин | C_2H_7N | 0.720 | 0.03 |
| Дифениловый эфир | $C_{12}H_{10}O$ | 1.231 | 0.06 |
| Метиленидиметилгидразин | $C_3H_8N_2$ | 1.072 | 0.09 |
| Морфолин | C_4H_9NO | 1.741 | 0.02 |
| Несимметричный диметилгидразин | $C_2H_8N_2$ | 0.781 | 0.33 |
| Уксусный ангидрид | $C_4H_6O_3$ | 1.008 | 0.09 |
| Хлорпентафторэтан | C_2F_5Cl | 1.130 | 0.11 |
| Этилгептаноат | $C_9H_{18}O_2$ | 1.004 | 0.026 |
| Этилбутаноат | $C_5H_{12}O_2$ | 1.079 | 0.059 |

П р и м е ч а н и е: X_{\max} соответствует интервалу, в который попадает наибольшее число X_n .

Многопараметрический бутстреп

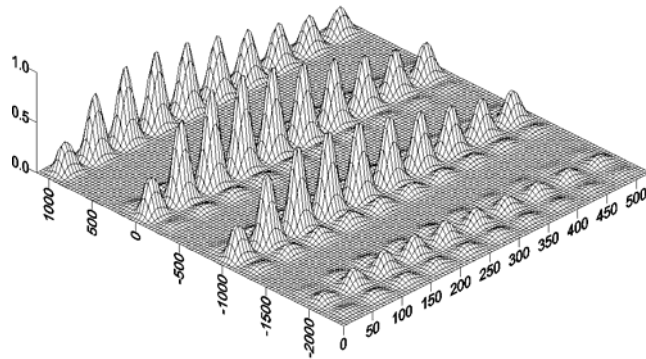


$$y = k(x - a)^2 \quad \text{Неизвестные } k \text{ и } a.$$

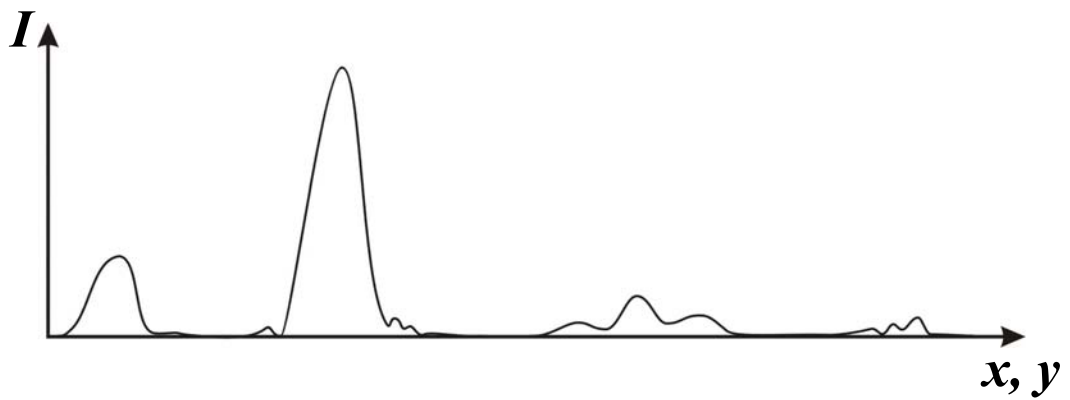
$$F = \sum_n \left(y_{n,\text{эксп}} - y_{n,\text{расч}} \right)^2$$

Минимизация по методу Монте-Карло

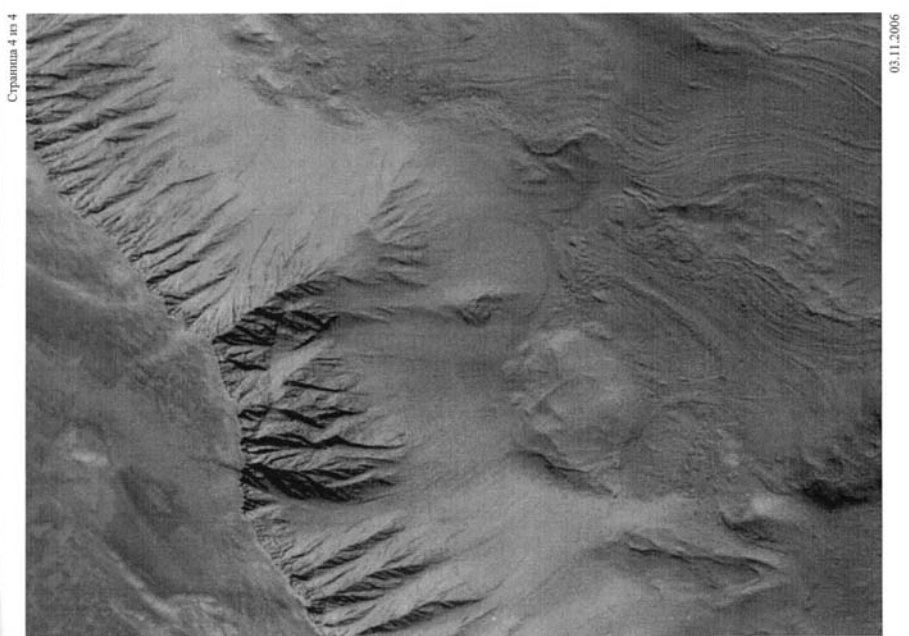
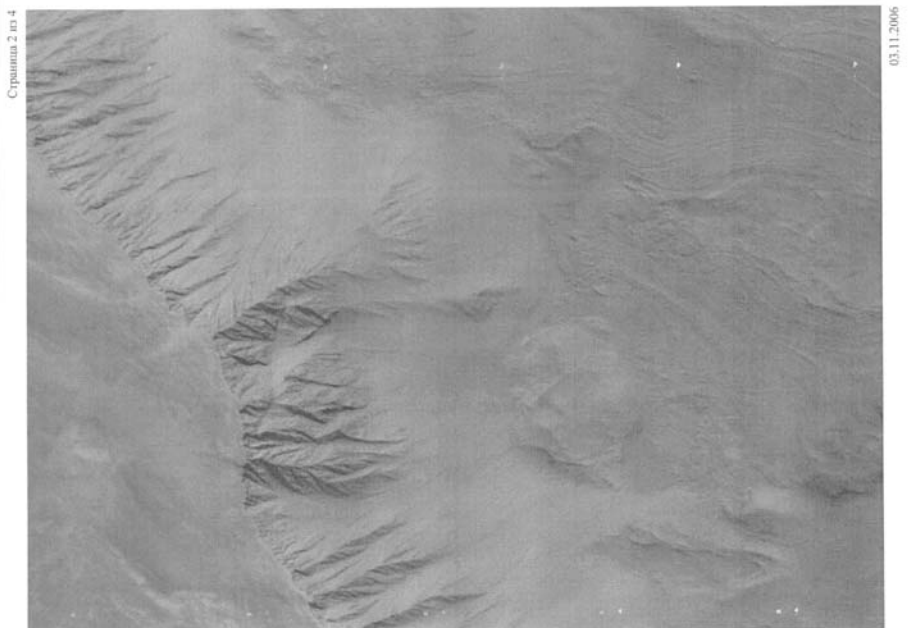
Трёхмерный спектр



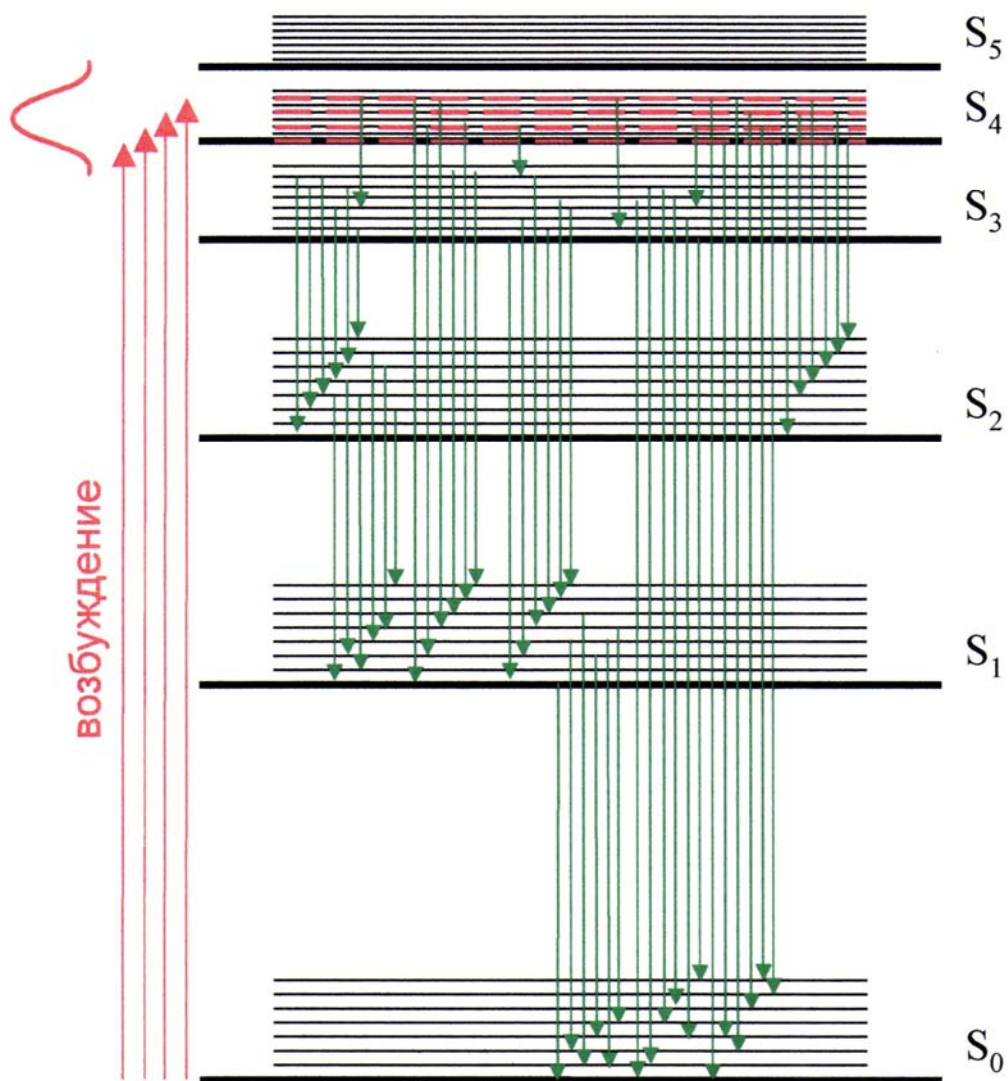
Сечение под углом 40°



Увеличение контрастности изображений (программа Photoshop)



Система вибронных уровней и переходов многоатомной молекулы



Количество переходов

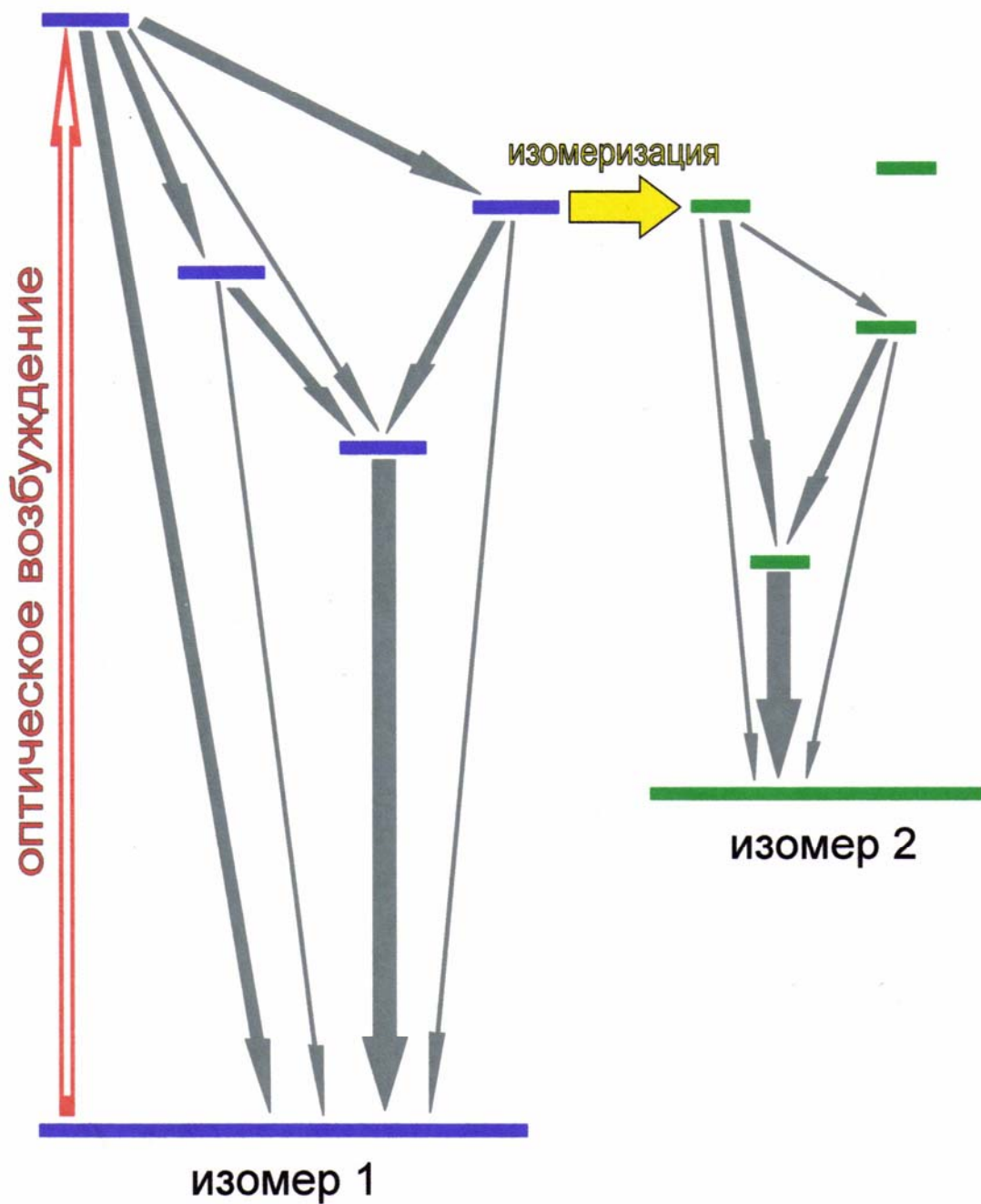
$$N_{\text{w}} > 10^6$$

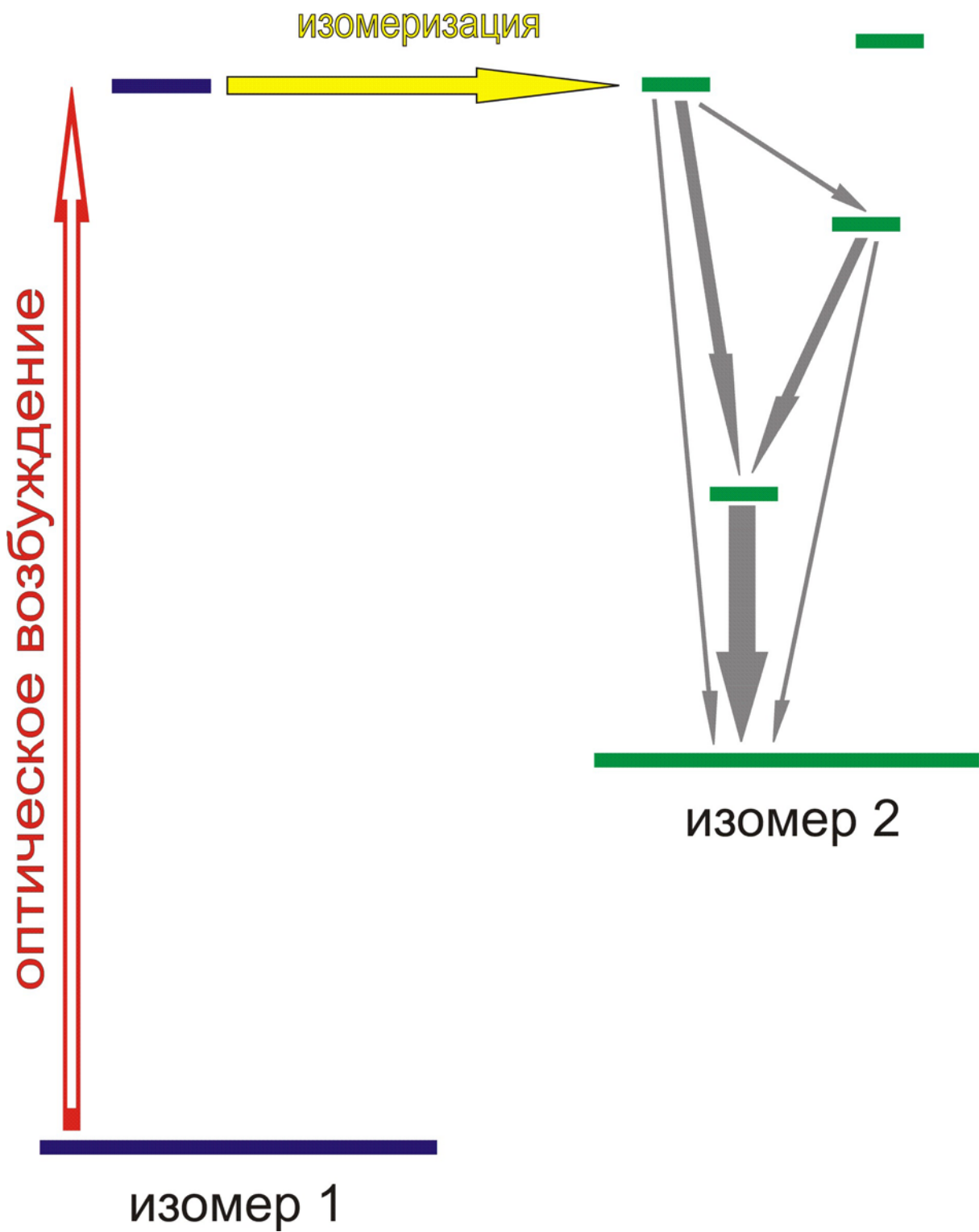
“Многомерная” спектроскопия

“Спектральные” переменные:

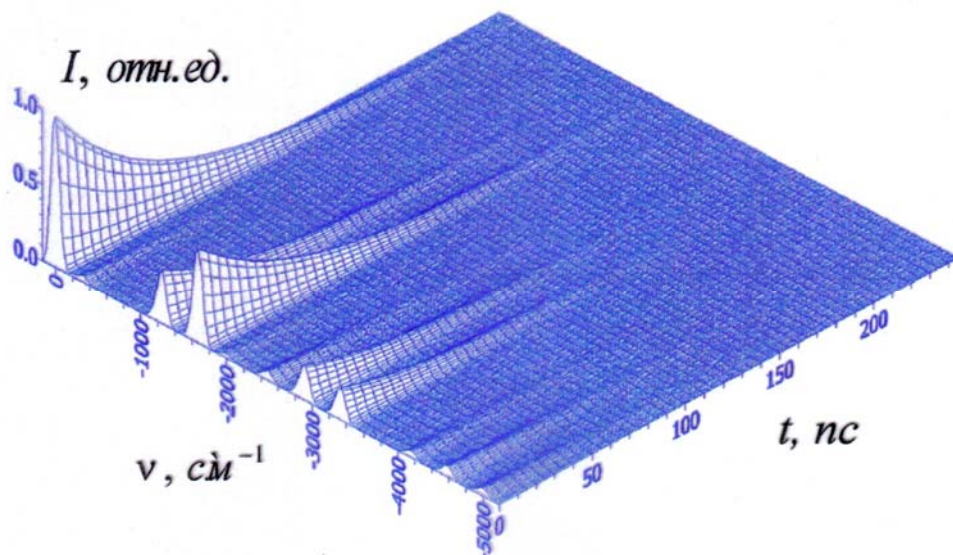
- 1. Значение интенсивности линии поглощения или излучения.**
- 2. Значение длины волны (волнового числа) линии поглощения или излучения.**
- 3. Значение длины волны возбуждения.**
- 4. Время релаксации при импульсном возбуждении.**
- 5. Значение длительности импульса возбуждения.**
- 6. “Спектральная” форма импульса возбуждения.**

Схема диссипации энергии в системе с возможной изомеризацией

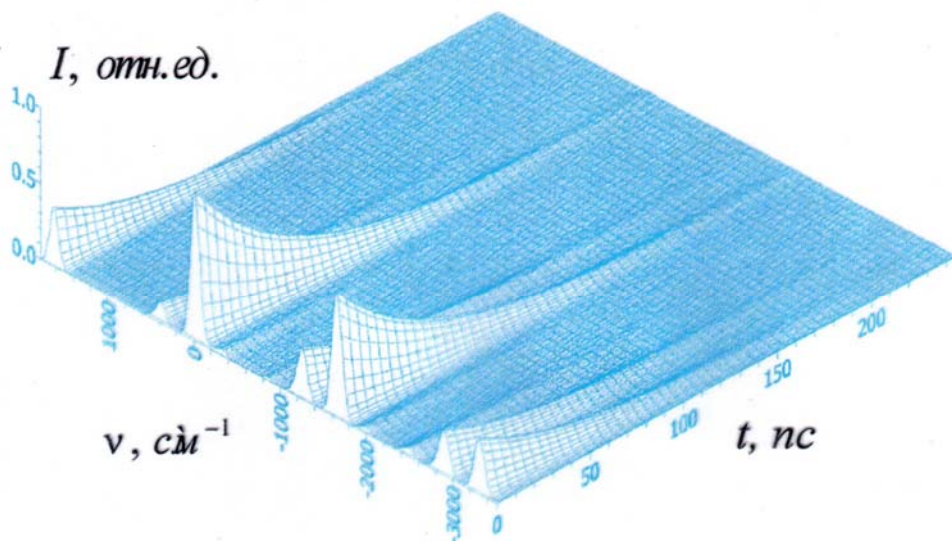




Спектры при различных условиях возбуждения (1)

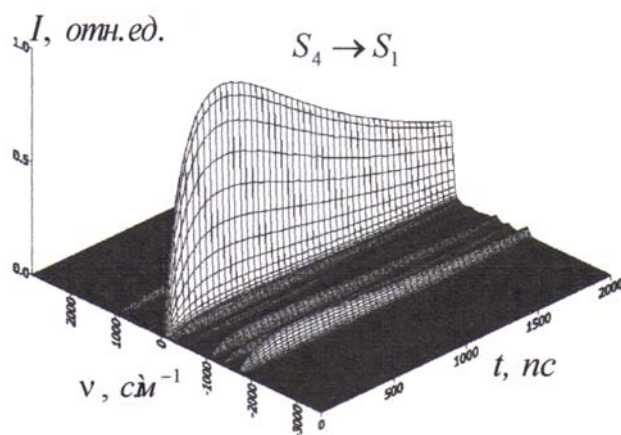
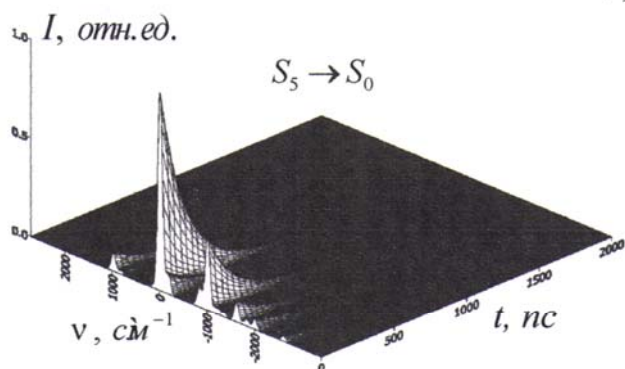
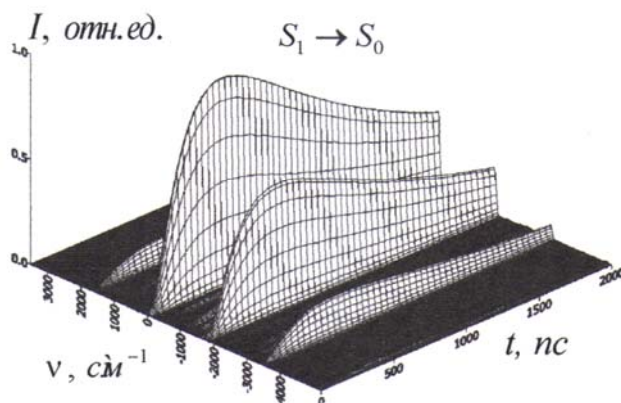
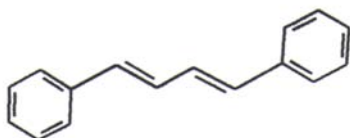


Резонансное возбуждение S_1



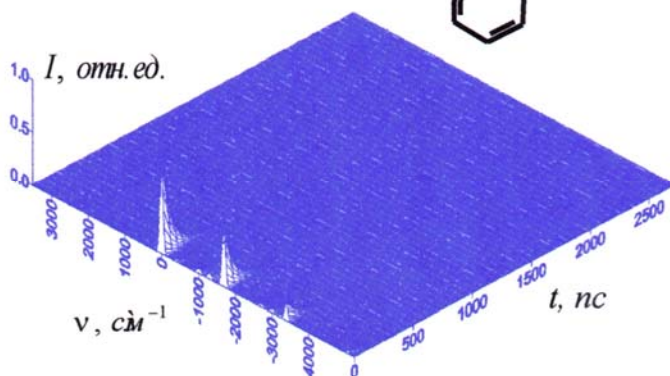
Резонансное возбуждение $S_1 + 1620 \text{ cm}^{-1}$

Спектры при различных условиях возбуждения (2)

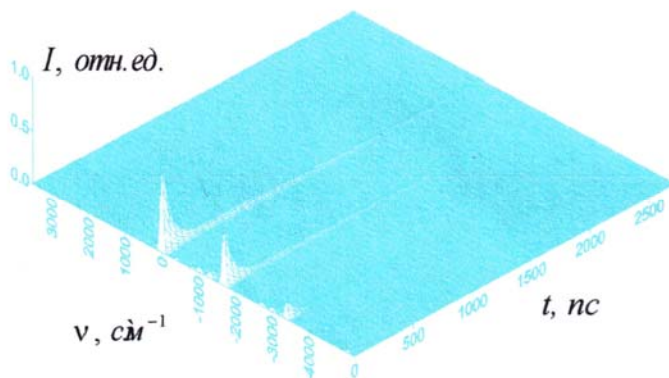


Широкополосное возбуждение колебательных подуровней состояния S_5

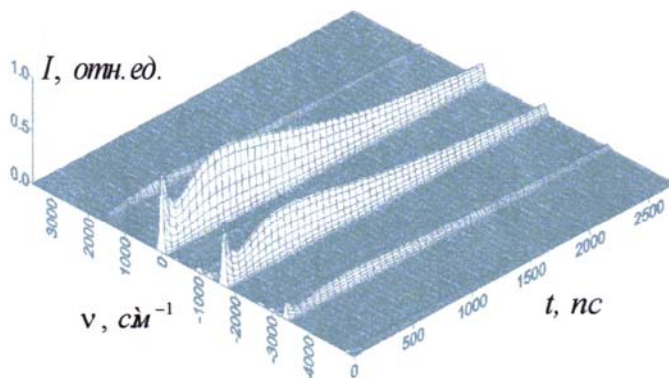
Спектры при различных условиях возбуждения (3)



Возбуждение состояния S_1



Одновременное возбуждение состояний S_1 и S_5 ($I_1 / I_5 = 100$)



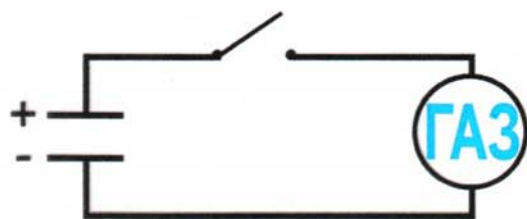
Одновременное возбуждение состояний S_1 и S_5 ($I_1 / I_5 = 1000$)

Получение коротких вспышек света



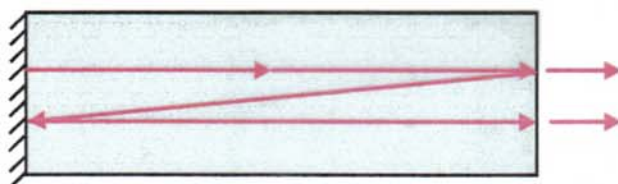
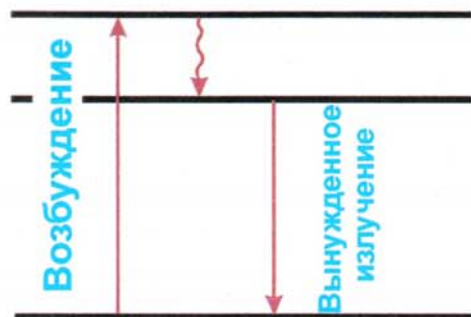
10^{-2} сек

Прерывание
непрерывного пучка
света (опыт Физо)



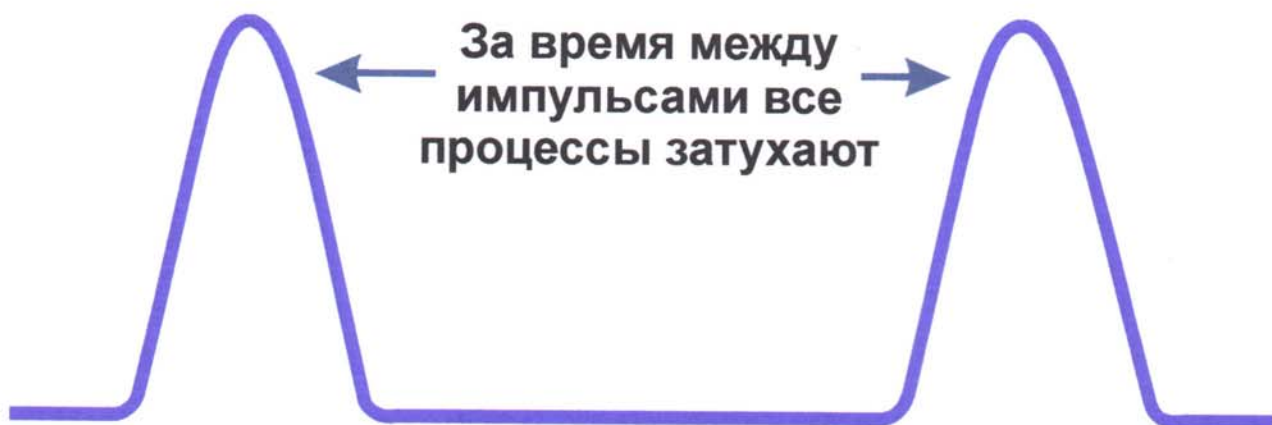
10^{-5} сек

Газовый разряд от
конденсатора (флаш-
фотолиз)

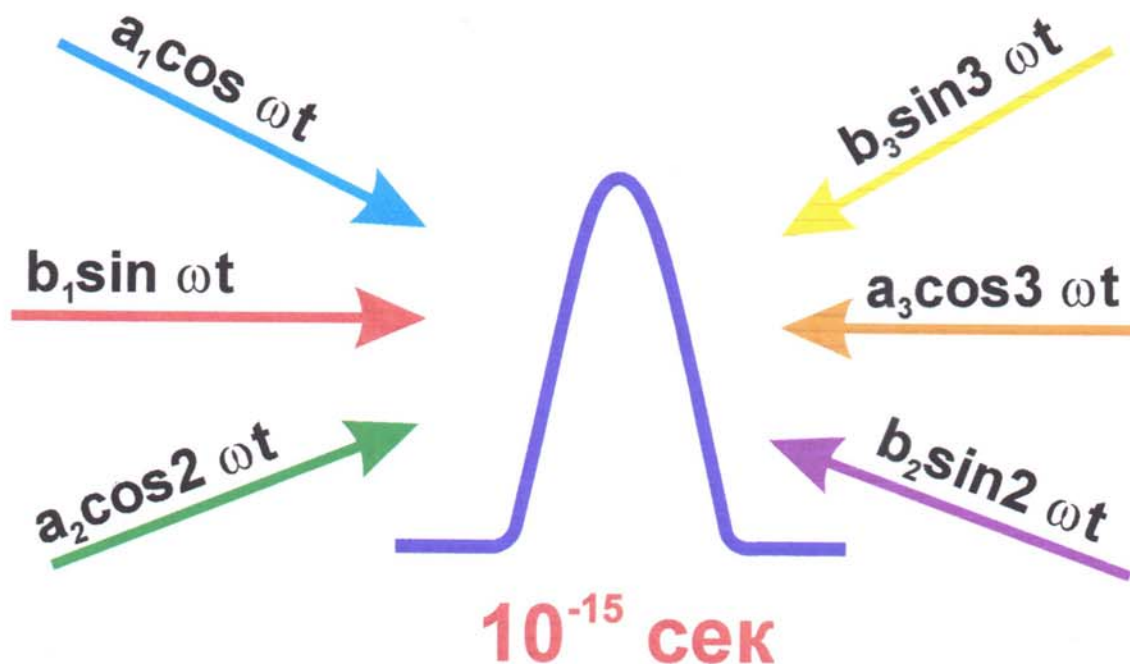


Импульсные лазеры:
Наносекундный
диапазон - 10^{-9} сек
Пикосекундный
диапазон - 10^{-12} сек

Идея получения сверхкоротких импульсов

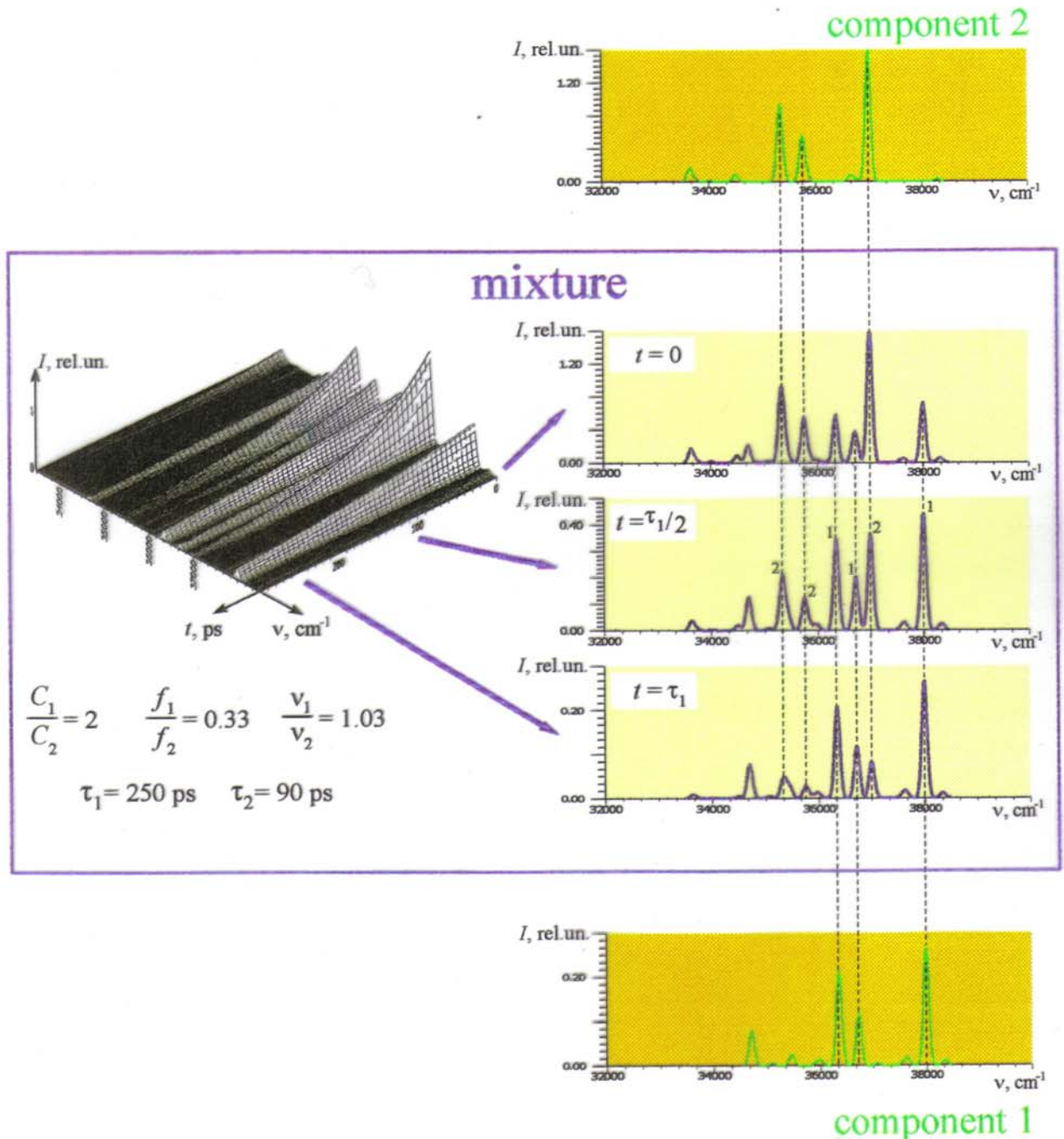


$$f(t) = \frac{a_0}{2} + \sum_k [a_k \cos(k\omega t) + b_k \sin(k\omega t)]$$

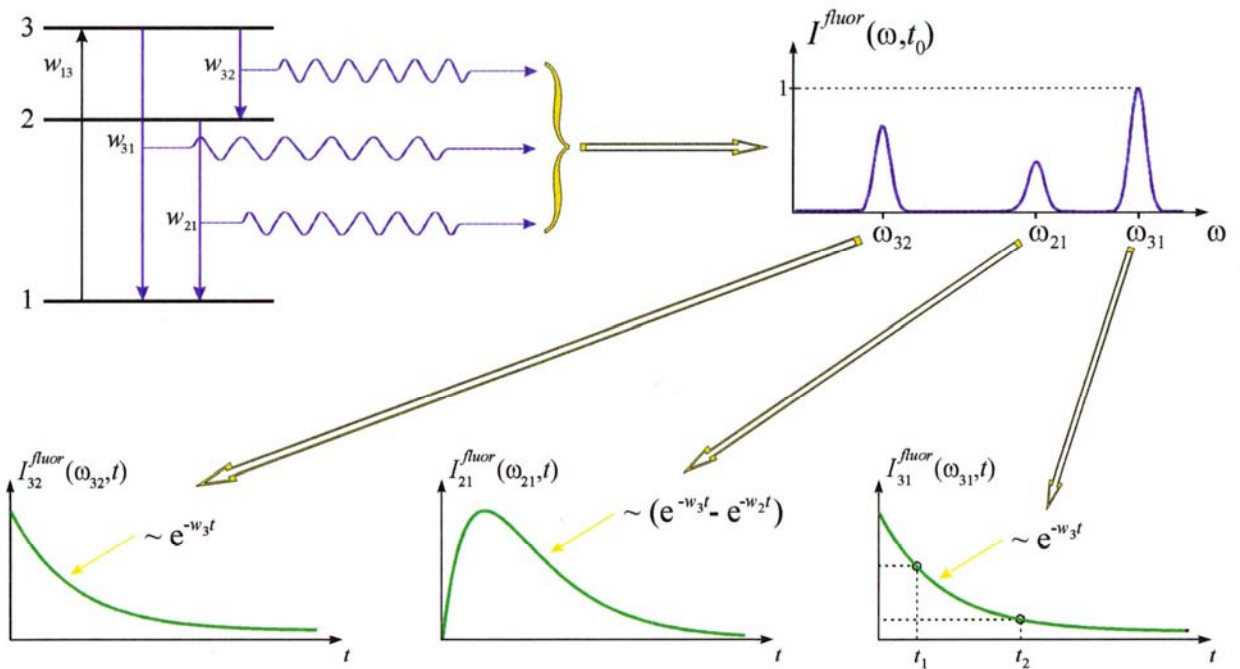


Фемтосекундный диапазон

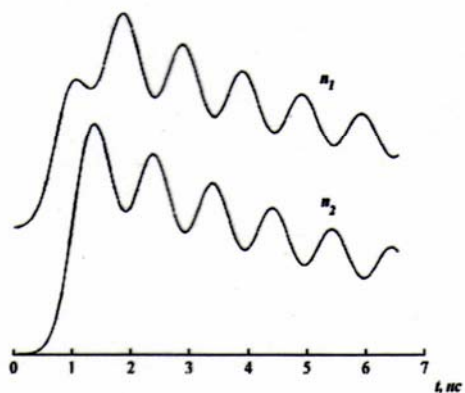
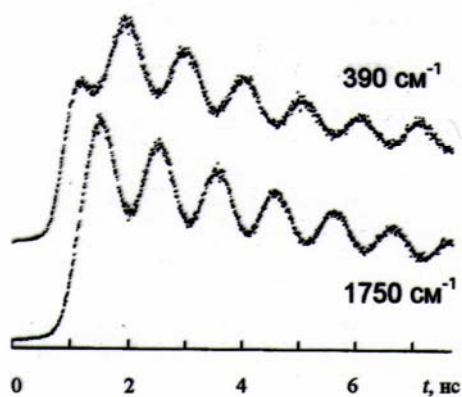
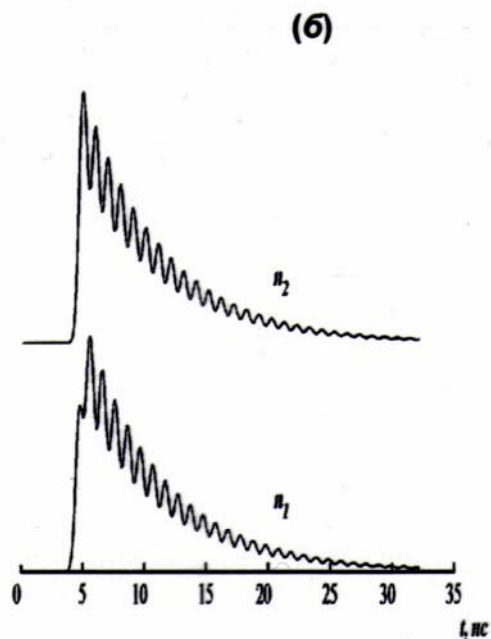
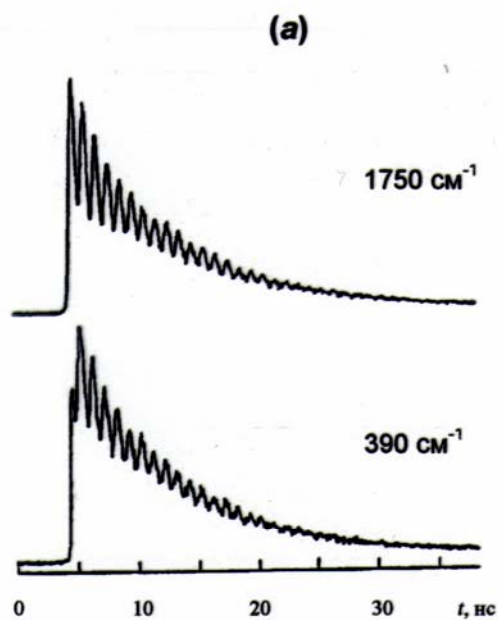
Time-resolved fluorescence spectrum of two-component mixture



Three-level system



МОДЕЛИРОВАНИЕ КИНЕТИКИ ВНУТРИМОЛЕКУЛЯРНЫХ ПРОЦЕССОВ
И СПЕКТРОВ С ВРЕМЕННЫМ РАЗРЕШЕНИЕМ С УЧЕТОМ
БЕЗЫЗЛУЧАТЕЛЬНЫХ ПЕРЕХОДОВ

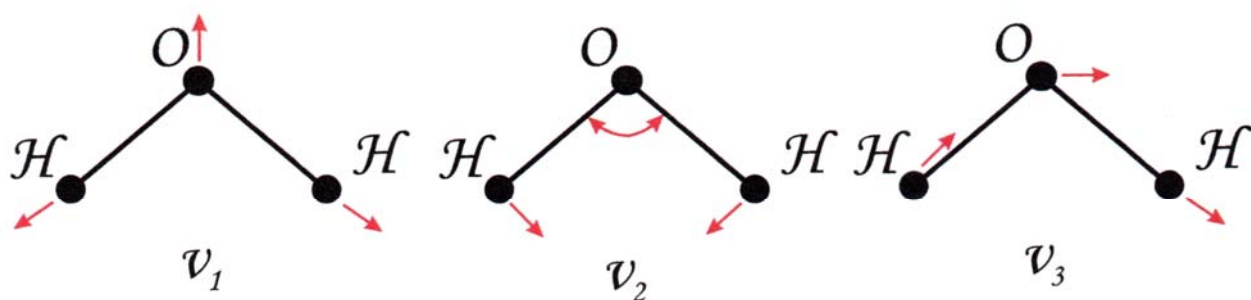


Экспериментальные (а) и рассчитанные (б) временные зависимости спектров флуоресценции из "резонирующих" вибронных уровней (1750 и 390 cm^{-1}).

Временное разрешение ~10 пс.

Происхождение обертонов

Нормальные колебания молекулы H_2O

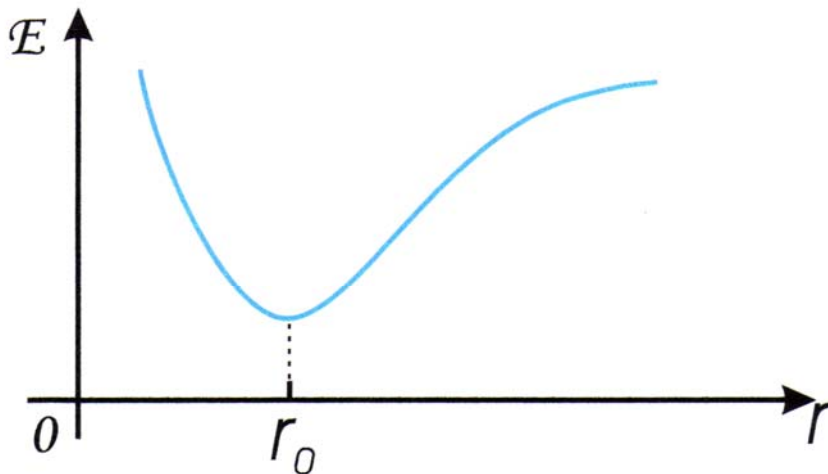


| v_1 | v_2 | v_3 |
|-------|-------|-------|
| 0 | 0 | 0 |
| 1 | 0 | 0 |
| 0 | 1 | 0 |
| 0 | 0 | 1 |
| 2 | 0 | 0 |
| 0 | 2 | 0 |
| 0 | 0 | 2 |
| 1 | 1 | 0 |
| 1 | 0 | 1 |
| 0 | 1 | 1 |

$$\omega(v_1, v_2, v_3) = \sum_k \omega_k v_k$$

С ростом суммы квантовых чисел спектр усложняется и смещается в область ближнего ИК

Потенциальная кривая для основного состояния



r_0 - равновесная геометрия

Волновая функция Ψ зависит от r

Интенсивности полос в ИКС

$$J_i \sim \left(\frac{\partial \vec{\mu}}{\partial Q_i} \right)_0 \text{ - для основных колебаний}$$

$$J_i \sim \left(\frac{\partial^2 \vec{\mu}}{\partial Q_i^2} \right)_0 \text{ - для обертонов}$$

$$J_{ij} \sim \left(\frac{\partial^2 \vec{\mu}}{\partial Q_i \partial Q_j} \right)_0 \text{ - для составных частот}$$

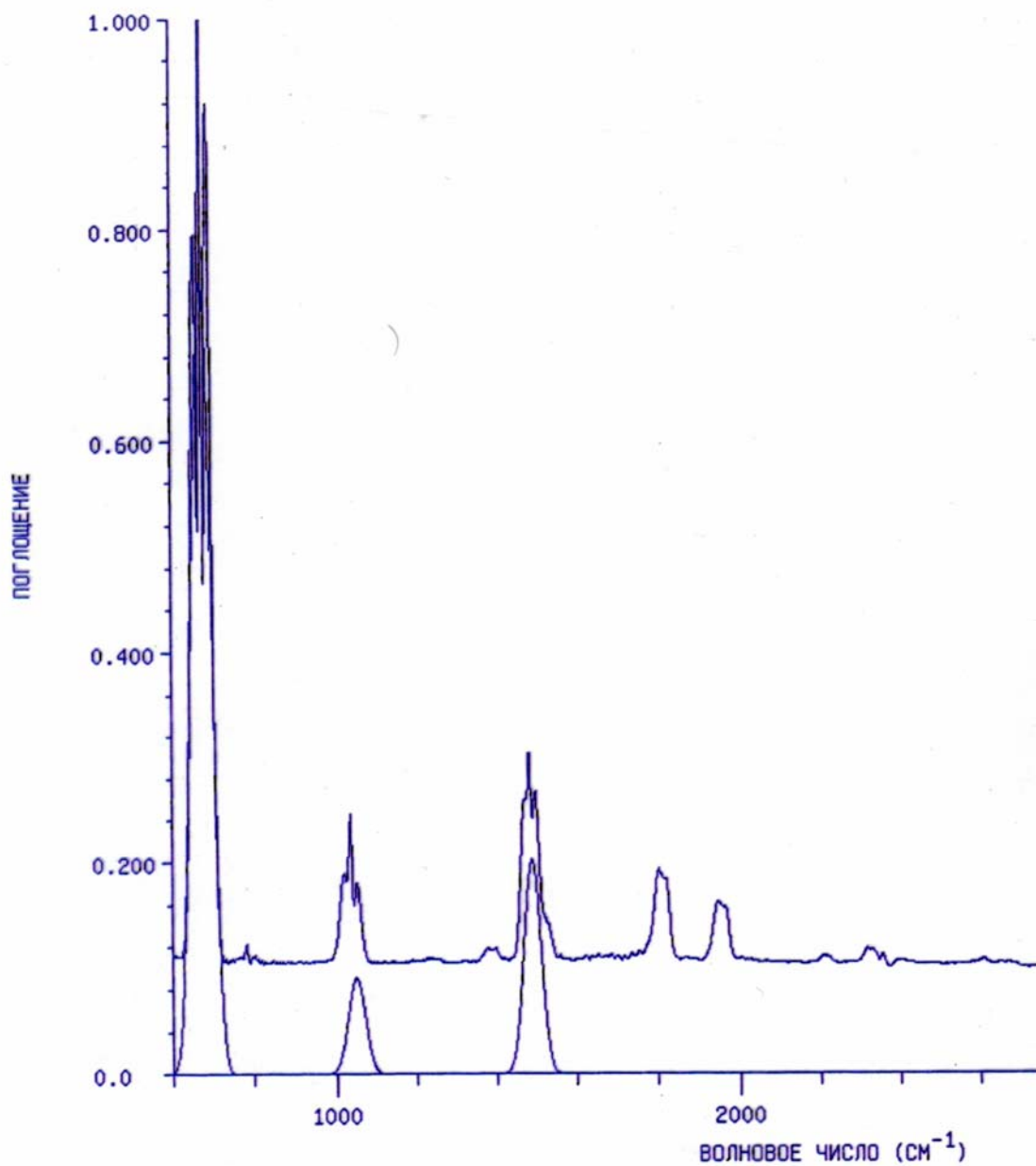
$\vec{\mu}$ - дипольный момент; Q_i - нормальная координата

Индекс i означает сочетание квантовых чисел

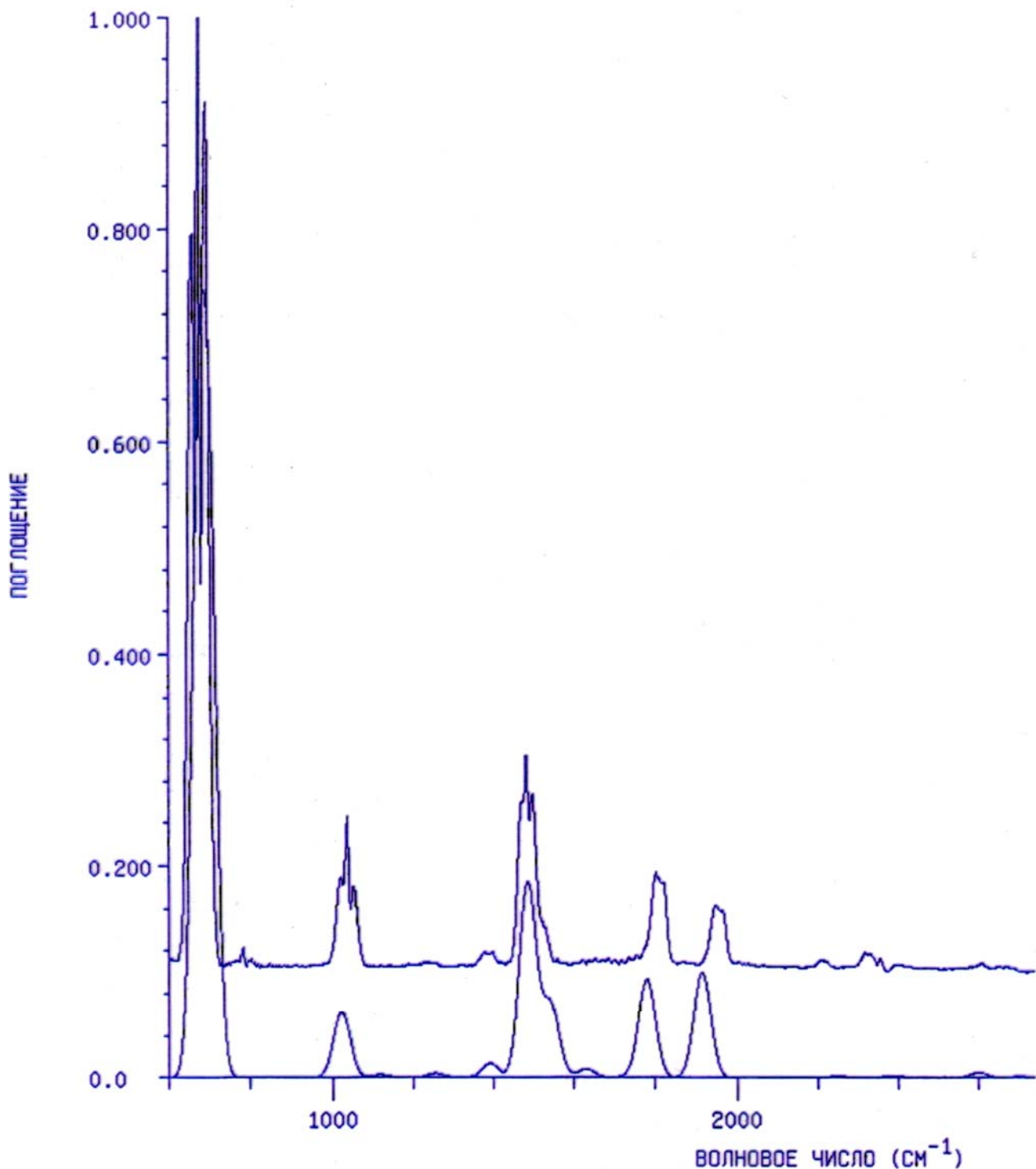
**Экспериментальные и вычисленные уровни энергии в см⁻¹
для молекулы ¹²CO₂***

| Сим-метрия | Отнесение | | | Эксперимент | Расчет | Сим-метрия | Отнесение | | | Эксперимент | Расчет |
|-----------------------|----------------|----------------|----------------|-------------|--------|-----------------------|----------------|----------------|----------------|-------------|---------|
| | v ₁ | v ₂ | v ₃ | | | | v ₁ | v ₂ | v ₃ | | |
| <i>E_{2u}</i> | [1 | 2 | 2] | 6103,7 | 6091,7 | <i>E_{1u}</i> | [2 | 2 | 1] | 7743,7 | 7732,6 |
| <i>E_{2u}</i> | [1 | 1 | 4] | 6288,5 | 6284,2 | <i>E_{1u}</i> | [2 | 1 | 3] | 7901,5 | 7897,9 |
| <i>E_{2u}</i> | [1 | 0 | 6] | 6474,5 | 6476,8 | <i>E_{1u}</i> | [2 | 0 | 5] | 8056,0 | 8058,8 |
| <i>E_{1u}</i> | [2 | 1 | 1] | 6538,0 | 6534,2 | <i>A_{1u}</i> | [3 | 1 | 0] | 8192,6 | 8189,8 |
| <i>E_{1u}</i> | [2 | 0 | 3] | 6679,7 | 6680,4 | <i>A_{1u}</i> | [3 | 0 | 2] | 8294,0 | 8294,7 |
| <i>E_{1g}</i> | [1 | 3 | 1] | 6688,2 | 6670,0 | <i>E_{2u}</i> | [3 | 0 | 2] | 8232,9 | 8231,9 |
| <i>E_{1g}</i> | [1 | 2 | 3] | 6863,6 | 6854,9 | <i>E_{1g}</i> | [1 | 2 | 5] | 8250,6 | 8242,1 |
| <i>E_{1g}</i> | [1 | 1 | 5] | 7023,7 | 7021,8 | <i>E_{1g}</i> | [1 | 1 | 7] | 8425,0 | 8424,7 |
| <i>E_{1g}</i> | [1 | 0 | 7] | 7203,8 | 7208,5 | <i>E_{1g}</i> | [1 | 0 | 9] | 8628,7 | 8638,0 |
| <i>E_{1g}</i> | [3 | 1 | 1] | 8803,3 | 8798,1 | <i>A_{1u}</i> | [5 | 0 | 0] | 11496,4 | 11498,8 |
| <i>E_{1g}</i> | [3 | 0 | 3] | 8944,1 | 8945,0 | <i>E_{1g}</i> | [5 | 0 | 1] | 12101,6 | 12102,0 |
| <i>E_{3g}</i> | [3 | 0 | 3] | 8863,7 | 8861,7 | <i>A_{1u}</i> | [5 | 1 | 0] | 12672,3 | 12668,6 |
| <i>A_{1g}</i> | [4 | 0 | 0] | 9246,9 | 9248,2 | <i>A_{1u}</i> | [5 | 0 | 2] | 12774,7 | 12776,9 |
| <i>A_{1u}</i> | [1 | 4 | 2] | 8676,7 | 8650,7 | <i>A_{1g}</i> | [6 | 0 | 0] | 13721,0 | 13725,0 |
| <i>A_{1u}</i> | [1 | 2 | 6] | 8963,9 | 8960,0 | <i>E_{1u}</i> | [6 | 0 | 1] | 14313,9 | 14314,9 |
| <i>A_{1u}</i> | [1 | 1 | 8] | 9137,8 | 9139,6 | <i>A_{1u}</i> | [7 | 0 | 0] | 15921,0 | 15926,9 |
| <i>A_{1u}</i> | [3 | 2 | 0] | 9389,0 | 9379,5 | <i>E_{1g}</i> | [7 | 0 | 1] | 16501,6 | 16503,4 |
| <i>A_{1u}</i> | [3 | 1 | 2] | 9517,0 | 9514,3 | <i>A_{1g}</i> | [8 | 0 | 0] | 18096,3 | 18104,4 |
| <i>A_{1u}</i> | [3 | 0 | 4] | 9631,3 | 9633,8 | <i>E_{1u}</i> | [8 | 0 | 1] | 18664,6 | 18667,4 |
| <i>E_{2u}</i> | [3 | 1 | 2] | 9419,2 | 9411,3 | <i>A_{1u}</i> | [9 | 0 | 0] | 20246,9 | 20257,7 |
| <i>E_{2u}</i> | [3 | 0 | 4] | 9589,9 | 9590,6 | <i>E_{1g}</i> | [9 | 0 | 1] | 20803,0 | 20807,1 |
| <i>E_{1u}</i> | [4 | 0 | 1] | 9864,4 | 9864,5 | <i>A_{1g}</i> | [10 | 0 | 0] | 22373,9 | 22386,9 |

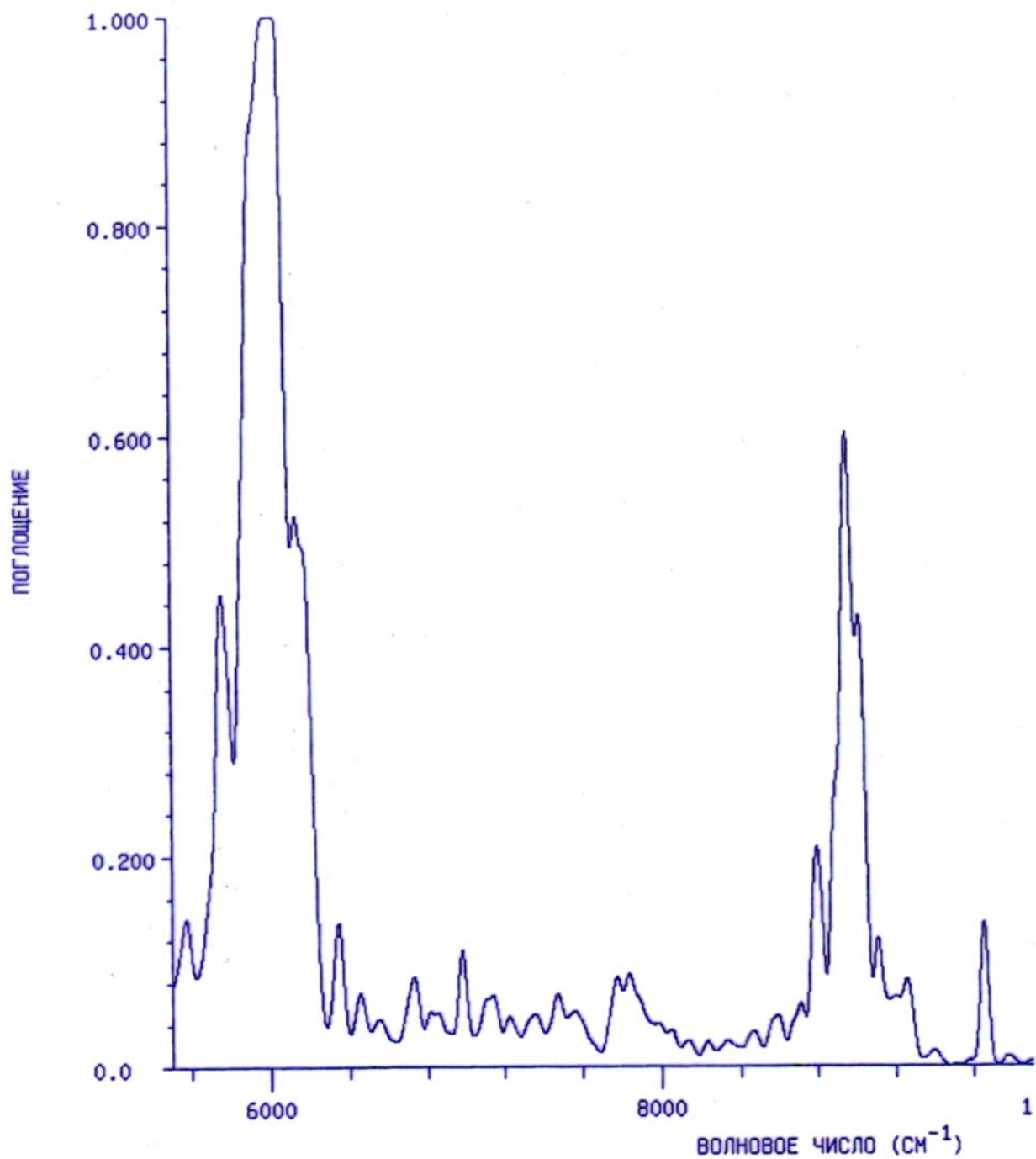
* Расчет выполнен в базе 27 405 смешанных Морзевско-гармонических функций. Экспериментальные уровни энергии взяты из работы [25]. Здесь и далее в квадратных скобках – уровни, участвующие в резонансе Ферми.



Экспериментальный (верхняя кривая) и расчётный в гармоническом приближении спектры (нижняя кривая) бензола. В расчётном спектре отсутствуют дублет в области 1450 см^{-1} и полосы выше 1800 см^{-1} .

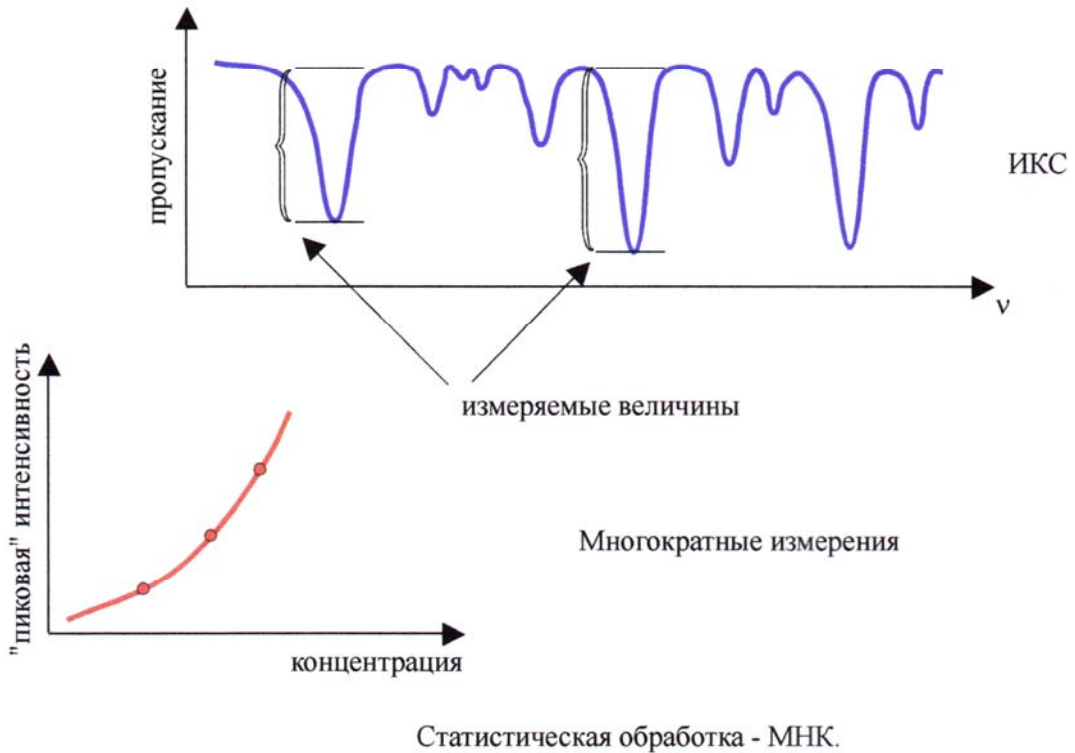


Экспериментальный (верхняя кривая) и расчётный в ангармоническом приближении (нижняя кривая) спектры бензола. Воспроизводится дублет и полосы выше 1800 см⁻¹.

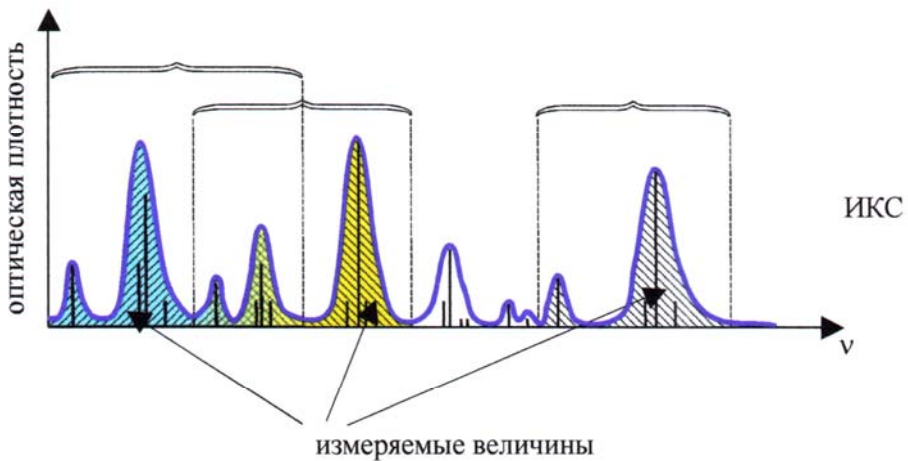


Теоретический спектр поглощения бензола в оберточной области.

Количественный спектральный анализ на базе образцов стандартного состава



"Безэталонный" анализ



Однократное измерение. Слабая чувствительность к точности вычисления положения полос.

Статистическая обработка - бутстреп.

Table 3.

Examples of quantitative analysis of individual compounds based on bootstrap processing.

ETHYL FORMATE (c=30ppm)

The number of sets M=5000

The number of variables N=22

BOOTSTRAP gives distribution with parameters:

$X_{main}=36$, STANDARD=3

GISTOGRAM for X_{mean} from random sets

| | | | |
|----|----|------|-------|
| 27 | 28 | .1 | |
| 29 | 30 | 1.3 | * |
| 31 | 32 | 5.3 | ***** |
| 33 | 34 | 12.1 | ***** |
| 35 | 36 | 20.3 | ***** |
| 37 | 38 | 21.8 | ***** |
| 39 | 40 | 17.7 | ***** |
| 41 | 42 | 11.5 | ***** |
| 43 | 44 | 6.1 | ***** |
| 45 | 46 | 2.4 | ** |
| 47 | 48 | .8 | |
| 49 | 50 | .2 | |

Table 2.

Results of IR quantitative analysis of some organic compounds based on *ab initio* absolute intensities of bands calculations.

| Substance and injected concentration (ppm) | Number of spectral intervals and their combinations | Results of concentration determinations based on bootstrap method | Error (standard deviation) ($\pm\Delta$) (ppm) |
|--|---|---|--|
| p-Xylene (190) | 10 | 165 | 9 |
| m-Xylene (180) | 21 | 180 | 6 |
| Ethylbenzene (250) | 22 | 227 | 13 |
| Methylcyclohexane (500) | 9 | 453 | 106 |
| Methyl Acetate (30) | 22 | 32 | 1 |
| Ethyl Formate (30) | 22 | 36 | 3 |
| Ethyl Acetate (25) | 22 | 31 | 2 |
| Hexane (400) | 16 | 424 | 27 |
| o-Xylene (150) | 28 | 139 | 8 |
| Butyl Acetate (30) | 29 | 29 | 2 |
| p-Ethyltoluene (120) | 21 | 114 | 7 |
| Toluene (150) | 21 | 165 | 13 |
| 1,2,4-Trimethylbenzene (200) | 15 | 181 | 21 |
| Ethylene (450) | 4 | 411 | 10 |
| Butadiene (300) | 16 | 311 | 6 |
| Propylene (450) | 11 | 462 | 6 |
| Trans-2-Butene (450) | 4 | 437 | 5 |
| Cys-2-Butene (450) | 16 | 657 | 11 |
| 1-Hexene (300) | 16 | 309 | 4 |
| 1,4-Pentadiene (300) | 15 | 311 | 23 |
| Trans-2-Pentene (180) | 11 | 236 | 19 |
| Allyl chloride (300) | 29 | 248 | 22 |
| Allyl bromide (300) | 16 | 218 | 3 |
| Allyl fluoride (195) | 11 | 211 | 4 |
| 1,1-Difluoro-1-chloro ethane (150) | 22 | 41 | 5 |

Table 5.

Standartless molecular spectral analysis of mixtures based on mathematical modelling (*ab initio* calculations of IRS band intensities).

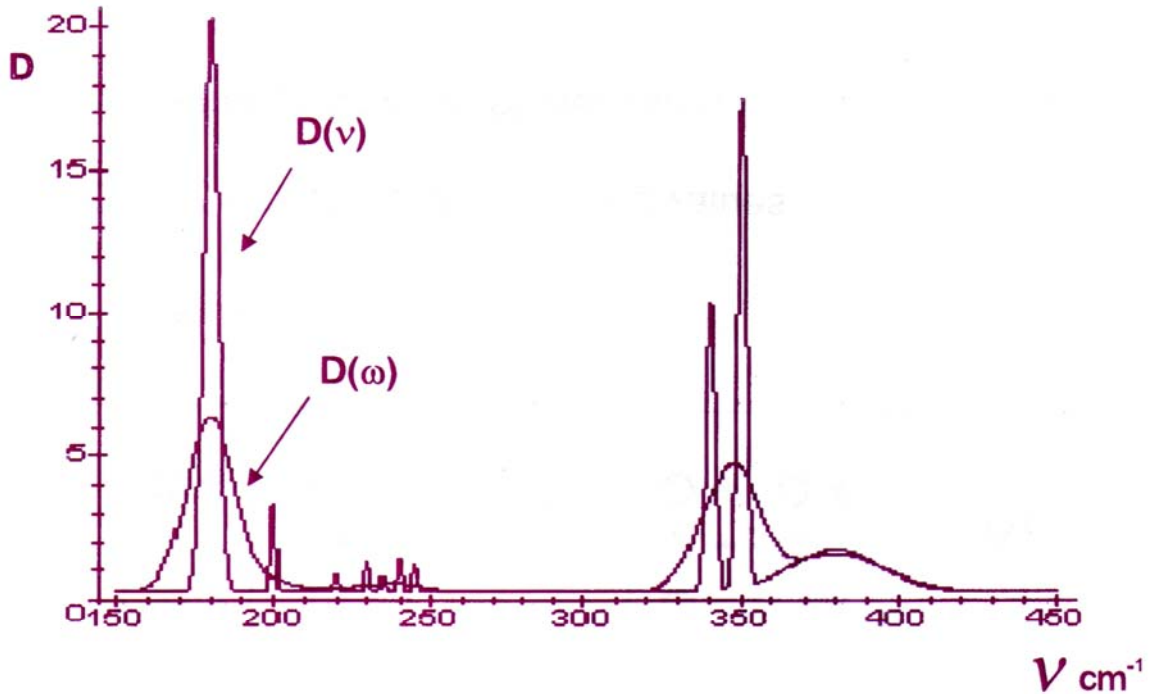
| Substance and injected concentration (ppm) | Number of spectral intervals and their combinations | Results of concentration determinations based on bootstrap method | Error (standard deviation) ($\pm\Delta$) (ppm) |
|---|---|---|--|
| MIX 1: Hexane (502) o-Xylene (63.4) p-Ethyltoluene (185) | 17 | 493 56 183 | 46 3 6 |
| MIX 2: Hexane (471) Toluene (38.6) 1,2,4-Trimethylbenzene (88.9) p-Ethyltoluene (150) | 17 | 472 31 93 122 | 33 2 21 12 |
| MIX 3: Hexane (593) o-Xylene (87) Butyl Acetate (20.9) | 16 | 666 100 17 | 28 5 1 |

How to compare experimental and calculated spectrum?

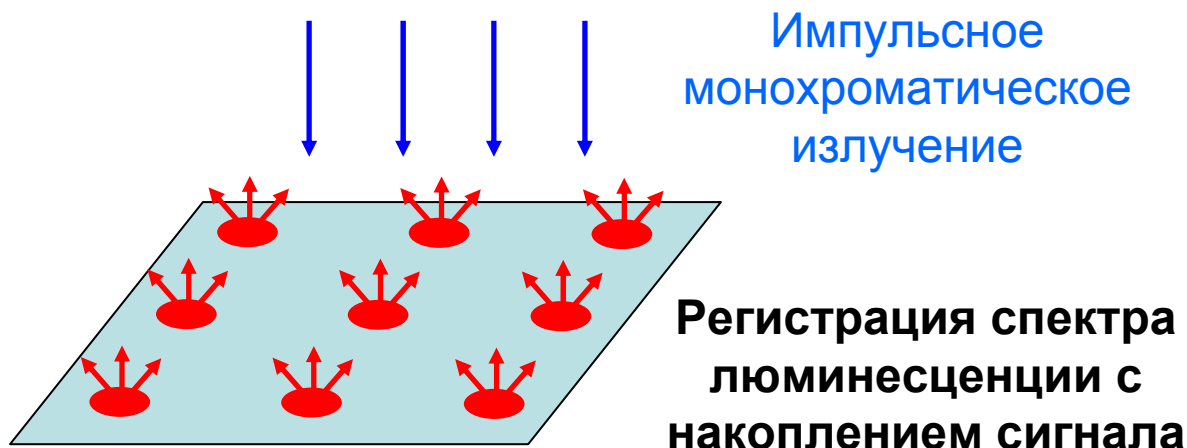
It is possible only to calculate positions of spectral bands and their integral intensities.

This result leads to transformation of usual spectral curve $D(\nu)$ to the curve $D(\omega)$ - spectral distribution of integral value $D_{int}(\omega)$:

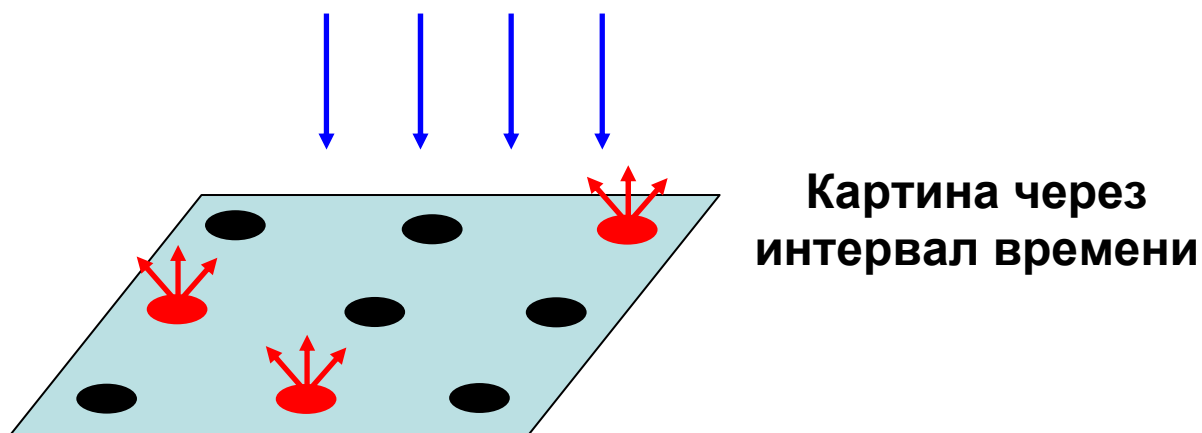
$$D_{int}(\omega) = \frac{1}{\delta \sqrt{\pi}} \int_{-\infty}^{\infty} e^{-\frac{(\omega - \nu)^2}{\delta^2}} D(\nu) d\nu$$



Спектры одиночных молекул



Тушение люминесценции при химическом превращении



Перспектива – создание спектральных микрочипов