

スペクトルデータ情報の共有化

—共通フォーマット化赤外スペクトルデータのパソコン上での利用—

伊藤 朋恭*, 田代 敏彦**, 鹿野比呂子*

要 約

パソコンの普及は目覚ましく、様々な測定機器にもパソコンが付属するようになってきた。これらの機器によって測定されたデータはデジタル化されて、付属のパソコンに保存される。このデジタル化されたデータには、誰もが同じ精度でデータを見ることができるという利点があり、情報の共有化という点において、研究者にとって大変都合が良い。しかし測定機器付属のパソコンには、データが機器独自の内部形式で保存されるため、このデータを他のパソコンで利用することができない。そこで赤外スペクトル測定装置のメーカーを中心に、どのコンピュータでも赤外スペクトルデータを扱えることを念頭に、JCAMP-DX フォーマットという共通のデータ保存形式が定められた。そこで本稿ではこのJCAMP-DX フォーマットの詳細について述べると共に、汎用パソコン上でこのデータを使用することを目的とした、JCAMP-DX フォーマットからのデータ復元プログラムの作成を試みた。そのためのソフトとしては、表計算ソフトの中で最も人気が高いと考えられるマイクロソフト社製Excelと、これに搭載されているマクロ言語であるVBAを使用した。また併せて、復元された赤外スペクトルデータを総合的に処理するプログラムの作成についても言及した。

1. はじめに

自然科学の研究は、まず対象物を観察することから始められる。この観察の過程においては、しばしば様々な測定機器が用いられ、対象物の持つ性質を計測する。センサーによって計測されたデータは電流や電圧などの電気信号に変換されるが、コンピュータが普及する以前は、この電気信号をメー

ターなどを使って目で読み取るか、あるいは付属するレコーダーで紙に記録し、そこからデータを読み取っていた。この際、人によって大なり小なりの誤差を生じる可能性がある。しかしこの誤差は測定とは何の関係もないものであり、これをできる限り最小限に押さえなければ、測定の努力が

*大妻女子大学 社会情報学部

**東京都立八王子工業高等学校定時制課程

報われないことになってしまう。

近年安価で高性能なパソコンが出回るようになり、各種測定装置にもパソコンが付属するようになってきた。この場合、電気信号に変換された情報はA/Dコンバーターによってデジタル化され、数値情報としてパソコンに記録・蓄積される。このように測定データがデジタル化されたことは、前に述べたデータ読み取りの際の誤差がなくなり、誰もが同じデータを共有出来るようになったことを意味しており、コンピュータのネットワーク化の時代を迎えたことと合わせて、特筆すべきことである。

二酸化炭素や窒素の酸化物など大気中の微量成分の濃度変動は、地球環境情報の一つとして重要である。これらの物質の濃度測定法の一つに赤外分光装置による赤外スペクトルの測定がある。赤外スペクトルの測定では、測定する波数範囲や分解能によっても異なるが、測定された波数とその波数での赤外線の吸収強度とを一組として、一回の測定において数千組のデータが得られる。これらを文字情報としてそのままディスクに保存するのは効率が悪いので、通常は内部フォーマットと呼ばれる、各測定機器メーカーが独自に定めたフォーマットでデータを圧縮して保存する。測定機器に付属しているパソコンには、スペクトルデータに関係する一連のソフトウェア（保存、表示、印刷、加工等）がインストールされており、それを使えば内部フォーマットによって保存されたデータを、一通り扱うことができる。しかし、付属のソフトウェアは必ずしも使用者のニーズを満たしているとは言えない。しかも、データは内部フォーマット形式で保存されているため、残念ながらこれを他のコンピュータでそのまま扱うことは不可能である。これでは、誰もが同じスペクトルを共有出来る、デジタル化されたスペクトルデータ情報の利点が全く活かされていない。

ここ数年パソコンの性能があがる一方で、DOS/V機の出現によりパソコンの価格は下がり、我々は以前よりも安い価格で、高性能のパソコンを手に入れることができるようになった。パソコンの高性能化に伴って、ワープロや表計算などの市販

のアプリケーションソフトも高機能化しており、これらのソフトが持つ機能を活用することによって、測定機器に付属するソフトと比べても優るとも劣らないデータ処理を行うことができる。しかしスペクトルデータの内部フォーマットという壁が存在するため、測定機器に付属するパソコン以外でスペクトルデータを利用することは容易ではない。従って、もし各社各様の内部フォーマットで保存されている赤外スペクトルのデータを、これらアプリケーションソフトで呼び出せる共通なフォーマットに変換できれば、デジタル化されたデータの利点を最大限に活かすことができる。

このような必要性から、JCAMP (Joint Committee on Atomic and Molecular Physical Data) という組織が赤外分光機器メーカーの援助のもと、機器メーカー、あるいはコンピュータの違いという壁を越えて、お互いに赤外スペクトルデータのやり取りができるようするための、スペクトルデータの外部フォーマットの基準 (JCAMP-DX) を定めた^{1), 2)}。このフォーマットは、もともと赤外スペクトルデータのために定められたものであるが、他のスペクトルデータ (ラマン、可視・紫外、核磁気共鳴³⁾、質量分析、粉末X線回折、ガス・液体クロマトグラフ等々) でも利用することができる他、化学構造を表現することも可能である (JCAMP-CSフォーマット)⁴⁾。これによって、JCAMP-DXフォーマットで保存されたスペクトルデータ情報は、呼び出し・保存のソフトさえ準備すれば、すべての測定機器、あるいはコンピュータで利用することが可能になった。このようにJCAMP-DXフォーマットの出現によって、本当の意味でのデジタル化されたスペクトルの共有化が実現したと言っても過言ではない。

本稿では、測定データ情報の普遍化・共通化という視点から、赤外スペクトルデータを対象としてJCAMP-DXフォーマットの内容について述べ、合わせてこの書式で記述されているデータを、汎用パソコン上の市販アプリケーションソフトで使用可能な書式に変換するための方法について述べる。さらにこれら変換後のデータを用いて、使用者の目的に合ったスペクトルやグラフ作

成等のデータ処理を便利に行うためのプログラム作成について言及する。なお市販アプリケーションソフトとしては、現在最も人気が高いと考えられる表計算ソフトMicrosoft Excel for Windows 95 Version 7.0 (以後Excelと略す)を使用した。

なお、本稿の基礎となる赤外スペクトル自身は、⁸⁾日本電子製フーリエ変換型赤外分光光度計JIR-6500型で測定した。しかし当然のことながら、ここで述べる内容はこの機種に固有ではなく、いかなる機種にも共通に適用可能である。

2. 表計算ソフトExcelと付属プログラミング言語VBAについて

JCAMP-DX フォーマットについて説明する前に、データ処理に利用する表計算ソフトExcelと、それに付属するプログラミング言語Visual Basic Programming System Application Edition (以後VBAと略す)について簡単に触れたい。

2.1 Excelについて^{5), 6)}

Excelは表計算ソフトとして分類されているが、実際には表計算以外にも、データのグラフ化、統計解析、データベース、シミュレーション演算、複素数や行列などの科学技術演算など、統合ソフトと言った方が正しいほど、豊富な機能を持っている。これらのうち、実際に赤外スペクトルデータを処理する場合には、スペクトルを描くためのグラフ描画機能や、スペクトル同士を演算するための演算機能を使うことが中心となる。これらの機能は表計算ソフトの最重要部分であり、特に充実した機能を持っている。またデータベース機能を活用することによって、蓄積される多数のスペクトルデータの整理も可能であり、高価である専用ソフトを除けば、表計算ソフトはパソコン上でのスペクトルデータの処理には最適のソフトと言うことができる。

2.2 VBAについて^{7), 8), 9)}

Excelは汎用ソフトであるため、使用者の行いたい作業がコマンドひとつで実行できるということはほとんどない。従って、Excelの持つ機能を組み

合わせて順次実行させていくことによって、使用者の意図する処理を行わせる必要がある。一般にスペクトルデータの処理においては、データが変わっても行う処理は基本的には変わらない。そこで毎回同じ操作を繰り返さなくても良いように、汎用ソフト(特に表計算)にはマクロ機能が搭載されている。マクロとは行いたい作業をソフトウェアに指示する命令の集まりで、ひとつの命令にはひとつの操作が対応する。つまりマクロは一種のプログラムであり、予めマクロを作成しておき、それを実行することによって、一連の操作を自動的に行うことができるようになる。ExcelのVer.5.0から搭載されたVBAは、Windows上でのBasicプログラミング言語として普及しているVisual Basic本来のコードと、Excelの操作に対するコードを併せ持った高機能プログラミング言語である。従ってこのVBAによってマクロプログラムを作成することにより、複雑な作業も効率的に実行することができるようになった。

JCAMP-DX フォーマットで保存されたスペクトルデータを数値に変換する場合、JCAMP-DX フォーマットに関する仕様は公表されているが、実際に変換を行うプログラムに関してはその機種ごとの様々な条件(OS, 使用機種等)が異なるため、各自で用意することになっている。最近の赤外分光装置には、付属するデータ処理装置に予めこの機能が搭載されている。一方、我々が通常使用するパソコン上でJCAMP-DXフォーマットのデータを取り扱うためには、そのプログラムを自分で組まなければならない。VBAはその目的に合ったプログラミング言語であり、またVBAによって作成されたプログラムは、Excelを使える環境ならば機種を問わず利用することができる。

3. JCAMP-DX フォーマットについて

3.1 JCAMP-DX フォーマットの設計思想¹⁾

JCAMP-DX フォーマットはこれまで述べたように、基本的には、赤外スペクトルとそれに関係する様々な情報を、使用機種の違いにとらわれずにやり取りするための標準フォーマットとして開発された。異機種間でも自由に、且つ精度が損な

われないようにデータをやり取りするために、実際の開発は次のようなコンセプトに基づいて行われた。

- (1)赤外スペクトル・インターフェログラム・フーリエ変換スペクトル・ピークテーブルなどを含んだ、分散型分光計・フーリエ変換型分光計・波長可変フィルター・レーザー分光計のいずれの機器からのデジタル出力も正確に表現できること。
- (2)赤外スペクトルのデータと同様に、使用したサンプルに関する情報（どのように調製された等）、測定条件などのパラメーター、計算の手順に関する情報も表現できること。
- (3)すべてのデータが目で読める形で表現されていること。
- (4)どのようなコンピュータ、通信体系、記録媒体にも受け入れられること。
- (5)各データフィールドのデータ長に拡張性があること。そして必要があれば、いつでも新しいデータフィールドが追加できること。
- (6)作成したJCAMP-DXファイルがいかなるシステムにおいても使用できるように、同ファイル作成のためのソフトウェアの仕様が正確に定められていること。

以上述べたコンセプトを十分に満たす形で、JCAMP-DX フォーマットが定められた。

3.2 JCAMP-DX フォーマットの構造¹⁾

ここでは、JCAMP-DX フォーマットバージョン4.24に基づいて、JCAMP-DX ファイルの構造とデータの格納方法について詳しく述べる。表1には、JCAMP-DX フォーマットの一形式であるDIF形式（3.3節参照）で保存された赤外スペクトルの実際のデータ例を示した。ここからの説明はこの表を参考に述べていく。なお表1の各行の左端に付いている行番号は、便宜上後から書き加えたものであり実際のデータには存在しない。また表1において、DIF形式で記述されたスペクトルデータは20～63行目に相当するが、参考として表1の20行目の一連のデータを、波数と透過率のデータに復元したものを表2に示した。

まずJCAMP-DX フォーマットでは、ファイルの基本要素はラベルの付いたデータレコード(Labeled-Data-Record, 以後LDRと略す)である。LDRが集まるとブロックを形成し、一つ、あるいはいくつかのブロックから一つのファイルが構成される。通常の赤外スペクトル測定では、一つのファイルは一つのブロックからなる。

一つのブロックは、'##TITLE=' (表1, 1行目) というLDRで始まり、'##END=' (表1, 64行目) で終わる。このブロックの中には、スペクトルデータを含めた測定に関するあらゆるデータがLDRとして入っている。予約語として'##'に続くラベルが多数用意されており(表3参照)¹⁾、これらの一部、あるいはすべてを使って測定に関する様々なデータ(以後スペクトル仕様データと呼ぶ)を書き込むことができる(表1, 1～19行目)。予約語にないものについては、ユーザー定義ラベルとして'\$'に続けて書き込むことができる。データの記述には、本フォーマットではASCII文字セットを使う約束になっているが、この中にはIBMの大型計算機で使われるEBCDIC文字セットにない文字がある。従って、データの互換性を保つため、記述にはそれら以外の文字を使わなければならない(表4参照)¹⁾。これらの文字のみを用いて表現されたデータは、すべてのコンピュータで扱えるばかりでなく、データ自身を画面あるいはプリンターで紙に印字した状態でも確認することが可能である。またコンピュータ間で記録媒体などに互換性がなくても、印字されたデータをOCRなど文字認識ソフトを用いて読み込ませることができる。

スペクトルデータ自身は、スペクトル仕様データの記述が終了後、データの構造を表すLDRである'##XYDATA=' (表1, 19行目)の次の行から記述される(表1, 20～63行目)。このデータの構造を表すLDRは'##XYDATA=(X++(Y..Y))'と記述されることが多く、各行先頭にまず1個のX座標の値が記述され、引き続いて数十個(1行80文字に収まる範囲)のY座標の値が圧縮された形で記述される。この先頭のX座標の値はすぐ右隣のY座標の値に対応し、それ以降のY座標

には、行頭のX座標の値に'##DELTA X='(表1, 10行目)というLDRに書かれている値を順次加算した値を持つX座標が対応する。なお赤外スペクトルでは、X座標には測定された波数(単位は cm^{-1} , カイザー), Y座標にはその波数における赤外線

透過率(単位は%)や吸光度(単位はない)などが対応するが、実際にこれらX, Y座標にどのような単位が使われたかは、それぞれ'##XUNITS='(表1, 11行目), '##YUNITS='(表1, 12行目)というLDRに明記されている。

表1. JCAMP-DX ファイル (DIF 形式で記録された赤外スペクトルデータ)

```

1: ##TITLE= CO ads on MgO
2: ##JCAMP-DX= 4.24
3: ##DATA TYPE= INFRARED SPECTRUM
4: ##ORIGIN= OTSUMA WOMEN'S UNIVERSITY
5: ##OWNER= T.ITO
6: ##DATE= 97/03/18
7: ##TIME= 16:32:29
8: ##SPECTROMETER/DATA SYSTEM= JEOL JIR series
9: ##RESOLUTION= 4.0
10: ##DELTA X= 1.928490E+00
11: ##XUNITS= 1/CM
12: ##YUNITS= TRANSMITTANCE
13: ##XFACTOR= 1.000000E+00
14: ##YFACTOR= 4.577637E-05
15: ##FIRSTX= 900.605
16: ##LASTX= 3999.689
17: ##NPOINTS= 1608
18: ##FIRSTY= 1.088355E+00
19: ##XYDATA= (X++(Y..Y))
20: 900.605B3775m4463p26N9j21k55K78j26O6N8j16M7k8j88J74J22o2j9r7J5J29J1KL6
21: 946.889B222714I5L1Jk1Rj0q5n2KK8N8Nn3L0j3n7m1NlrJ1M9J0o1ok5i2OL9L2k6oj5
22: 1014.386B2012J9K3j9i9LiK2K2qj1oj9j3KJ2li0j0%K0J6KPoJ3PL4nJ%KorRjJ7J5J7
23: 1089.597B2014nj4j2k3m0n7n0k6j2j5j8j4NRJ2Pkk4i1m0m1o0q3j07j29j28j01q0pJ7
24: 1147.452B0846o5pP4J02POL4Pk15m4p4q6q8p0m1jN8J06J20J20J15J14J18R8O2O0N4
25: 1199.521B1626L8LrK5N1L7K5J5m18k0J1L7L2JjJ1mklrnKMNqj7k3k3k9m1o0o4p9j05j27
26: 1268.947B1252j37j28r4k1N5J18J66J49P9J8k0n1q6j34j66j80j69j19j6Q8K02K69K55
27: 1313.302B1330J84J17P2M8K5J2J4K3J8Kj0Nmj4j9j2j5j0j1k4m4o2q5j01j46j64j12o3
28: 1367.300B0954r5j88j94i7j57K21J62Q1K2l8j16j52r6l2p1j82k16r2J03K74L11K50J92
29: 1411.655B1218J6J249J40J13O2K7L3K4J3MNPQJk%kIQN%onj4i%2mMmqokjKJNmj2rNJ
30: 1492.651B1897LnLNLPI4j4RJ2j0iORLnmjqk0J5Lj3r%RjnJ0MLj6J7mrjilPKLrkJJ
31: 1585.219B1899N9%LPkqrrKoPJokj0nj5Jrmj4oj6MmNll0Klrj4nj7qk4j5m1m4p2m4
32: 1672.001B1472p4o5j07j11j72k22i88m68o66p33q44i310j676p77P2K37i51o78j254j142
33: 1710.571A0933k72O75J055J008N57K40q3m87o40j99M76J034J045O64N11M21L53K97M41
34: 1747.212A8029O53N37K59J23J44J67J51K64L18K71J80R6M1K8J0J0qk4k7m2m4o4o8
35: 1793.496B0994p0n4m0k1m1o3r7j56k31j58P9K17J20nn8p4j86m29n97i21J57L85K81J25
36: 1839.780A9757N8r8j94K1K23K40J50J6N2q3j18P3J92R8P4J64J54J46R7n2j65j7j6j0
37: 1884.135B0613o0i7J05J84O2N8J19P4M6J54K47J48N3M1L3K0J5ML%OOPNlnKJlnJ0OPK
38: 1949.704B1918LolJlJo%JMLN%Lkp% % % % MnLLKjJ%O%JLlJKJMLlJ%LMnjKjKj%pMLmkj
39: 2061.556B1924%Jj%Omm%LlK0Olk%KjJlJlJmLk%lklLLMpq%Mml0kRlJlJlJm%kjoM
40: 2171.480B1901L0nj0J1N%KjllkjJkomRNjilKqrK0J0K%Onj3oknK%mnJ3J1MmJjinnM
41: 2271.761B1891JMK%NjJ3mnj0mlj4j4p5j5k3i5k5k8op4m9l9k2k0l49nJl0nJ3j3j6
42: 2343.116B1384lK0J0j8q1o9m1Kk8k7L7O6O2P5PO3O6N9M9M5K2RNQM%lkLlpLJomLNK
43: 2418.327B1890JkklLkMj%KKNKqjKjM%pLlNkKjmmmmKOJ0lKlnLMlK%KJmnlJmKojMjinnm
44: 2528.251B1881J0OjL0k%lK0kiojPmnlj%nj1LLLJl%LLMlJn%KM%kPppJ0QLkKmj5n
45: 2634.318B1897MPKj0lJ3lrjN%jil%LmLmj1qMQkpJomJ0mj0k0nj%LnrJ2Kj3nQLqm
46: 2732.671B1852K%K0OqokJ3poMKqjoo%JOJ0%j3j0lmlJ7NjKjJlNlMjJjJ3JrmQNO%J0P
47: 2831.024B1881mj3j1LJ3NnJlmKQqNj1lLkqoJ4Lj8j1MNLmMMLj%mjmr0OJkKqJ3Lo
48: 2927.448B1859okQmqoRj4j4j6KJ2%o%Olkjlnj2j4PnlrmlNlKnlLqMj2j2LJ5Pnp%J0
49: 3020.016B1831j3k0K5j1j%J1Jpml5jllJ4nj6KJ0QpoJ0ORoMjKpJ7j4J1J4Jp6pk3
50: 3102.941B1833Qj1lPomJ0loojOJ6RpnJ4%MKpikrQJ3P4lL7K6j5j0J7j1pqKojJ5j1j2
51: 3187.794B1866ON%K2J7nj1LlpJ2J4J0nkj1moJ8QoqQ0k2rK7j1Rk0KnmLJ8Jpmqk1LK8
52: 3272.648B1924J0J2mj0k2j2Mj4J2J0j1j6opOKj1qmlj8mKQ0J2j6onj8j2jMj6J3J0m
53: 3347.859B1852Lj8j9MQNj8j8Rj1Ml6j2J5KpNrk4J2Pj0k2m3k5j4oj3J0K5K6K2J7K1
54: 3417.285B1741k1Jj0j4j6j6K1J3nJ4j8k4MNj5Q6kmk0LmKmk4Oqrj0k9j7QL2K6j3k0r
55: 3486.710B1606J7Ok7j8J5K4J3k4r%2k0j1j8Rk1OJ3j0j0K7Nfj4l1pk4pQkJ5L8j1k3J0
56: 3554.208B1556M4kpLk0j4rMM5J3j3j6j2K1L4%n4MM5nOk0QM1k1lK3j3n6J0L6RmMn45
57: 3623.633B1551N6L1Rj4i7j9K3OJ7J4l1m2N9J4i3Kl9mJ5pnm0n0l%MJ8kMj9L1O0l0m7Nq
58: 3691.130B1572K9l0j0Qn8qm4j0k2k7L0N5K6k5m5j0K0P2j4J2%k7O3Mm9k9K3PJ7J0O
59: 3752.842B1639K9k2m7QN1j6l5nK0Qoj4K7M1Lj5oK9k4JL0rj95OL4j2rK3J6j4J5m1n2
60: 3818.411B1625PK2M4J2m5K3J4K9j6k8nK4kpk2L7L0k1J6M4n6j0M7pJ8j3j0P9K8n0i8
61: 3882.051B1761rMK5J9k1j3L0J5J1J6k9m7Nj5N7O7i7j02rQ2K9j3K2k2m7J9J8J6j6i2M9
62: 3941.834B1863J2m5J8J5K9N3m9p4k4M4M9K2Mk2j0mj8l%n%J8J2KpN3j1k0Kk0
63: 3999.689B1906
64: ##END=

```

表 2. JCAMP-DX ファイル (DIF 形式) とその復元後のデータ

DIF 形式で記述された JCAMP-DX ファイルのデータ (表 1, 20 行目)

900.605B3775m44l63p26N9j2lk55K78j26O6N8j|6M7k8j88J74J22o2j9r7J5J29J|KL6

その復元後のデータ

波数 (cm ⁻¹)	透過率 (%)
900.605	108.8333
902.533	106.8009
904.462	105.1392
906.390	101.8158
908.319	102.0859
910.247	101.5320
912.176	100.3647
914.104	101.6373
916.033	101.0605
917.961	101.3626
919.890	101.6281
921.818	101.0971
923.747	101.3123
925.675	101.1841
927.604	100.3235
929.532	101.1200
931.461	101.6785
933.389	101.3947
935.318	101.3077
937.246	100.8637
939.175	100.9323
941.103	101.5228
943.032	101.5732
944.960	101.5824
946.889	101.7471

表 3. JCAMP-DX ファイルの主な予約ラベル

##TITLE=	##FIRSTY=
##JCAMP-DX=	##RESOLUTION=
##DATA TYPE=	##DELTAX=
##BLOCKS=	##XYDATA=
##END=	##XYPOINTS=
##XUNITS=	##PEAK TABLE=
##YUNITS=	##PEAK ASSIGNMENTS=
##FIRSTX	##CLASS
##LASTX=	##ORIGIN=
##MAXX=	##OWNER=
##MINX=	##DATE=
##MAXY=	##TIME=
##MINY=	##SOURCE REFERENCE=
##XFACOR=	##CROSS REFERENCE=
##YFACOR=	##SPECTROMETER/DATA SYSTEM=
##NPOINTS=	##INSTRUMENT PARAMETERS=

本フォーマットの場合、スペクトルデータ（各 X, Y 座標値）は基本的には圧縮された形で表現される。以下にデータ圧縮の手順を示す。

- (1)赤外スペクトルのデータはすべて小数を含んでいるため、まずこれらのデータを共通の数字で割り算することによって、すべて整数に直す。この割り算に使われた共通の数字は、X 座標のデータの場合は '# # XFACTOR =' (表 1, 13行目), Y 座標のデータの場合には '# # YFACTOR =' (表 1, 14行目) という LDR に記述されている。なお X 座標については、本例のように圧縮を行わない (XFACTOR の値が 1) 場合もある。
- (2)(1)で加工された各データを、選択された保存形式 (3.3 節参照) によって、数字そのもの、あるいはデータの区切り・符号・数字自身を兼ねた擬数字を使って表現し直す (表 5 参照)¹⁾。なお擬数字は表現する数字の先頭桁の

みで使い、表現する数字が 2 桁以上になるときは 2 桁目以降は通常の数値を使う。

3.3 JCAMP-DX フォーマットによる保存形式の種類^{1), 10)}

スペクトルデータの保存形式には具体的には以下の方法があり、利用者は自由に選択することができる (表 6 参照)¹⁾。

(1)FIX 形式

整数化された数値をそのまま使い、桁の多少に関わらずデータを一定の間隔 (7 桁) ごとに揃えて並べる。

(2)PAC 形式

整数化された数値をそのまま使い、データとデータの間を符号 (+, -) かスペース一つで区切って並べる。

(3)SQZ 形式

整数化された数値の先頭桁を擬数字に置き

表 4. JCAMP-DX ファイル内で使用される ASCII 文字*

!"#\$%&'()*+,-./0123456789:;<=>?
@ABCDEFGHIJKLMNPQUSTUVWXYZ^_`
abcdefghijklmnopqrstuvwxyz

* [*] '{ } ~ の 7 文字は EBCDIC 文字セットにないため使えない。

表 5. 数字とそれに対応する擬数字

実際の ASCII 数字	0	1	2	3	4	5	6	7	8	9
プラスの SQZ 形式	@	A	B	C	D	E	F	G	H	I
マイナスの SQZ 形式*		a	b	c	d	e	f	g	h	i
プラスの DIF 形式	%	J	K	L	M	N	O	P	Q	R
マイナスの DIF 形式*		j	k	l	m	n	o	p	q	r
プラスの DUP 形式		S	T	U	V	W	X	Y	Z	s

*マイナスの SQZ、DIF 形式の場合擬数字はその数字とマイナスの符号の両方を表す (例: c→-3、p→-7)。

表 6. 実際のデータとそれぞれの形式による表現の例

実際のデータ	1, 2, 3, 3, 2, 1, 0, -1, -2, -3, -25
FIX 形式	1 2 3 3 2 1 0 -1 -2 -3 -25
PAC 形式	1+2+3+3+2+1+0-1-2-3-25
(または)	1 2 3 3 2 1 0-1-2-3-25
SQZ 形式	1BCCBA@abc5
DIF 形式	1JJ%jjjjjk2
DIFDUP 形式	1JT%jXk2

換え、それを隙間なく並べる。

(4)DIF形式

整数化された数値の隣同士の差をとり、この差の先頭桁の数字を擬数字に置き換え、それを隙間なく並べる。なおDIF形式の場合、確認のために行末のY座標の値と次の行の行頭のY座標の値が等しくなっており、かつ行頭のY座標のみはSQZ形式で記述されている。

(5)DUP形式

SQZ形式とDIF形式においては、同じ値のデータが続く場合は繰り返しの擬数字を使ってそれらをまとめて表現する、SQZDUP形式とDIFDUP形式も存在する。

表6¹⁾から分かるように、これらの保存形式の中ではDIFDUP形式で保存されたファイルのサイズが一番小さくなる。

4. JCAMP - DX フォーマットからのデータの復元

3.で述べたJCAMP-DXフォーマットについては、保存形式さえ選べばあとは赤外分光装置付属のパソコンが自動的にその形式でスペクトルデータを保存してくれる。これによって生成されたファイルを我々が通常使用する汎用パソコンで利用する場合、測定に関する情報については各LDRから文字として直接情報を読み取ることができるが、スペクトルデータについては、圧縮されたデータを復元しなければならない。本稿では3.3で述べた表現形式のうち、当研究室で通常使用しているDIF形式について、Excelのマクロ言語であるVBAを使って実際にデータを復元する方法について述べる。他の形式からのデータの復元についても、表5¹⁾の変換テーブルに基づいて、このプログラムの一部を書き換えることによって対応することができる。

まず、データ復元のアルゴリズムは、大まかに分けて次の6つの過程から構成される。

(1)目的のファイルのオープン

(2)スペクトル仕様データの読み取り

(3)スペクトルデータの読み取り・変換一行頭のX, Y座標値

(4)スペクトルデータの読み取り・変換一行頭以外のX, Y座標値

(最終行まで(3), (4)の処理を繰り返す)

(5)Y座標が透過率の場合の変換

(6)ファイルのクローズ

実際のプログラムを表7に示したので、表7のプログラムに沿って上記の6つの過程を各々簡単に補足説明する。

・前準備：2～20行目

ファイルのオープンに先立ち、プログラムで使用する変数とその型宣言等を行う。なおここでは、スペクトルデータ(X, Y座標)の収容個数を2000個とした(14, 15行目)。

(1)目的のファイルのオープン：22, 24行目

画面でデータを復元したいファイルを選択させ、そのファイルをオープンする。

(2)スペクトル仕様データの読み取り：25～46行目

スペクトルデータ自身より前に記述されているLDRから、スペクトル仕様データ(スペクトルに関連する様々な情報)を読み込む。なお、以下のスペクトルデータの復元処理に最低限必要な変数の内容は、以下の通りである。

①DeltaX：X座標の増し分

②YUnits：測定が透過率か吸光度のどちらで行われたか

③YFactor：Y座標の値を復元する際に掛ける数 (※XFactorが1でない場合は、XFactorも必要)

④NPoints：全データ数

続いて50行目のDo文と138行目のLoop文の間でスペクトルデータの具体的復元作業を行うが、行頭のX, Y座標の読み取り(51～95行目)と行頭以外のX, Y座標値の読み取り(97～136行目)の2段階に分けて処理することが必要である。

(3)スペクトルデータの読み取り・変換一行頭のX, Y座標値：51～95行目

スペクトルデータを1行分読み取り、行頭から1文字ずつ調べていき、数字と小数点以外の文字が初めて現れる文字(場所)を見つ

表 7. DIF 形式から数値へのデータ復元プログラム

```

1:Sub データ復元()
2:  Dim N As Integer
3:  Dim K As Integer
4:  Dim NPoints As Integer
5:  Dim NChar As Integer
6:  Dim NPos As Integer
7:  Dim NPosBef As Integer
8:  Dim LChar As Integer
9:
10:  Dim DeltaX As Single
11:  Dim YFactor As Single
12:  Dim X As Single
13:  Dim Y As Single
14:  Dim XData(2000) As Single
15:  Dim YData(2000) As Single
16:
17:  Dim FName As String
18:  Dim FileAtr As String
19:  Dim CharData As String
20:  Dim CharDataP As String
21:
22:  FName = Application.GetOpenFilename("IR 測定データ ファイル(*.dx), *.dx", , "ファイルの選択")
23:
24:  Open FName For Input As #1
25:  Do
26:      Line Input #1, CharData
27:      Select Case Mid(CharData, 3, 6)
28:          Case "DELTAx"
29:              DeltaX = Val(Right(CharData, 12))
30:          Case "YUNITS"
31:              CharData = Right(CharData, Len(CharData) - 10)
32:              CharData = Left(CharData, 3)
33:              Select Case CharData
34:                  Case "TRA"
35:                      FileAtr = "T% "
36:                  Case "ABS"
37:                      FileAtr = "ABS"
38:              End Select
39:          Case "YFACTO"
40:              YFactor = Val(Right(CharData, 12))
41:          Case "NPOINT"
42:              NPoints = Val(Right(CharData, 4))
43:          Case "XYDATA"
44:              CharData = "END"
45:          End Select
46:      Loop While CharData <> "END"
47:
48:      N = NPoints
49:
50:      Do '----- <Outer Loop>
51:          Line Input #1, CharData
52:          NChar = Len(CharData)
53:          NPos = 0
54:
55:          Do
56:              NPos = NPos + 1
57:              Loop While InStr(1, "1234567890.", Mid(CharData, NPos, 1), 0) <> 0
58:
59:              X = Val(Left(CharData, NPos - 1))
60:
61:              NPosBef = NPos
62:
63:              Do
64:                  NPos = NPos + 1
65:                  If NPos > NChar Then
66:                      Exit Do
67:                  End If
68:              Loop While InStr(1, "1234567890.", Mid(CharData, NPos, 1), 0) <> 0
69:
70:              CharDataP = Mid(CharData, NPosBef, NPos - NPosBef)
71:              LChar = InStr(1, "@ABCDEFGHIIabcdefghi", Left(CharDataP, 1), 0)
72:
73:              Select Case LChar
74:                  Case 1 To 10
75:                      CharDataP = Chr(&h2f + LChar) + Right(CharDataP, Len(CharDataP) - 1)
76:                  Case 11 To 19
77:                      CharDataP = "-" & Chr(&h30 + LChar) + Right(CharDataP, Len(CharDataP) - 1)

```

```

78:      End Select
79:
80:      Y = Val(CharDataP) * YFactor
81:
82:      Select Case N
83:      Case NPoints
84:          XData(N) = X
85:          YData(N) = Y
86:          N = N - 1
87:      Case 0
88:          XData(1) = X
89:          YData(1) = Y
90:          N = N - 1
91:          GoTo EndP
92:      Case Else
93:          XData(N + 1) = X
94:          YData(N + 1) = Y
95:      End Select
96:
97:      Do '----- <Inner Loop>
98:          NPosBef = NPos
99:
100:         Do
101:             NPos = NPos + 1
102:             If NPos > NChar Then
103:                 Exit Do
104:             End If
105:             Loop While InStr(1, "1234567890.", Mid(CharData, NPos, 1), 0) <> 0
106:
107:             CharDataP = Mid(CharData, NPosBef, NPos - NPosBef)
108:             LChar = InStr(1, "JKLMNOPQRjklmnopqr%", Left(CharDataP, 1), 0)
109:
110:             Select Case LChar
111:             Case 1 To 9
112:                 Select Case Len(CharDataP)
113:                 Case 1
114:                     CharDataP = Chr(&h30 + LChar)
115:                 Case Else
116:                     CharDataP = Chr(&h30 + LChar) + Right(CharDataP, Len(CharDataP) - 1)
117:                 End Select
118:             Case 10 To 18
119:                 Select Case Len(CharDataP)
120:                 Case 1
121:                     CharDataP = "-" & Chr(&h30 + LChar - 9)
122:                 Case Else
123:                     CharDataP = "-" & Chr(&h30 + LChar - 9) +
124:                                     Right(CharDataP, Len(CharDataP) - 1)
125:                 End Select
126:             Case 19
127:                 CharDataP = "0"
128:             End Select
129:
130:             Y = Val(CharDataP) * YFactor
131:             XData(N) = XData(N + 1) + DeltaX
132:             YData(N) = YData(N + 1) + Y
133:
134:             N = N - 1
135:
136:             Loop While NPos <= NChar '----- <Inner Loop>
137:         EndP:
138:         Loop While N >= 0 '----- <Outer Loop>
139:
140:         If FileAtr = "T%" Then
141:             For K = 1 To NPoints
142:                 YData(K) = YData(K) * 100
143:             Next K
144:         End If
145:
146:         Close #1
147:
148:     End Sub

```

ける。行頭からこの文字の左隣までに出現した数値列を、この行の最初の X 座標の値として取り込む (55~59行目)。

この X 座標の右隣の文字から行末までの文字列が、この行に順次格納されている Y 座標の値となる。ただし、各行の最初の Y 座標の値のみは DIF 形式ではなく SQZ 形式で格納されているので、最初にこの Y 座標の値のみを、3.3 節で述べた SQZ 形式の保存形式と表 5¹⁾にある擬数字対応表に基づいて数値に変換する。この変換に際して本稿では、数字の文字コードの連続性を利用した。数字 0~9 には、文字コード 48~57 (16進法では &H30~&H39) が連続して割り当てられている。そこで数字の 0 から 9 に対応する擬数字をその順番で並べた文字列を作り、調べたい文字がこの中の何番目の文字と一致するかを調べる (70, 71行目)。そして何番目と一致したかという数を、数字 0 の文字コード (48, 16進法で &H30) より一つ少ない 47 (16進法で &H 2 F) に加えたものが、目指す数字の文字コードになる (75行目)。そしてこの文字コードから数字を復元し、この復元された値に YFactor の値を掛けることによって、各行最初の Y 座標の値が得られる (80行目)。

こうして得られた各行先頭の X, Y 座標の値を配列変数に格納する (82~95行目)。そして処理している行がこのスペクトルデータの最終行でなければ次の(4)の処理に進み、最終行の場合は(5)の処理に進む。

(4) スペクトルデータの読み取り・変換—行頭以外の X, Y 座標値 :97~136行目

行頭の X, Y 座標の後には、行末まで Y 座標の値が前の Y 座標の値との差として DIF 形式で格納されている。そこでまず DIF 形式の Y 座標を取り出して表 5¹⁾にある擬数字対応表に基づいて数値に変換する。この変換には(3)と同様に、文字コードを利用した。こうして得られた値に YFactor の値を掛け、これをさらに前の Y 座標の値に加えることによって真の Y 座標の値が得られる。これに対応する X

座標の値は、前の X 座標の値に DeltaX の値を加えることによって得られる。

そしてこの処理を行末まで行ったら、(3)に戻って操作を繰り返す (138行目)。

(5) Y 座標が透過率の場合の変換

このスペクトルデータの Y 座標の値が透過率として格納されている場合は、Y 座標の値を 100倍して % に変換する (140~144行目)。

(6) ファイルのクローズ

ファイルを閉じて、処理を終了する。

以上の処理によって変換された X 座標, Y 座標の値はそれぞれ配列 XData, YData に格納されたが、このデータを使って Excel 上で実際に赤外スペクトルを描かせる場合には、配列中のデータをシート上のセルに書き込まなければならない。その場合はファイルをクローズした後に (146行と148行目の間) 以下のような文を加えればよい。これは Sheet 1 の左側の 2 列 (A と B 列) に、1 行目から順番に配列のデータを書き込む命令である。

For K=1 to Npoints

Sheets ("Sheet1"). Cells (K, 1) = XData (K)

Sheets ("Sheet1"). Cells (K, 2) = YData (K)

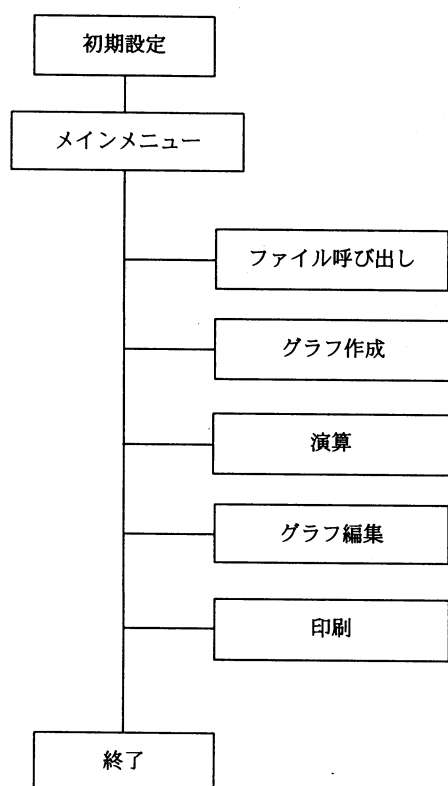
Next K

シート上のセルに書き込まれたデータを使ってスペクトル (グラフ) を描かせたりする事に関しては、Excel 自身の操作に関するたくさんの本が出版されているので、そちらを参照されたい。^{5) 6)}

5. 赤外スペクトル処理用汎用プログラム¹⁾

表 7 に示したプログラムによって、JCAMP-DX 形式で圧縮・保存された赤外スペクトルデータを、誰でもが使用可能な共有データに復元することが出来た。そこで一つの応用例として、この復元データを汎用パソコン上で自由に使用するために、Excel 上で赤外スペクトルの演算やスペクトル図の作成を効率よく行える汎用プログラムの作成を試みた。使用言語はデータ復元のときと同じ VBA とし、実際にはデータ復元部とそれ以外のデータ処理部とを合わせて 1 個のプログラムとなっている。図 1¹⁾ にプログラム全体の大まかなフローチャートを示

図1. 赤外スペクトル処理用汎用プログラム全体のフローチャート



したので、これに沿ってプログラムの構成を簡単に説明する。

(1)初期設定

プログラムで使用する変数の宣言，スプレッドシート・グラフシートの準備と初期化，等を行う。設定終了後は，メインメニューの画面を表示させる。

(2)メインメニュー

ここではメインメニュー画面を表示し，そこから以下の①～⑤，及び終了の処理を選択させる。①～⑤の処理を選んだ場合，処理終了後はここに戻って再びメインメニュー画面を表示する。

①ファイル呼び出し

表7に示したデータ復元のプログラムがここに相当し，スペクトルデータファイルの選択，データ復元，及び復元されたデータの

スプレッドシートへの書き込みを行う。

②グラフ作成

スプレッドシートに書き込まれたデータから，グラフを作成してグラフシートに表示させる。

③演算

赤外スペクトルの処理に必要な，スペクトル同士の四則演算，スペクトルのベースラインの補正及び移動，透過率と吸光度の変換，等の処理を選択させて行う。

④グラフ編集

表示したグラフを体裁よくするために必要な，グラフの大きさの設定，X・Y軸の設定，凡例の表示の有無，文字や矢印などの書き込み，等の処理を選択させて行う。

⑤印刷

表示されたグラフシートを印刷する。

(3)終了

作成したグラフシートを保存し，処理を終了する。

このプログラムの作成によって我々は，それまで赤外分光装置付属のパソコンでお仕着せの方法でしかできなかったデータ処理を，手元のパソコンで自由に，しかも半自動的に行えるようになった。特に印刷にレーザープリンターが使えるようになったことで，格段にきれいなスペクトル図を出力することができるようになった。

6. 終わりに

コンピュータの普及は情報のデジタル化を推し進め，様々な情報がデジタル化されて蓄積されている。そしてインターネットの拡大は，それらの蓄積された情報への手軽なアクセスを，世界的な規模で可能にしつつある。本稿で述べたJCAMP-DXフォーマットは，それぞれで扱っている機種の違いを超えて，デジタル化されたスペクトルデータ情報の普遍化・共有化を可能にした。但し一方では，スペクトルデータ情報の処理には基本的には分光装置付属のパソコンのみでも用が足りてしまうため，JCAMP-DXフォーマット自身がそれほど普及しているとはいえない状況もある。しか

し本来スペクトル情報の共有化と相互利用は、科学の進歩と発展には必要不可欠の要素である。現在、高性能パソコンの普及、インターネットを含めたネットワークの拡大によって、情報の共有化に関する条件は整い、同時に関心も高まりつつある。JCAMP-DX フォーマットの利用者が増えて、スペクトルデータ情報を更に有効に活用できるよう、本稿がお役に立てば幸いである。

参考文献

- 1) R.S.McDonald et al., JCAMP-DX: A Standard Form for Exchange of Infrared Spectra in Computer Readable Form, Appl.Spectrosc., 42, 151-162, 1988.
- 2) 宮田 壽, 触媒, 36, 352-356, 1994.
- 3) A.N.Davies et al., JCAMP-DX for NMR, Appl. Spectrosc., 47, 1093-1099, 1993.
- 4) J.Gasteiger et al., JCAMP-CS: A Standard Exchange Format for Chemical Structure Information in Computer-Readable Form, Appl. Spectrosc., 45, 4-11, 1991.
- 5) 島谷明男, Excel95操作ハンドブック, ナツメ社, 1996.
- 6) 佐藤謙一ら, できる Excel 95 Windows版 改訂版, インプレス, 1997.
- 7) 横井与次郎, Excel 95 VBA プログラミング入門, ソフトバンク, 1996.
- 8) 藤森洋志ら, EXCEL VBA スタandard プログラミング, オーム社, 1995.
- 9) 相沢文雄, Excel 95 VBA マクロハンドブック, ナツメ社, 1996.
- 10) JIR-6500 JX + ソフトウェアシステム, 日本電子, 10-7.
- 11) 鹿野比呂子, Visual Basic 言語による赤外分光用データ処理プログラムの作成, 大妻女子大学社会情報学部平成8年度卒業研究論文, 1996.

Exchange of spectral data information —Utilization of infrared spectral data in standard form on a personal computer—

TOMOYASU ITO*, TOSHIHIKO TASHIRO** and HIROKO KANO*

* *School of Social Information Studies, Otsuma Women's University*

** *Tokyo Metropolitan Hachioji Technical High School*

Abstract

Personal computers are currently going to be incorporated into many instruments used in natural science and technology. Digitalized data stored in a computer are very useful from a stand point of exchange of measured data among many persons since digitalized data can be used without loss of exactness. However, these digitalized data are usually stored in an encoded form specific to each instrument. This prevents ones from exchanging their data and from utilizing them on his own computer. The JCAMP-DX format is a standard encoding form for infrared spectral data and makes it possible to exchange exact spectral data information on all kinds of computers.

In this paper, firstly, the concept of the JCAMP-DX format is explained, and secondly, the idea for decoding infrared spectral data recorded in the JCAMP-DX format is shown and a sample program is prepared by VBA on Microsoft Excel to realize this decoding. Lastly, a total program to treat infrared spectrum using these decoded data is outlined.

Key Words (キーワード)

standard data file (標準データファイル), infrared spectrum (赤外スペクトル), personal computer (パーソナルコンピュータ), JCAMP-DX format (ジェイキャンブーディーエックス書式), data exchange (データ交換), Microsoft Excel (マイクロソフトエクセル), VBA (ヴィビーエイ・マクロ言語)