

Qualitative Phase Analysis KVALITATIVNÍ FÁZOVÁ ANALÝSA

Jaroslav Fiala a Michal Kolega

Nové technologie - Výzkumné centrum, Západočeská univerzita v Plzni

Keywords

phase analysis, powder diffraction

Abstract

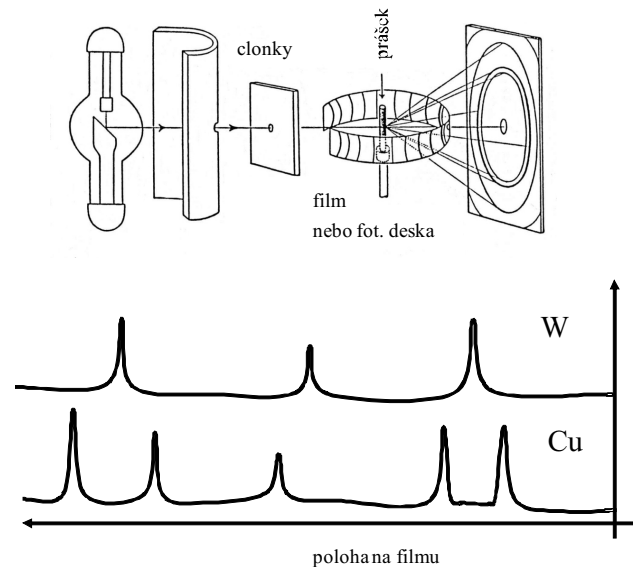
The determination of phase composition is one of the most frequently used applications of X-ray diffraction. It is based on comparison of powder diffraction pattern of analysed compound with the database of reference diffraction patterns. Ordering of the database and the way how it is used for the determination of phase composition determine the efficiency of qualitative phase analysis. Methodic development of this discipline can be based on the mathematical model of X-ray diffraction identification which is build on the basis of linear algebra. This model also theoretically explains the reasons of practical difficulties encountered at the analysis of phase composition of multi-component mixtures. And it is also a base for overcoming of these difficulties and successful identification of complicated mixtures by factor analysis of their diffraction patterns.

Abstrakt

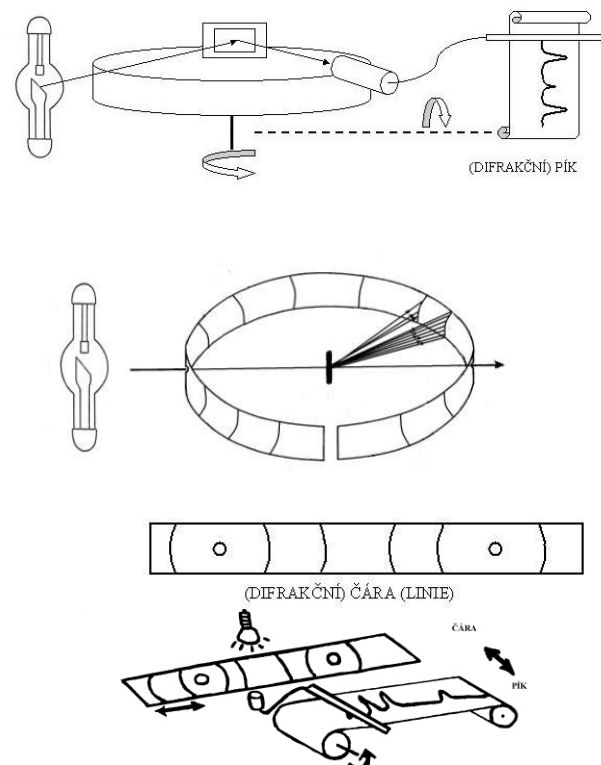
Určování fázového složení je jednou z nejvíce využívaných aplikací rtg difrakce. Vychází z porovnání difraktogramu analysované látky s databází referenčních difraktogramů. Uspořádání té databáze a způsob, jakým se s ní při určování fázového složení pracuje, určují efektivnost kvalitativní fázové analýsy. Předpokladem metodického rozvoje této disciplíny je matematický model rtg difrakční identifikace vybudovaný na bási lineární algebry. Tento model také teoreticky vysvětluje příčiny praktických potíží, na které narážíme při rozboru fázového složení mnohokomponentních směsí. A je východiskem pro překonání těch potíží a úspěšnou identifikaci složitých směsí faktorovou analýsou jejich rtg difraktogramů.

1. Úvod

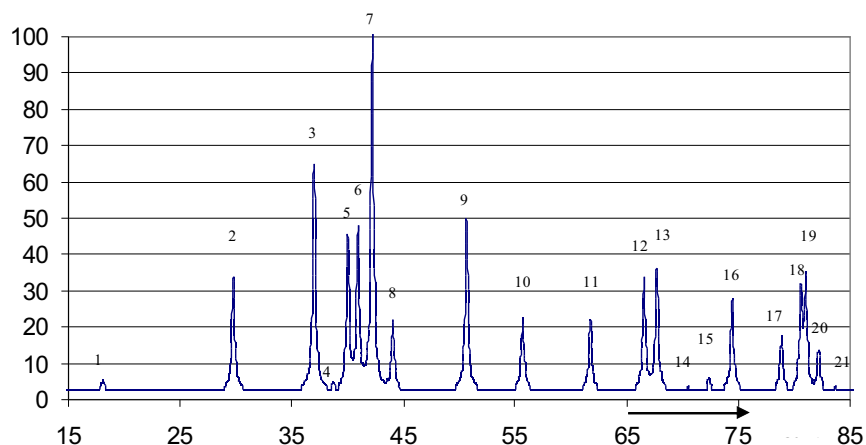
Rtg záření procházející polykrystalickým materiálem se rozptyluje (difraktuje) a směrová distribuce intenzity difraktovaného záření, registrovaná jako t.z. difraktogram (angl. diffraction pattern) nese informaci o fázovém složení tohoto materiálu [1-3] – obr. 1. Vedle toho záleží vzhled difraktogramu samozřejmě také na experimentálním uspořádání: tak při registraci na film sestává difraktogram z difrakčních čar (linií), jejichž fotometrováním dostáváme píky (maxima) zčernání; ty pak odpovídají píčkám směrové distribuce intenzity difraktovaného záření, měřené při elektronické detekci přímo (obr. 2). Při jednom a tomtéž experimentálním uspořádání je však vztah mezi vzhledem difraktogramu daného materiálu a jeho fázovým složením (až na isomorfní látky, což je zvláštní případ, o kterém se zmíníme v dalším) vzájemně jednoznačný: různým látkám odpovídají různé difraktogramy, které jsou



Obrázek 1. Rtg záření se průchodem skrze zkoumaný materiál ohýbá (difraktuje) a směrová distribuce intenzity difraktovaného záření, registrovaná například na filmu jako tzv. difraktogram, nese informaci o fázovém složení toho materiálu.



Obrázek 2. Rtg difraktogram se registruje buď přímo, elektronicky (a), nebo na film (b), který se po vyvolání fotometruje (c) za účelem zjištění plošné distribuce zčernání použitého filmu.



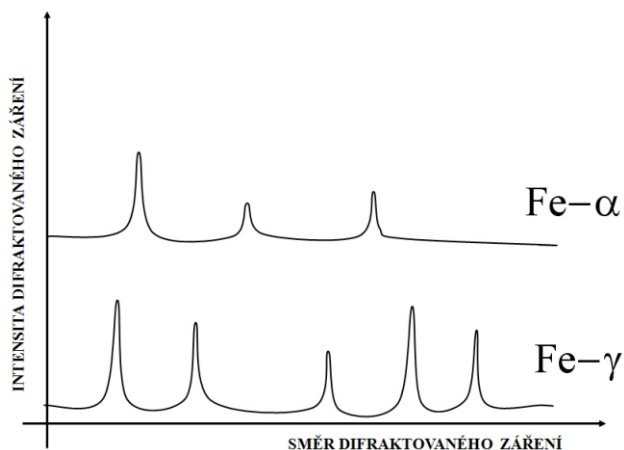
číslo piku	d[A]	I/I ₁	ZnO		Al ₂ O ₃	
			d[A]	I/I ₁	d[A]	I/I ₁
1	5.68	2.5	-	-	-	-
2	3.48	31	-	-	3.48	70
3	2.814	62	2.814	57	-	-
4	2.698	2	-	-	-	-
5	2.612	42	2.603	44	-	-
6	2.557	43	-	-	2.551	97
7	2.482	100	2.476	100	-	-
8	2.384	19	-	-	2.379	42
9	2.089	47	-	-	2.085	100
10	1.914	20	1.911	23	-	-
11	1.743	19	-	-	1.740	42
12	1.629	31	1.625	32	-	-
13	1.605	33	-	-	1.601	82
14	1.549	1	-	-	1.546	2
15	1.514	3	-	-	1.511	7
16	1.478	25	1.477	29	-	-
17	1.408	15	1.407	4	1.405	30
18	1.382	26	1.378	23	-	-
19	1.376	30	-	-	1.374	45
20	1.360	11	1.358	11	-	-
21	1.340	1	-	-	1.336	1

Obrázek 6. Difraktogram směsi korundu (- Al₂O₃) a zinkitu (ZnO) spolu s tabulkou, kde je vyznačeno, které difrakční linie odpovídají korundu a které zinkitu. Záření CoK .

difraktogram nese informaci o poloze atomů, jinými slovy je na geometrii, rozložení atomů ve struktuře citlivý. Proto se difraktogramy různých modifikací látek, které mají stejné prvkové složení, liší (obr. 7). Různé polymorfní modifikace látek můžeme tedy rtg difrakcí od sebe dobře rozpoznat. To je velmi užitečné, nebo polymorfie jest mezi materiály hojně rozšířena [9].

Na druhé straně existují páry materiálů, které označujeme jako isomorfní: jsou to látky, které mají různé prvkové složení, ale geometrické rozložení jejich atomů je shodné (např. BN a C nebo AlPO₄ a SiO₂). Protože vzhled difraktogramu závisí na geometrickém uspořádání atomů mnohem více než na tom, které atomy jsou v příslušných polohách rozmístěny, nemůžeme většinou isomorfní látky pomocí rtg difrakce rozpoznat. To je jistě nedostatkem rtg difrakční techniky fázové analýsy, ale tento nedostatek není zas až tak závažný, protože se dá

snadno obejít, když určíme prvkové složení příslušného preparátu např. rtg fluorescenční spektrální analýsou [10]. Mnohem závažnější je to omezení rtg difrakce, které spočívá ve skutečnosti, že touto technikou můžeme určovat pouze látky krystalické (na rozdíl např. od infračervené spektroskopie nebo diferenciální termické analýsy). S tímto nedostatkem musíme počítat a uvědomit si, že provedeme-li pomocí rtg difrakce rozbor fázového složení daného materiálu, dostaneme tím způsobem informaci pouze o jeho krystalických komponentách. Vedle nich však ten materiál může obsahovat ještě několik dalších fází, které jsou amorfní a jejichž podíl může být velký. Tak třeba aluminosilikátové sklo představuje 40-50 % objemového podílu čediče ale na rentgenogramu ho „nevidíme“, protože veškeré difrakční linie na difraktogramu čediče pocházejí pouze od jeho krystalických složek (většinou jsou to pyroxen augit a minoritní magnetit).



Obrázek 7. Vzhled difraktogramu závisí mnohem více na geometrii (prostorové distribuci) atomů v analysovaném krystalu, než na jejich chemii (na tom, o jaké se jedná atomy). Pomocí rtg difrakce můžeme proto dobře rozlišit jednotlivé polymorfní modifikace látek, které mají totéž prvkové složení.

2. Database referenčních difraktogramů

K tomu abychom mohli látky pomocí rtg difrakce identifikovat, potřebujeme databasi referenčních difraktogramů. Tu si tvoří každá laboratoř jako sbírku difraktogramů látek svého analytického programu. Traduje se, že první takovou sbírkou (databasí) byl soubor difraktogramů 51 minerálů vyskytujících se v měděných rudách, které posbíral a r. 1935 publikoval A. W. Waldo [11]. Jinou takovou známou databasí z pionýrské éry rtg difrakční kvalitativní fázové analýzy byla sbírka difraktogramů 142 minerálů sestavená a publikovaná r. 1939 A. K. Boldyrevem et al. [12]. Při úzkém a málo proměnném analytickém programu specializované sbírky referenčních difraktogramů, které si tvoří jednotlivé uživatelské laboratoře „ke svému obrazu“, dobře slouží. Obsahují zpravidla několik desítek až několik stovek referenčních difraktogramů, jak to může ilustrovat příklad database 298 difraktogramů vytvořené pro potřeby Dow Chemical Company ve státě Michigan, USA, která se tam velmi dobře osvědčila a dlouhá léta byla ke spokojenosti využívána [13] nebo database asi 350 difraktogramů, kterou si vytvořil a intenzivně používal pisatel těchto řádek během své 37-leté praxe v rentgenografické laboratoři plzeňské Škodovky.

Pro účely laboratoří, které mají široký a/nebo proměnný analytický program, bude ovšem užitečné mít databasi referenčních difraktogramů jež by byla co největší, s co nejširším pokryvem a soustavně aktualizovaná. Toto povědomí vzniklo již před druhou světovou válkou a vykristalovalo v úsilí o založení instituce, jež by takovou databasi vytvořila a nadále obhospodařovala. Jako první vzala na sebe tento úkol „národní“ (federální) organizace Spojených států amerických – ASTM (American Society for Testing and Materials); to bylo v roce 1941. Později se připojily další organizace a r. 1969 byla jako zastrešující instituce zabývající se obhospodařováním univerzální database referenčních difraktogramů pro laboratoře na celém světě ustavena Joint Committee on Powder Diffraction Standards (JCPDS) jako nezisková

organizace státu Pennsylvania. A v roce 1978 byla pak tato organizace přejmenována na International Centre for Diffraction Data (ICDD) což má lépe vyjadřovat její globální poslání v rámci mezinárodní vědecké komunity. I když dvakrát změnila jméno, duch této instituce se v průběhu let nezměnil: se svými necelými třemi desítkami zaměstnanců a 280 vědeckými spolupracovníky z 38 zemí světa, ICDD sbírá, edituje a publikuje referenční difrakční data pro identifikaci látek, přičemž dbá o úplnost a vysokou kvalitu těchto dat a podporuje rozvoj identifikační metodiky a jejich aplikací. V současnosti obsahuje database referenčních difraktogramů ICDD – Powder Diffraction File – skoro 350 000 difraktogramů a je ji možno objednat na adrese International Centre for Diffraction Data, 12 Campus Boulevard, Newtown Square, Pennsylvania 19073 – 3273 U.S.A.; tel. 610 325 9814; fax 610 325 9823; e-mail: info@icdd.com; Web: www.icdd.com.

Stojí za zmínku, že základem této database se stala sbírka jednoho tisíce difraktogramů látek, o jejichž identifikaci pomocí rtg difrakce měla zájem průmyslová společnost Dow Chemical Company. Ta tuto sbírku difrakčních údajů publikovala r. 1938 [14] a v r. 1941 ji dovolila reprintovat ve formě kartotéčních lístků jako první vydání, tedy úplný počátek database Powder Diffraction File. V parafrázi k výroku L.J.Hendersona z roku 1917, že totiž věda vděčí parnímu stroji za víc, než zač vděčí parní stroj vědě [15], můžeme v této souvislosti konstatovat, že rtg difrakční kvalitativní fázová analýza vděčí průmyslu za víc, než zač vděčí průmysl rtg difrakční kvalitativní fázové analýze... .

První vydání database referenčních difraktogramů PDF bylo tedy realizováno v podobě kartotéčních lístků (obr.8) což je praktické, nebo takovou sbírku dat můžeme snadno rozšiřovat a přeorganizovávat. Aby se v ní dalo i dobře hledat (a aplikovat tehdy hojně používanou mechanizaci hromadného zpracování dat pomocí děroštitkových analyzátorů), přešlo se záhy na obvodově děrované kartotéční listky (edge codeable cards) [16]. A později se database začala vydávat také ještě v knižní podobě, na mikrofiších, magnetických páskách, magnetických discích [4] a posléze na optických discích [17], což je nejmmodernější velice efektivní a dnes snad nejpoužívanější forma.

V průběhu let se měnila nejen forma nosiče, na němž je database ICDD – Powder Diffraction File distribuována, ale také způsob, jakým jsou referenční difrakční údaje získávány. Vyvíjí a zdokonaluje se experimentální technika rtg práškové difraktometrie a to samozřejmě ovlivňuje také kvalitu difrakčních dat. Data, jež jsou ukládána do database referenčních difraktogramů dnes se svou kvalitou, přesností, úplností ale i správností liší od dat, která byla do database uložena před padesáti lety. Tak třeba přesnost, jež se dnes vyžaduje při určování polohy difrakčních linií referenčních difraktogramů je $0,02^\circ 2\theta$, zatímco před padesáti lety se vystačilo s přesností $0,12^\circ 2\theta$. Čím úctyhodnější je dojem, který database na první pohled vzbuzuje svou velikostí a starobylostí svého původu, tím větší je nebezpečí, že data jež obsahuje budou různorodá a těžko srovnatelná jak mezi sebou, tak s aktuálně změněným difraktogramem analysované látky jejíž identitu na základě porovnání s tou databasí chceme

d	5.5	2.85	3.03	d in Å λ = .708	I I ₁	d in Å λ = .708	I I ₁
I ₁	1.00	1.00	0.80	9.5	0.06	1.72	0.02
I	50	50	40	8.0	0.05	1.74	0.20
$Ce_2(SO_4)_3$				6.7	0.04	1.77	0.12
				5.5	0.50	1.69	0.04
				4.33	0.04	1.56	0.14
				3.50	0.50	1.495	0.04
				3.03	0.90	1.443	0.04
				2.85	1.00	1.430	0.04
				2.71	0.07	1.390	0.20
				2.60	0.02	1.27	0.05
				2.47	0.14	1.260	0.07
				2.37	0.14	1.220	0.06
				2.27	0.14	1.195	0.02
a =	b =	c =		2.15	0.35	1.145	0.04
A =	C =			2.05	0.04	1.135	0.02
D =				2.01	0.07	1.075	0.02
n =	ω =	ε =		1.93	0.06	1.074	0.04
				1.77	0.60	1.052	0.04

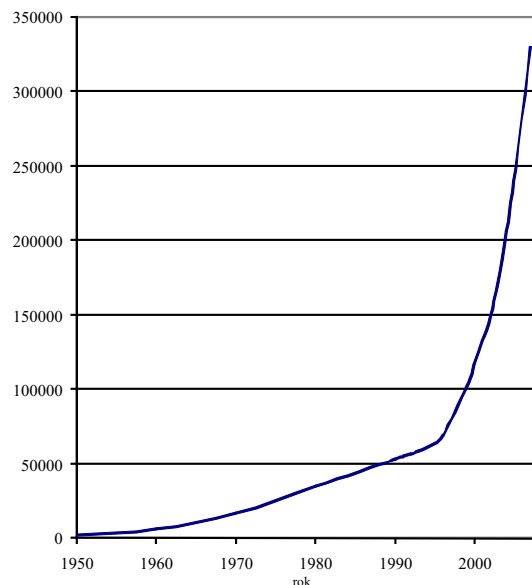
1- 208

Ce ₂ (SO ₄) ₃				dÅ	Int	hkl	dÅ	Int	hkl
Cerium Sulfate				9.5	6		1.56	14	
				8.0	2		1.50	4	
				6.7	4		1.46	4	
				6.1	50		1.43	4	
				5.5	100		1.37	4	
Rad. MoKα ₁ λ 0.709 Filter Beta d-sp				4.85	4		1.32	20	
Cut off Int. I/I ₁				4.33	16		1.28	8	
Ref. Hanawalt, J., Rinn, H., Frevel, L., <i>Anal. Chem.</i> , 10 457 (1938)				3.50	50		1.26	8	
S.G.				3.03	80		1.22	6	
a b c				2.85	100		1.19	2	
α β γ				2.71	4		1.17	4	
Z mp				2.60	2		1.14	2	
D ₄ 3.912 SS/FOM				2.47	14		1.11	2	
Color Colorless, green				2.37	14		1.09	4	
CAS# 13454-94-9; Color and measured density from Data on Chemicals for Ceramic Use, <i>National Research Council Bulletin 107</i> . Decomposition temperature is 920 C at 746 nm.				2.27	14		1.05	4	
				2.15	35				
				2.08	4				
				2.01	8				
				1.95	6				
				1.87	60				
				1.82	2				
				1.76	12				
				1.71	20				
				1.68	12				
				1.63	4				

Obrázek 8. Dvě ukázky referenčních difrakčních údajů. Originální záznam pro síran ceritý z databáse jednoho tisíce referenčních difraktogramů, kterou vybudovali r. 1938 Hanawalt, Rinn a Frevel (a). Moderní prezentace těchto údajů v databáse Powder Diffraction File-International Centre for Diffraction Data (b)

určit. ICDD se s tímto nebezpečím vyrovnala jednak hloubkovou revisí dat uložených v databáse Powder Diffraction File, která proběhla na začátku osmdesátých let [18], jednak zavedením konsistentního souboru přísných pravidel odpovídajících mezinárodním standardům kontroly a řízení jakosti ISO, podle kterých je od té doby hodnocena kvalita každého nového difraktogramu, o němž se uvažuje, že by měl být zařazen do ICDD – Powder Diffraction File.

V posledních 50 letech se také velice rozvinula výpočetní technika, což značně ovlivnilo jak vlastní databáse referenčních difrakčních údajů, tak i způsob získávání těchto údajů (nástup automatických samočinných počítačů řízených difraktometrů) a práci s databáse (database management system). Jedním zvláštním důsledkem rozvoje výpočetní techniky, který ovlivnil ICDD – Powder Diffraction File, se stalo zařazení vypočtených difraktogramů (calculated patterns) do databáse, jež byla původně sbírkou změřených (experimentálně získaných) difrakčních údajů. Na křivce znázorňující jak se zvětšoval počet referenčních difraktogramů v ICDD – Powder Diffraction File (obr. 9) se to projevilo skokovitým nárůstem rozsahu databáse, ke kterému došlo v letech 1996-2000, kdy bylo do Powder Diffraction File zařazeno velké množství (syntetických) difraktogramů vypočtených (nasimulovaných) ze strukturních údajů převzatých z databáse Inorganic Crystal Structure Database (Fachinformationszentrum Physik v Karlsruhe), Crystal



Obrázek 9. Největší databáse referenčních difraktogramů, která v současnosti obsahuje difrakční údaje 350 000 látek, obhospodařuje instituce International Centre for Diffraction Data (ICDD) ve státě Pennsylvania v USA. Na obrázku je znázorněno jak se zvětšoval počet referenčních difraktogramů, které tato databáse obsahuje, od r. 1950 do současnosti.

Data (National Institute of Standards and Technology, Washington D.C.) a Cambridge Crystallographic Database (U.K.).

Přes to, že ICDD – Powder Diffraction File je největší a nejlépe obhospodařovanou databáse pro určování látek (když to porovnáme s databáse infračervených spekter, hmotnostních spekter nebo jaderných magnetických rezonančních spekter) [19, 20], její obsah a přírůstky dramaticky zaostávají za počtem látek registrovaných v databáse Registry File (Chemical Abstracts Service) : 22 609 624 látek 8. ledna 2004 a 46 030 630 látek 22. dubna 2009. Většinu známých látek (látek zaregistrovaných v databáse Chemical Abstracts Service – Registry File) nemůžeme pomocí rtg difrakční kvalitativní fázové analýzy určit, protože difraktogramy těchto látek nemáme k dispozici (nejsou obsaženy v databáse ICDD – Powder Diffraction File). A je malou útěchou, že 75 % všech registrovaných látek je citováno v literatuře jen jedenkrát a že naopak 1 % všech registrovaných látek představuje 45 % všech literárních referencí.

Určitým projevem snahy vymanit se z té zoufalé nedostatečnosti rozsahu ICDD – Powder Diffraction File je prohlubování množství informací, které tato databáse o jednotlivých látkách nese. Vedle difraktogramu, pomocí kterého můžeme danou látku určit (což bylo původním a stále samozřejmě je nejdůležitějším posláním Powder Diffraction File) uvádí se v nové verzi této databáse (Powder Diffraction File – 4) navíc ještě mnoho dalších údajů o struktuře a vlastnostech té které látky, jakož i údaje bibliografické, což z Powder Diffraction File činí univerzální pomůcku materiálového (chemického) inženýrství.

Таблица 16.2. Фрагмент указателя («Ключ») рентгенометрического определителя Михеева

$\frac{d}{n}$	l	$\frac{d}{n}$	l	$\frac{d}{n}$	l	$\frac{d}{n}$	l	$\frac{d}{n}$	l	Минерал	Номер карточек
2,022	9	2,552	10	1,439	9	1,393	9	1,019	9	Литистый турмалин (NaLi _{1,5} Al _{4,5} [B ₃ Al ₃ Si ₆ O ₂₇ F ₃]	722
2,022	9	1,168	10	1,430	7	1,012	7	—	—	Железо (Fe)	16
2,02	10	3,32	9	1,726	8	1,169	5	1,431	4	Берцелинит (Cu _{2-x} Se)	58
2,02	8	3,00	10	2,06	10	3,21	8	4,26	7	Англезит (PbSO ₄)	463
2,02	4	2,96	10	1,813	10	1,547	8	2,57	6	Шватцит (Sb, As) ₂ S ₃ × 3(Cu ₂ , Hg)S	171
2,02	8	2,95	10	3,40	8	2,80	6	2,14	6	Козалит (Pb ₂ Bi ₂ S ₅)	206
2,02	4	2,77	10	2,11	5	2,31	3	1,897	3	Петцит (Ag ₃ AuTe ₂)	56
2,02	10	2,70	10	1,717	10	1,447	9	1,211	6	Маухерит (Ni ₃ As ₂)	68
2,02	8	2,50	10	2,65	8	2,57	8	1,77	8	Пирсонит (Na ₂ CO ₃ × CaCO ₃ · 2H ₂ O)	447a
2,02	2	2,12	10	2,61	7	2,88	5	2,48	4	β-энкайрит (AgCuSe)	66
2,02	10	1,742	8	1,054	8	1,234	6	0,804	6	Никель (Ni)	8
2,020	9	1,427	10	1,053	10	2,441	9	1,552	9	Шпинель (MgAl ₂ O ₄)	312
2,016	9	3,25	10	1,655	10	1,131	8	1,305	7	Штольцит (PbWO ₄)	524
2,016	5	2,892	10	1,804	10	3,707	6	1,509	5	Норденшильдит (CaSnV ₂ O ₆)	534
2,015	4	2,883	10	1,785	6	2,191	5	1,110	5	Доломит [CaMg(CO ₃) ₂]	425
2,013	6	2,914	10	2,615	10	3,88	6	1,434	4	Клейнит (Hg ₄ Cl ₂ O ₃)	397

Обр́азек 10. Ukázka z indexu (klíče) Michejevovy sbírky referenčních difrakčních údajů 1176 minerálů.

3. Vyhledávání v databásích referenčních difraktogramů

Je-li database referenčních difraktogramů malá, nejsou s vyhledáváním v takové databázi žádné problémy. Obsahuje-li však database více než 1000 – 2000 referenčních difraktogramů, neobejdeme se už bez určitých pomůcek a zvláštních postupů, které vyhledávání ve velkých databásích usnadňují [4].

Jako první se uplatnila pomůcka využívaná při hledání ve spektroskopických koincidenčních tabulkách – tak zvaný klíč. V aplikaci na určování látek pomocí rtg difrakce je to soubor všech difrakčních linií všech referenčních difraktogramů dané database, uspořádaný sestupně podle hodnoty mezirovinné vzdálenosti (tedy podle polohy difrakční linie, která jest odpovídající hodnotou mezirovinné vzdálenosti kódována). Je tudíž uspořádání difrakčních údajů v klíči inverzní vůči uspořádání difrakčních údajů ve vlastní databázi, kde jsou difrakční data uspořádána podle látek (standardů). Proto se klíči také někdy říká inverzní soubor (a vlastní databázi pak označujeme jako přímý soubor). Klíčem je opatřena například Michejevova database referenčních difraktogramů minerálů [21], z níž je také převzata ukázka na obr. 10. Z toho obrázku je patrné že pro vyhledávání ve větších databásích ztrácí používání klíče účinnost, nebo spektrum difrakčních linií se stává příliš husté. Don Hanawalt proto v roce 1936 vyvinul jinou pomůcku, jíž se říká Hanawaltův index [22]. Hanawaltův index se tvoří ze seznamu referenčních difraktogramů, které jsou uspořádány podle hodnoty mezirovinné vzdálenosti d své nejintenzivnější difrakční linie. Tento seznam je pak rozčleněn na několik desítek malých, zhruba stejně početných dílčích seznamů, tzv. Hanawaltových grup, jejichž hodnoty d pro nejintenzivnější difrakční linii na sebe navazují (dokonce se mírně překrývají aby se omezila možnost selhání, ke kterému by při vyhledávání v databázi referenčních difraktogramů mohlo dojít v důsledku nepřesnosti udávaných hodnot d). Každá Hanawaltova grupa je uspořádána podle hodnoty d druhé nejintenzivnější linie

(obr.11). Aby se omezil vliv nepřesnosti stanovování intenzit difrakčních linií, každý referenční difraktogram vstupuje do Hanawaltova indexu několikrát (ale ne více než čtyřikrát) v preuspořádaném pořadí svých nejintenzivnějších difrakčních linií. Hanawaltův index se znamenitě osvědčil [23] a dodnes se úspěšně využívá při hledání v databázi referenčních difraktogramů PDF-ICDD.

Po druhé světové válce se začala rychle rozvíjet technika zpracování hromadných dat a to se samozřejmě uplatnilo i při vyhledávání v databásích referenčních difraktogramů. První počítačový systém pro identifikaci krystalických látek pomocí rtg difrakce vytvořil Ludo Frevel roku 1965 [24]. Pracoval s databází referenčních difraktogramů 1359 látek, která byla prohledávána programem napsaným v jazyce Algol pro počítač Burroughs 5000. První velký počítačový práškový difrakční identifikační systém, který pracoval s databází asi 12000 difraktogramů, vytvořili Vladimír Vand (krystalograf českého původu) a Gerald G. Johnson (profesor výpočetní techniky na Pensylvánské státní universitě) r. 1967 [25]; program byl napsán ve Fortranu (ale jádro nejdůležitější procedury bylo v IBM-assembleru) pro počítač IBM 7074. Výpočetní technika se rychle rozvíjela a vznikaly nové a nové počítačové systémy pro automatisovanou identifikaci fázového složení pomocí rtg difrakce. V roce 1986 vyšla monografie popisující deset nejznámějších počítačových rtg difrakčních systémů, jež byly ve světě v té době používány [26] a r. 1991 byl publikován seznam 17 programů pro práškovou difrakční identifikaci látek [27], který byl v následujících letech opakovaně aktualizován. Dokud byly počítače pomalé a měly malou operační paměť, vyvíjely se důmyslné algoritmy, které paměť a výpočetní čas šetřily [28-30].

V současné době se počítače zdokonalily natolik, že na tyto aspekty již není třeba brát zřetel. A objevily se nové možnosti: dnes se v programech pro identifikaci fázového složení krystalických látek na základě porovnání difraktogramu zkoumané látky s databází referenčních difraktogramů využívají různé indexy, což bylo dříve

s ohledem na omezenou paměť počítačů nemožné. Každý index představuje jakousi repliku databáze referenčních difraktogramů, uspořádanou nějakým zvláštním způsobem, který je pro ten index specifický: například podle prvkového složení nebo určitým způsobem zvolené a vhodně komprimované kombinace příznaků odvozených z difrakčních údajů. Indexy jsou kódovány většinou jako binární mapy, tj. matice, jejichž prvky jsou nula nebo jednička podle toho, zda se příslušný příznak (na určité hladině) u daného standardu (referenčního difraktogramu) vyskytuje čili nic. Zařazení indexů do identifikačního systému vedle vlastní databáze sice představuje dodatečné nároky na paměť, rychlost identifikačního procesu se tím však zvětšuje. Množství dat uchovávaných pro daný standard v nějakém indexu je totiž mnohem menší než množství dat zaznamenaných pro ten standard v kmenové databázi. Proto trvá prohledávání indexu mnohem kratší dobu než prohledávání celé databáze. Na druhé straně však výsledky prohledání indexu přinesou mnohem menší informaci než prohledání celé databáze. Ten informační zisk (v bitech) se počítá jako

$$\log_2 \frac{A}{B}$$

kde A je počet alternativ (látek, které přicházejí v úvahu, že by mohly být identické s neznámou, tj. určenou látkou) před identifikačním aktem a B jest počet alternativ po identifikačním aktu. Pro určitý identifikační systém jest A rovno počtu referenčních difraktogramů v databázi toho systému, a to bez ohledu na to, zda identifikační akt spočívá v prohledání databáze nebo jenom určitého indexu (nebo v jakémkoli indexu jsou samozřejmě zahrnuty všechny standardy, které pokrývá ta databáze). Hodnota B bude však v případě, že prohledáváme index větší nebo rovna hodnotě B , kterou dostaneme prohledáním celé databáze. V indexu se totiž vyskytuje jen část údajů, které o jednotlivých standardech pokrývá celá databáze a proto bude výsledek identifikace provedené na základě indexu méně určitý než výsledek identifikace, při níž jsme použili celou databázi. Když však prohledávání indexu pojmem jenom jako počáteční akt identifikace, ve které pak pokračujeme dále ještě prohledáním celé databáze, můžeme využít informaci, získanou prohledáním indexu k tomu, abychom ve vlastní databázi prohlíželi údaje jenom o těch standardech, které byly vytipovány na základě prohledání indexu. Tím se ovšem doba prohledávání databáze velice zkrátí, což více nebo méně kompenzuje ztrátu, již představuje jednak doba nezbytná k prohledávání indexu, jednak paměť nezbytná k uchování indexu v identifikačním systému. Většinou se používá několik indexů, čímž se účinnost takto koncipovaných systémů ještě mnohokrát zvyšuje. Budeme to ilustrovat na konkrétním příkladu identifikačního systému, pracujícího s databází 47847 referenčních difraktogramů a třemi indexy, v kterých jsou standardy pokryté databází utříděny jednak dle svého typu (anorganické látky, organické látky, minerály, kovy a slitiny a j.), dále dle prvkového složení a posléze podle hodnoty mezirovinné vzdálenosti tří nejintenzivnějších difrakčních linií. Určovaná neznámá látka byla nalezena v přírodě a předpokládalo se tedy, že to

je nerost; dále se rtg fluorescenční spektrální analýzou zjistilo, že ta látka obsahuje křemík a když se pořídil difraktogram, našla se na něm nejintenzivnější difrakční linie odpovídající mezirovinné vzdálenosti 3,35 Å. Pomocí prvního indexu zjistíme, že databáze pokrývá 3544 minerálů, prohlédnutím druhého indexu se dozvíme, že v databázi existuje 2895 látek obsahujících křemík a posléze ve třetím indexu nalezneme 816 standardů na jejichž difraktogramu se mezi třemi nejintenzivnějšími difrakčními čarami vyskytuje linie, jejíž mezirovinná vzdálenost leží mezi 3,32 Å a 3,36 Å. To všechno zjistíme velmi rychle, nebo ty tři indexy obsahují mnohem méně dat než vlastní databáze. Velmi snadno (rychle) také zjistíme, že všech minerálů které obsahují křemík, je 1004 a že mezi těmito (křemík obsahujícími) 1004 minerály je pouze 27 takových, které mají intenzivní difrakční linii v intervalu (3,32 Å; 3,36 Å). Teď už stačí prozkoumat jen těchto 27 referenčních difraktogramů z celé databáze (obsahující 47847 referenčních difraktogramů) abychom neznámou látku (nižší křemen) identifikovali. Což bude trvat mnohem kratší dobu, než kdybychom prozkoumávali a s difraktogramem neznámé (určované) látky porovnávali čáru po čáře všechny referenční difraktogramy (v počtu 47847) jež vlastní databáze obsahuje. První softwarový systém této koncepce (intenzivně využívající vedle vlastní databáze ještě celou řadu indexů) v aplikaci na databázi referenčních difraktogramů ICDD – Powder Diffraction File se jmenoval POWDER SUITE [31] a pracoval v prostředí VAX/VMS. Současná verze tohoto softwaru se jmenuje ICDD SUITE a pracuje v prostředí Windows.

Moderní automatické práškové difrakční identifikační systémy pracují velmi rychle; mnohem rychleji než v šedesátých letech, kdy se s nimi začínalo, a to přesto, že tenkrát bylo k prohledávání a porovnávání mnohem méně referenčních difraktogramů než dnes. Identifikační účinnost práškových difrakčních identifikačních systémů je však kupodivu stále poměrně nízká, zejména když se jedná o identifikaci složitějších směsí. Se směsí více než pěti složek si žádný ze současných automatických počítačových práškových difrakčních identifikačních systémů spolehlivě neporadí. V ideové koncepci, v samotné myšlenkové podstatě těchto systémů bude asi nějaká slabina, která se nasazením nejmodernější výpočetní techniky vykompenzovat nedá.

Proto se zkouší a vyvíjejí stále nové identifikační systémy a laboruje se s netradičními postupy. Objevily se například identifikační systémy, které využívají fuzzy logiku [32–34], čímž by se mělo vyloučit nebo aspoň potlačit nebezpečí, že k neúspěchu identifikace dojde v důsledku rigidnosti kritérií na základě kterých shodu resp. rozdílnost porovnávaných difraktogramů hodnotíme. Slibné jsou také pokusy se „strukturně citlivými“ algoritmy [35–38], jimiž by se měly eliminovat nepříjemnosti související s tuhými roztoky a strukturně příbuznými látkami. Takovéto látky se vyskytují často spolu a jejich určení ve směsi představuje pro konvenční identifikační systém většinou nerozlousknutelný oříšek, a to nejen v případě, že se jejich referenční difraktogramy v používané databázi nenalézají.

Ani jedním z těchto přístupů se však zatím nepodařilo účinnost identifikačních systémů při určování složitějších



3.39 – 3.32 (± .02)										File No.	I/Ic
i	3.38 _x	8.58 _x	3.04 _o	4.11 _o	3.18 _o	1.69 _o	2.65 _o	1.88 _o	(Mg,Fe) ₂ Al ₂ Si ₂ O ₁₀ /Cordierite, ferroan	9– 472	
	3.33 _x	6.72 _o	3.19 _o	8.09 _o	3.28 _o	5.18 _o	3.10 _o	4.30 _o	C ₁₅ H ₁₉ N ₇ O ₆	29–1716	0.20
i	3.31 _o	6.40 _x	6.10 _o	3.85 _o	2.77 _o	6.70 _o	3.48 _o	2.64 _o	C ₁₂ H ₉ Cl ₃	17–1054	
	3.38 _o	6.13 _x	8.66 _o	3.20 _o	3.29 _o	9.70 _o	3.46 _o	4.57 _o	C ₁₁ H ₁₁ N ₂ :HCl	28–1749	
i	3.34 _x	5.93 _o	5.19 _o	3.77 _o	3.65 _o	3.51 _o	2.94 _o	1.67 _o	C ₆ H ₈ N ₂ O ₂	26–1863	3.30
	3.37 _x	5.85 _o	3.86 _o	3.72 _o	3.52 _o	3.03 _o	2.70 _o	7.72 _o	C ₆ H ₉ N ₃ O ₂ :HCl	5– 459	
*	3.31 _o	5.73 _x	3.43 _o	3.59 _o	3.19 _o	4.36 _o	4.19 _o	3.27 _o	C ₆ H ₃ NO ₂	30–1845	1.00
	3.30 _x	5.44 _o	5.63 _o	3.24 _o	4.97 _x	6.58 _o	4.58 _o	3.15 _o	(NH ₄) ₄ P ₂ O ₇	20– 102	
	3.38 _x	5.30 _x	3.49 _x	5.90 _o	3.67 _o	3.26 _o	3.18 _o	2.99 _o	KH ₂ P ₂ O ₇	15– 509	
	3.35 _x	5.21 _o	4.86 _o	4.33 _o	4.04 _o	3.90 _o	3.55 _o	2.73 _o	β-C ₆ H ₁₁ NO ₂	22–1874	
i	3.40 _x	5.01 _o	3.09 _o	4.10 _o	3.00 _o	4.03 _o	6.74 _o	3.45 _o	C ₂ H ₄ N ₄	24–1654	1.10
*	3.30 _x	4.76 _o	4.18 _o	5.73 _o	2.92 _o	3.98 _o	2.38 _o	3.35 _o	C ₈ H ₄ O ₄	37–1919	
	3.31 _x	4.71 _o	3.50 _o	5.56 _o	3.84 _o	3.03 _o	7.02 _o	2.30 _o	C ₈ H ₂ N ₂ O ₂ :HCl	29–1827	
*	3.39 _x	4.48 _o	3.43 _o	3.01 _o	4.09 _o	2.98 _o	2.78 _o	3.18 _o	NaHSO ₄	25– 833	
	3.34 _o	4.42 _x	10.1 _o	1.48 _o	2.56 _o	1.68 _o	1.28 _o	1.23 _o	Al ₂ Si ₂ O ₄ (OH) ₄ ·2H ₂ O/Halloysite-10A	9– 451	
i	3.40 _o	4.38 _x	2.88 _o	5.76 _o	2.61 _o	4.09 _o	2.76 _o	1.76 _o	V ₂ O ₅ /Shcherbinaite, syn	9– 387	1.60
*	3.33 _x	4.30 _o	2.82 _o	6.08 _o	4.72 _o	1.71 _o	3.52 _o	2.15 _o	(NH ₄) ₂ Co ₂ (SO ₄) ₃	22–1037	2.30
i	3.37 _x	4.28 _o	1.84 _o	1.55 _o	2.47 _o	2.31 _o	1.39 _o	2.14 _o	AlPO ₄ /Berlinite, syn	10– 423	
*	3.34 _x	4.26 _o	1.82 _o	1.54 _o	2.46 _o	2.28 _o	1.37 _o	1.38 _o	SiO ₂ /Quartz, low, syn	33–1161	3.60
o	3.36 _x	4.23 _x	3.57 _o	5.27 _o	3.72 _o	4.04 _o	3.97 _o	7.19 _o	C ₆ H ₈ N ₂ O ₂ S	30–1944	
*	3.35 _x	4.22 _o	3.25 _o	4.43 _o	3.67 _o	6.22 _o	2.89 _o	3.55 _o	C ₆ H ₄ N ₂ O ₄	37–1915	
i	3.32 _x	4.22 _x	5.28 _o	4.97 _o	3.52 _o	3.48 _o	2.65 _o	2.41 _o	C ₆ H ₄ N ₄ O ₂ :HCl	25–1541	
i	3.35 _o	3.88 _x	3.73 _o	3.54 _o	2.91 _o	2.52 _o	2.32 _o	1.92 _o	C ₆ H ₈ NO ₄ :HCl	25–1925	
*	3.39 _x	3.87 _o	6.52 _o	2.36 _o	3.01 _o	2.04 _o	1.99 _o	2.41 _o	HgSO ₄	31– 867	
*	3.41 _o	3.84 _x	3.52 _o	3.26 _o	3.87 _o	3.03 _o	2.74 _o	2.37 _o	KHSO ₄ /Mercurite, syn	11– 649	
*	3.31 _x	3.77 _o	4.22 _o	3.24 _o	3.29 _o	2.99 _o	3.47 _o	2.90 _o	KAlSi ₃ O ₈ /Orthoclase	31– 966	
*	3.36 _o	3.52 _x	7.69 _x	6.16 _o	3.84 _o	3.14 _o	3.09 _o	3.09 _o	C ₆ H ₄ (CO) ₂ C ₂ H ₄ /Hoelite, syn	28–2002	
*	3.35 _x	3.50 _o	5.04 _o	3.56 _o	4.00 _o	3.15 _o	5.58 _o	2.48 _o	(NH ₄) ₂ S ₂ O ₈	31– 69	
*	3.36 _x	3.47 _o	6.52 _o	2.59 _o	3.02 _o	3.28 _o	3.56 _o	2.61 _o	BaAl ₂ Si ₂ O ₈ /Celsian, syn	38–1450	
*	3.36 _x	3.47 _o	6.52 _o	2.59 _o	3.02 _o	3.28 _o	3.56 _o	2.61 _o	BaAl ₂ Si ₂ O ₈ /Celsian, syn	38–1450	
*	3.30 _x	3.47 _o	3.66 _o	4.87 _o	3.06 _o	3.03 _o	2.86 _o	2.86 _o	K ₂ Cr ₂ O ₇ /Lopezite, syn	27– 380	0.63
	3.33 _x	3.46 _o	3.79 _o	3.26 _o	3.01 _o	2.58 _o	2.91 _o	2.77 _o	(K,Ba)(Si,Al) ₄ O ₈ /Orthoclase, barian	19– 3	
*	3.38 _x	3.45 _o	3.44 _o	2.28 _o	6.30 _o	2.22 _o	2.75 _o	5.07 _o	KPO ₃ /Potassium metaphosphate	35– 819	
i	3.39 _x	3.43 _x	2.21 _o	5.39 _o	2.54 _o	2.69 _o	1.52 _o	2.12 _o	Al ₂ Si ₂ O ₇ /Mullite, syn	15– 776	
	3.39 _x	3.41 _o	2.31 _o	2.84 _o	3.82 _o	2.14 _o	2.03 _o	2.00 _o	NaBF ₄ /Ferruccite, syn	11– 671	
	3.41 _o	3.39 _x	2.31 _o	2.84 _o	3.82 _o	2.14 _o	2.03 _o	2.00 _o	NaBF ₄ /Ferruccite, syn	11– 671	
i	3.38 _x	3.39 _o	2.53 _o	3.11 _o	2.29 _o	3.57 _o	2.41 _o	2.37 _o	Gd ₂ S ₃	20–1056	
i	3.39 _x	3.38 _o	2.53 _o	3.11 _o	2.29 _o	3.57 _o	2.41 _o	2.37 _o	Gd ₂ S ₃	20–1056	
*	3.30 _x	3.29 _x	4.76 _o	4.18 _o	5.73 _o	2.92 _o	3.98 _o	2.38 _o	C ₆ H ₈ O ₄	37–1919	
*	3.33 _x	3.28 _o	2.97 _o	3.83 _o	2.36 _o	2.34 _o	2.35 _o	2.21 _o	CdSO ₄	14– 352	

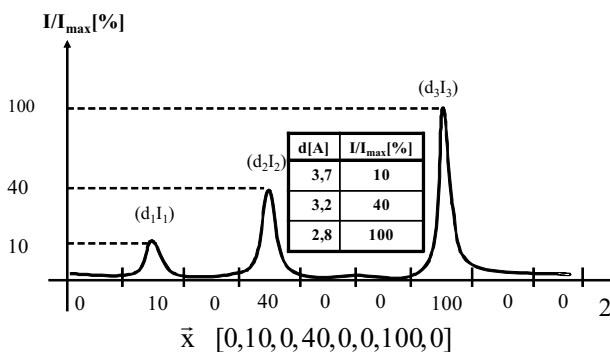
Obrázek 11. Ukázka z Hanawaltova indexu database referenčních difrakčních údajů Powder Diffraction File, obhospodařované ICDD.

směsí v obecném případě podstatně zlepšit. Proto v následujícím paragrafu odvodíme matematický model rtg difrakční kvalitativní fázové analýzy pomocí kterého lze mechanismus identifikačního procesu detailně a kvantitativně krok po kroku sledovat s cílem odhalit skutečnou příčinu nedostatečnosti jeho současné koncepce a najít způsob jak zvýšit účinnost rtg difrakční identifikace složitějších směsí.

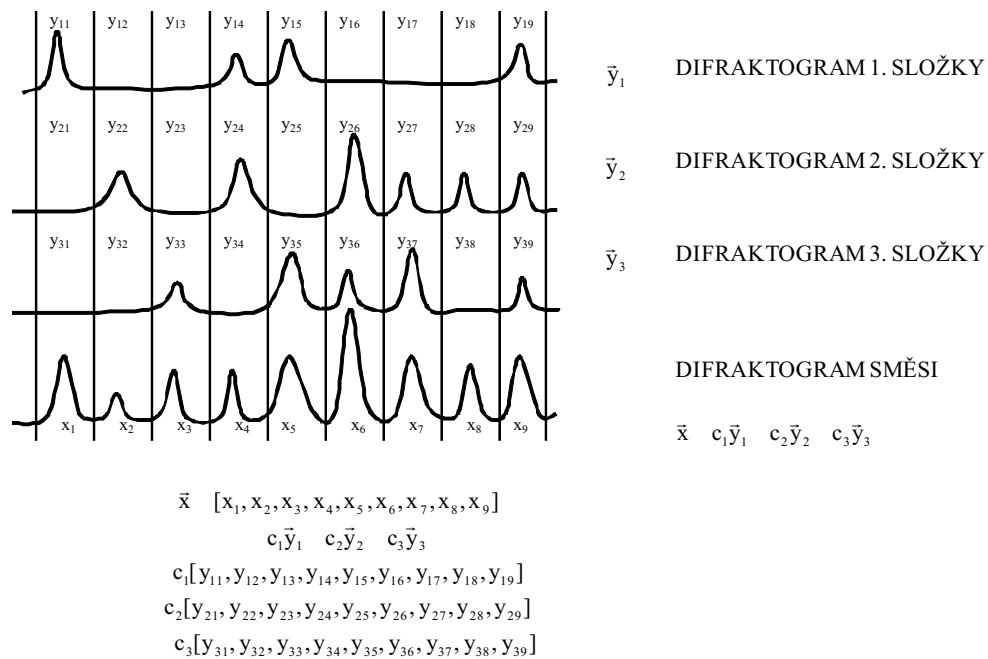
4. Identifikační algoritmy

Rozdělíme-li interval měřených difrakčních úhlů (nebo jim odpovídajících mezirovinných vzdáleností) na dílčí subintervaly, můžeme každý difraktogram chápat také jako číselný vektor, tedy n-tici čísel rovných intenzitám difrakčních linií, ležících v jednotlivých subintervalech (jestliže do nějakého subintervalu nepřípadně žádná difrakční linie, přiřadíme tomuto číslo nula) – obr. 12. V takto umělem zavedeném abstraktním prostoru číselných vektorů A můžeme potom jako v zrcadle sledovat operace uskutečňující se v reálném prostoru skutečných difraktogramů B při rentgenografické identifikaci krystalických látek. To s sebou přináší dvě důležité výhody: za prvé to umožňuje využít matematickou strukturu protiobrazného prostoru A difraktogramů kódovaných jako číselné vektory k algebraisaci postupů a konstrukcí používaných při kvalitativní fázové analýze; za druhé to umožňuje využít naši geometrickou zkušenost s (třírozměrným) vektorovým prostorem k hlubšímu pochopení smyslu těchto postupů a konstrukcí. Můžeme to ilustrovat na příkladu kritérií pro porovnávání (hodnocení podobnosti nebo odlišnosti) difraktogramů, což jest alfa a omega identifikace fázového složení pomocí rtg. difrakce. Takových kritérií byla zkonstruována řada, několik desítek

[39] a více nebo méně dobře se při identifikaci krystalických látek na základě porovnávání jejich rtg difraktogramů osvědčují. Ta kritéria byla ovšem vesměs odvozena empiricky (řekněme „citem“) aniž by bylo známo co přesně znamenají a proč fungují tak jak fungují – někdy lépe, někdy hůře. A vůbec už není známo jak by se na základě zkušeností se stávajícími kritérii mohla konstruovat kritéria nová, jež by byla účinnější. Pokud však pojmem difraktogramy jako vektory body (prvky) abstraktního prostoru difraktogramů A , stává se význam kritéria podobnosti (odlišnosti) difraktogramů jasný: jedná se o blízkost (odlehlost) a měření vzdálenosti (odlehlosti) dvou bodů, tj. o metriku v prostoru A . Definujeme-li



Obrázek 12. Každému práškovému difraktogramu lze přiřadit číselný vektor, tj. n-tici čísel vyjadřujících intenzity difrakčních linií ležících v subintervalech (kanálech), na které celkový úhlový rozsah difraktogramu rozdělíme.



Obrázek 13. Vektor, který popsaným způsobem přiřazujeme difraktogramu směsi, je superposicí (lineární kombinací) vektorů odpovídajících difraktogramům jednotlivých komponent té směsi s koeficienty úměrnými jejich koncentracím.

prostor A jako metrický, tj. zavedeme-li v něm podle pravidel funkcionální analýzy nějakou metriku [40,41], dostáváme tím do ruky nástroj pro kvantitativní porovnání difraktogramů. Různá kritéria podobnosti difraktogramů představují pak různé metriky v prostoru vektorové reprezentace difraktogramů A . Metrika je ústředním pojmem při klasifikaci [42-44], rozpoznání [45-48], shlukové [49-51] a diskriminantní analýze [52]. Zde všude se musí kvantitativně vyhodnocovat, tedy „měřit“ podobnost (odlišnost) tříděných, identifikovaných a dle různých hledisek rozlišovaných objektů. Objektů, jež jsou „mnohorozměrné“ v tom smyslu, že jsou určovány velkým množstvím příznaků (aspektů): obdobně jako je difraktogram definován velkým množstvím čísel vyjadřujících polohy a intenzity jednotlivých difrakčních maxim. Velké myšlenkové bohatství, jež se při budování teorie a praktické aplikaci třídění, identifikace a rozlišování objektů v nejrůznějších přírodních a technických disciplínách v průběhu posledního půlstoletí nashromáždilo, můžeme znamenitě využít pro porovnání difraktogramů kódovaných jako vektory.

Pokud bychom měli určovat (identifikovat) pouze látky, kterým se ve fyzikální chemii říká někdy jednoduché, tj. látky tvořené jedinou krystalickou fází, pak by to stačilo. Metoda „otisku palce“ vycházející z porovnání difraktogramu určované, neznámé látky s databází „vzorových difraktogramů“ látek známých, je v tomto případě adekvátní a dobře funguje. Potíže jsou s identifikací směsí. Zde musíme použít takový matematický model, který by vycházel z nejdůležitější fyzikální zákonitosti, jež se při tom uplatňuje: totiž difraktogram (analysované) směsi látek je lineární superposicí (kombinací) difraktogramů ji tvořících komponent (s koeficienty rovnými koncentracím těch komponent);

jestliže chceme na základě porovnávání difraktogramů určovat nejen jednoduché látky, ale i směsi, musí být abstraktní prostor difraktogramů A , který je nositelem našeho matematického modelu rtg difrakční kvalitatívni fázové analýzy, nejen metrický ale i lineární [53]: tak aby vektor, reprezentující v tomto modelu difraktogram směsi, byl lineární superposicí vektorů představujících difraktogramy složek té směsi (s koeficienty rovnými koncentracím těch složek) – obr. 13.

A úkolem kvalitatívni fázové analýzy formulovaným v kategoriích tohoto matematického modelu je pak nalézt hodnoty oněch koeficientů c_i ; $i=1,2,\dots,m$ ze vztahu

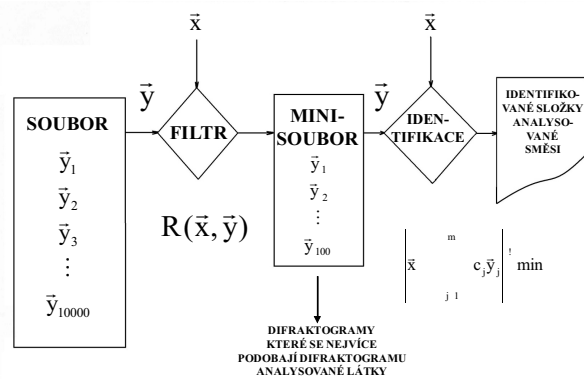
$$\bar{x} = \sum_{i=1}^m c_i \bar{y}_i \quad (1)$$

kde vektor \bar{x} představuje difraktogram analysované látky a vektory $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_m$ vyjadřují referenční difraktogramy použité database. Většina čísel c_1, c_2, \dots, c_m jejichž hodnoty se rovnají koncentracím toho kterého standardu v analysované směsi, bude ovšem v případě database obsahující velké množství referenčních difraktogramů nulová.

Vzhledem k měřickým chybám a také vzhledem k tomu, že difraktogramy různých preparátů jedné a téže látky se od sebe mohou lišit (např. vlivem textury) nebude rovnice (1) splněná přesně a to co hledáme ve skutečnosti jsou čísla c_i ; $i = 1, 2, \dots, m$ které tu rovnici splňují „co nejlépe“, tedy čísla c_i pro která residuální vektor

$$\bar{x} - \sum_{i=1}^m c_i \bar{y}_i \quad (2)$$

má co nejmenší normu



Obrázek 15. Rtg difrakční fázová analýza je dvojestupňová procedura. V první etapě (search) prohledáme soubor (databasi) referenčních difraktogramů a vyhledáme mezi nimi ty, které jsou co nejvíce podobné difraktogramu analysované látky. Z takto vytvořeného minisouboru referenčních difraktogramů pak vyjdeme v druhé etapě kvalitativní fázové analýzy (match), když se snažíme difraktogram analysované látky vyjádřit superpozicí difraktogramů té referenční minibase.

$$\frac{f}{c_i} \quad 0; i = 1, 2, \dots, m \quad (9)$$

kteřá je ekvivalentní s rovnicemi (7). Když difraktogram popisujeme jako n -tice čísel $\vec{x} = (x_1, x_2, \dots, x_n)$; $\vec{y} = (y_1, y_2, \dots, y_n)$ představujících intenzity difrakčních linií v jednotlivých kanálech, které charakterisují polohu těch difrakčních linií (obr.12), můžeme skalární (vnitřní) součin definovat tak, jak se to dělá v Euklidově prostoru:

$$(\vec{x}, \vec{y}) = ((x_1, x_2, \dots, x_n), (y_1, y_2, \dots, y_n)) = \sum_{i=1}^n x_i y_i$$

resp.

$$(\vec{x}, \vec{y}) = \sum_{i=1}^n w_i x_i y_i$$

kde $\vec{w} = (w_1, w_2, \dots, w_n)$ je váhový vektor. Někdy by se však mohla stát volba dělicích bodů, které definují rozdělení používaného intervalu difrakčních úhlů (potažmo mezirovinných vzdáleností) na dílčí subintervaly (kanály určující polohu difrakční linie) problematická: například u tuhých roztoků, kde se poloha difrakčních linií s měnícím se prvkovým složením mírně posouvá nebo u látek, které mají široké difrakční linie (jako jsou jílové minerály). Potom je lépe chápat difraktogram jako bod (vektor) nekonečněrozměrného prostoru, představovaný spojitou funkcí (závislostí) $I(2\theta)$ intenzity difraktovaného záření I na difrakčním úhlu 2θ

$$\vec{x} = x(2\theta); 2\theta \in \langle r, s \rangle$$

(obr. 15). V tom případě se pak součet resp. číselný násobek vektorů reprezentujících difraktogramy definuje jako

$$\vec{a} + \vec{b} = (a + b)(2\theta) = a(2\theta) + b(2\theta)$$

resp.

$$c\vec{a} = (ca)(2\theta) = c \cdot a(2\theta)$$

a skalární součin vyjadřujeme integrálem

$$(\vec{a}, \vec{b}) = \int_r^s w(2\theta) a(2\theta) b(2\theta) d(2\theta)$$

kde $w(2\theta)$ je váhová funkce. Anebo se skalární součin definuje pomocí normy vektoru

$$(\vec{a}, \vec{a}) = \frac{1}{4} \|\vec{a}\|^2 = \frac{1}{4} \int_r^s w(2\theta) a^2(2\theta) d(2\theta)$$

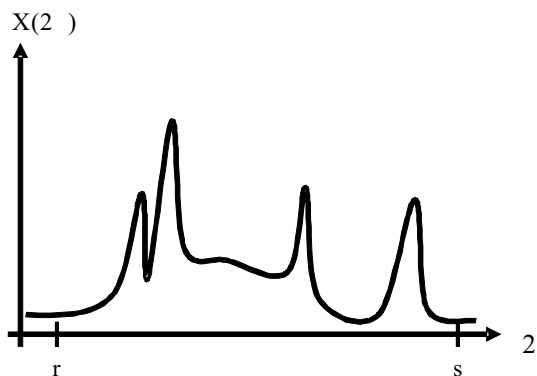
kteřá je vyjádřena jako

$$\|\vec{a}\| = \sqrt{\int_r^s w(2\theta) a^2(2\theta) d(2\theta)}$$

kde $a(2\theta)$ je intenzita difraktovaná ve směru určeném difrakčním úhlem 2θ .

Intenzita difrakčních linií je tedy důležitá nejen pro kvantitativní fázovou analýzu ale i pro fázovou analýzu kvalitativní. A problémy kolem intenzity difrakčních linií jsou příčinou závažných obtíží a nedostatků, se kterými se kvalitativní fázová analýza potýká. Především se jedná o to, že naměřené hodnoty intenzit difrakčních linií jsou méně přesné než hodnoty mezirovinných vzdáleností určujících polohu linií; a jsou také hodně citlivé na použité experimentální uspořádání. S tímto zlem se potýká zejména ICDD, která vyvíjí velké úsilí o zvýšení kvality referenčních difrakčních údajů. Druhým zásadním problémem kolem intenzit difrakčních linií je to, že tyto hodnoty nejsou – obecně vzato – správné: jsou totiž uváděny jako relativní čísla a vyjadřují poměr intenzity dané difrakční linie příslušné látky k intenzitě její nejsilnější difrakční linie. Tím mohou vznikat při identifikaci směsi systematické chyby vyjádřené stovkami procent jmenovitého podílu té které složky. Proti tomu se bojuje tak, že hodnoty intenzit difrakčních linií se v referenčních databasích uvádějí nikoli jako relativní, ale absolutní, vztažené například k nejintenzivnější linii korundu, jenž se k dané látce přimíchá v poměru 1:1 (říká se tomu RIR = reference intensity ratio [59, 60]). To, že intenzity difrakčních linií, jejich přesnost a správnost jsou pro kvalitativní fázovou analýzu velmi důležité plyne ze skutečnosti, že identifikovat směsi znamená určit kolik je v ní které složky, tedy jaké jsou hodnoty c_i ve vztahu (3). Tento vztah je matematickou formulací kvalitativní fázové analýzy a plyne z něj jasně že c_i jsou funkcí \vec{x} a \vec{y}_i , tj. (směrové distribuce) intenzity záření difraktovaného analysovanou směsí (\vec{x}) a jejími složkami (\vec{y}_i). To, že hodnoty intenzit difrakčních linií neznáme dost dobře, také úzce souvisí s vůbec největší slabinou kvalitativní fázové analýzy, totiž s tím, že není schopna spolehlivě identifikovat složitější směsi. V následující kapitole ukážeme, že odstranění tohoto tíživého nedostatku kvalitativní fázové analýzy předpokládá rigorosní, tzv. faktorovou analýzu difraktogramu zúčastněných látek, jejíž nezbytným předpokladem je korektnost intenzit difrakčních linií se kterými se přitom pracuje.

Identifikace (kvalitativní fázová analýza) je dvojestupňový proces (obr. 15). První krok, kterému se tradičně říká SEARCH, je diskriminační (analytický). Z database



Obrázek 16. Pokud jsou difrakční linie (píky na rtg práškovém difraktogramu) široké a/nebo se překrývají, stávají se rozdělení celkového úhlového oboru difraktogramu na dílčí subintervaly (kanály) a na to navazující reprezentace práškových difraktogramů jako vektorů Eukleidova prostoru problematické. V tomto případě můžeme při matematickém modelování kvalitativní fázové analýzy pojímat difraktogram jako bod (vektor) Hilbertova či Banachova prostoru, představovaný spojitou funkcí (závislostí) intenzity difraktovaného záření na difrakčním úhlu.

referenčních difraktogramů, již při identifikaci používáme, vybereme ty, které jsou difraktogramu neznámé, určené látky co nejpodobnější. V druhém kroku (MATCH) pak provedeme syntesu referenčních difraktogramů takto vytvořené minibase s cílem co nejlépe namodelovat difraktogram určené látky [61]. Duálnost identifikační procedury (SEARCH/MATCH algoritmu) reflektuje dva aspekty obecné metodologie rozpoznávání, nebo identifikace, chcete-li (pattern recognition). SEARCH jest typickou procedurou tzv. analytického rozpoznávání (někdy se hovoří o statistickém rozpoznávání nebo o rozpoznávání s hlediska teorie rozhodování), které funguje na principu „otisku palce“, tj. klasifikuje objekty dle existujících vzorů (ideálů) [48]. MATCH pak patří do kategorie syntaktického rozpoznávání [62], při němž hledáme vztahy mezi podmnožinami identifikovaného subjektu; v našem případě difraktogramy komponent analysované směsi, kterou pak identifikujeme v kategoriích těchto podmnožin (koncentracemi jejích komponent). Po pravdě řečeno, pro identifikaci krystalických látek pomocí rtg difrakce se pořádně nehodí ani jedno ani druhé (tedy ani analytické ani syntaktické rozpoznání). Analytické rozpoznávání se nehodí proto, že většina určených látek jsou směsi, a proto se difraktogram určené látky principiálně musí lišit od kteréhokoliv difraktogramu referenční databáze (nebo to jsou difraktogramy jednoduchých, tj. jednofázových látek). Syntaktické rozpoznávání se nehodí z toho důvodu, že struktura určeného objektu je příliš složitá. V aplikaci syntaktického rozpoznávání na difrakční identifikaci látek se totiž objektem rozumí difraktogram, jeho prvky jednotlivé difrakční linie a strukturou pak dispozice (distribuce) těchto linií na difraktogramu. Možných podmnožin je tolik, kolik difraktogramů obsahuje použitá referenční databáze – v případě PDF-4/Organics 2008/2009 341540 a to je opravdu příliš. Na identifikaci systému s tak velkým počtem parametrů dnes ještě žádnou fungující metodu neznáme. A tak se pro určení fázového složení krystalických látek používá mnoho různých algoritmů, které se liší tím, jak velký důraz

kladou na tyto dvě složky identifikačního procesu, SEARCH a MATCH a jakým způsobem je konkrétně realizují.

Během vyhledávací etapy identifikační procedury, tedy během toho, čemu se říká SEARCH se pro porovnání difraktogramu určené látky s difraktogramy referenční databáze využívají různá kritéria, jež jsou měrou podobnosti (odlišnosti) porovnávaných difraktogramů (tab.1). Je to vlastně různá metrisace abstraktního prostoru difraktogramů, anebo to můžeme chápat jako filtraci, již se z použité databáze referenčních difraktogramů vyberou („vyfiltrují“) difraktogramy co nejpodobnější difraktogramu neznámé (určené) látky. Jak už jsme říkali, je takové podnikání svým způsobem pochybné, nebo v případě, že je neznámá látka směsí, nebude se její difraktogram moc podobat žádnému z jejích komponent. A to je důvod toho, že se používají různá kritéria, různé filtry, a že se jich aplikuje celá řada, a to v rozmanitých kombinacích, postupně i souběžně v naději, že mezi mnoha nesprávnými výsledky se objeví i látka, které v analysované směsi skutečně jsou. Vhodnost kritériální funkce (filtru) závisí m.j. také na struktuře použité databáze referenčních difraktogramů. Jestliže se například v dané databázi vyskytuje mnoho referenčních difraktogramů, které mají silnější linie v určitém intervalu mezirovinných vzdáleností bude filtr, jenž třídí difraktogramy podle toho, zda mají (silnější) difrakční linie právě v tomto intervalu či nic, málo účinný. Když filtr přizpůsobíme struktuře referenční databáze, zvýší se účinnost identifikace [63].

Optimální řešení druhé etapy identifikace (MATCH) je to, které hovoří soustavě normálních rovnic (7). Ovšem pouze za předpokladu, že se nám mezi referenční difraktogramy $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_m$ jejichž superpozicí

$$\sum_{i=1}^m c_i \bar{y}_i$$

difraktogram neznámé, určené látky aproximujeme tak, aby

$$\left\| \bar{x} - \sum_{i=1}^m c_i \bar{y}_i \right\| \min \quad (8)$$

podají dostat difraktogramy všech komponent, které analysovaná směs skutečně obsahuje. Což závisí m.j. na úspěšnosti předešlé, první etapy identifikace (které říkáme SEARCH). Úspěch procedury SEARCH však není zaručen; naopak, je velmi pravděpodobné, že difraktogramy některých komponent určené látky procedura SEARCH nenajde. V takovém případě se může více přiblížit pravdě (o identitě určené látky) řešení, které se získá nějak jinak než ze soustavy normálních rovnic (7). Například tak, že se referenční difraktogramy vyhledané procedurou SEARCH nakonec uspořádají sestupně podle hodnoty nějakého kritéria, která je tím větší, čím je referenční difraktogram podobnější difraktogramu určené látky. A v tomto pořadí, tedy v pořadí

$$\bar{y}_1, \bar{y}_2, \dots, \bar{y}_m \quad (10)$$

Tabulka 1. Různé filtry používané při identifikaci fázového složení pomocí rtg difrakce, tj. výrazy, jejichž hodnota je funkcí podobnosti difraktogramu analysované látky a difraktogramu referenční látky.

$\frac{x_h y_h}{\sqrt{x_h^2 + y_h^2}}; \quad \frac{ x_h - y_h }{x_h + y_h}; \quad \frac{1}{n} \frac{ x_h - y_h }{x_h + y_h}$
$\frac{x_h}{\sqrt{x_1}}; \quad \frac{y_h}{\sqrt{y_1}}; \quad \frac{x_h}{\sqrt[3]{x_1}}; \quad \frac{y_h}{\sqrt[3]{y_1}}$
$\frac{x_h}{\sqrt{x_1^2}}; \quad \frac{y_h}{\sqrt{y_1^2}}$
$\frac{x_h y_h}{\sqrt{x_h^2 + y_h^2}}; \quad \frac{x_h}{\sqrt{x_1^2}}; \quad \frac{y_h}{\sqrt{y_1^2}}$
$\frac{x_h}{y_h}; \quad 1 - \frac{x_h y_h}{\sqrt{x_h^2 + y_h^2}}$
$\frac{x_h y_h^2}{x_h^2 + y_h^2}; \quad \frac{x_h \operatorname{sgn} y_h}{x_h}; \quad \frac{x_h 2 \operatorname{sgn} y_h - 1}{x_h}$
$\frac{x_h (1 - \operatorname{sgn} y_h)}{x_h}; \quad \frac{\operatorname{sgn} x_h \operatorname{sgn} y_h}{\operatorname{sgn} x_h \operatorname{sgn} y_h}; \quad \frac{\operatorname{sgn} y_h}{\operatorname{sgn} y_h}$
$\frac{y_h \operatorname{sgn} x_h}{y_h}; \quad \frac{y_h \operatorname{sgn} x_h}{y_h}; \quad \operatorname{sgn} a \quad \begin{matrix} 1 \dots a & 0 \\ 0 \dots a & 0 \end{matrix}$

kde

$$R(\bar{x}, \bar{y}_1) \quad R(\bar{x}, \bar{y}_2) \quad \dots \quad R(\bar{x}, \bar{y}_m) \quad (11)$$

se referenční difraktogramy postupně odčítají od difraktogramu neznámé látky s vahami $c_1^0, c_2^0, \dots, c_m^0$ určenými soustavou rovnic

$$\begin{aligned} (\bar{x} \quad c_1^0 \bar{y}_1, \bar{y}_1) &= 0 \\ (\bar{x} \quad c_1^0 \bar{y}_1 \quad c_2^0 \bar{y}_2, \bar{y}_2) &= 0 \\ (\bar{x} \quad c_1^0 \bar{y}_1 \quad c_2^0 \bar{y}_2 \quad c_3^0 \bar{y}_3, \bar{y}_3) &= 0 \end{aligned} \quad (12)$$

tak, aby příspěvek $c_k^0 \bar{y}_k$ k -tého referenčního difraktogramu \bar{y}_k byl ortogonálním průmětem residuálního vektoru

$$\bar{x} - c_1^0 \bar{y}_1 - c_2^0 \bar{y}_2 - \dots - c_k^0 \bar{y}_k$$

který zůstane po odečtení příspěvků prvních k referenčních difraktogramů, do \bar{y}_k :

$$(\bar{x} \quad c_1^0 \bar{y}_1 \quad c_2^0 \bar{y}_2 \quad \dots \quad c_k^0 \bar{y}_k, \bar{y}_k) = 0$$

Mluvíme o metodě postupného svlékání (stripping) difraktogramu neznámé látky, z kterého „odkládáme kus po kusu“ příspěvky referenčních difraktogramů v pořadí

určeném během první etapy identifikace procedurou SEARCH. V případě, že vektory $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_m$ jsou po dvojicích kolmé, jsou řešení soustav (7) a (12) ekvivalentní. Jinak bude řešení (12) odlišné od řešení (7), a tedy horší než toto řešení, tj

$$\left\| \begin{matrix} \bar{x} \\ c_i \bar{y}_i \end{matrix} \right\|_{i=1}^m \left\| \begin{matrix} \bar{x} \\ c_i \bar{y}_i \end{matrix} \right\|_{i=1}^m$$

za předpokladu, že SEARCH byl úspěšný. Nebylo-li tomu tak, je řešení (7) nesprávné, čímž ztrácí svoji vyjimečnost a může se stát, že řešení (12) bude v nějakém ohledu lepší než řešení (7). Jestliže například SEARCH nenalezne nějakou ze složek analysované směsi, řekněme tu, jejíž difraktogram jest \bar{y}_r , pak to v případě, že druhou etapu identifikace (MATCH) skutečně řešíme soustavou (7), ovlivní velikost vyčíslených příspěvků $c_i \bar{y}_i$; $i=1,2,\dots,m$; i r všech ostatních komponent, zatímco v případě, že MATCH realizujeme jako řešení soustavy (12), ovlivní takové pochybení v první fázi identifikačního procesu (SEARCH) velikost příspěvků $c_i \bar{y}_i$ pouze těch komponent pro něž je $i > r$.

A bude-li uspořádání (11) adekvátní podílům jednotlivých komponent ve směsi, pak to znamená, že budou

zkruseny pouze příspěvky komponent minoritních, jež jsou pro výsledek analýzy méně významné.

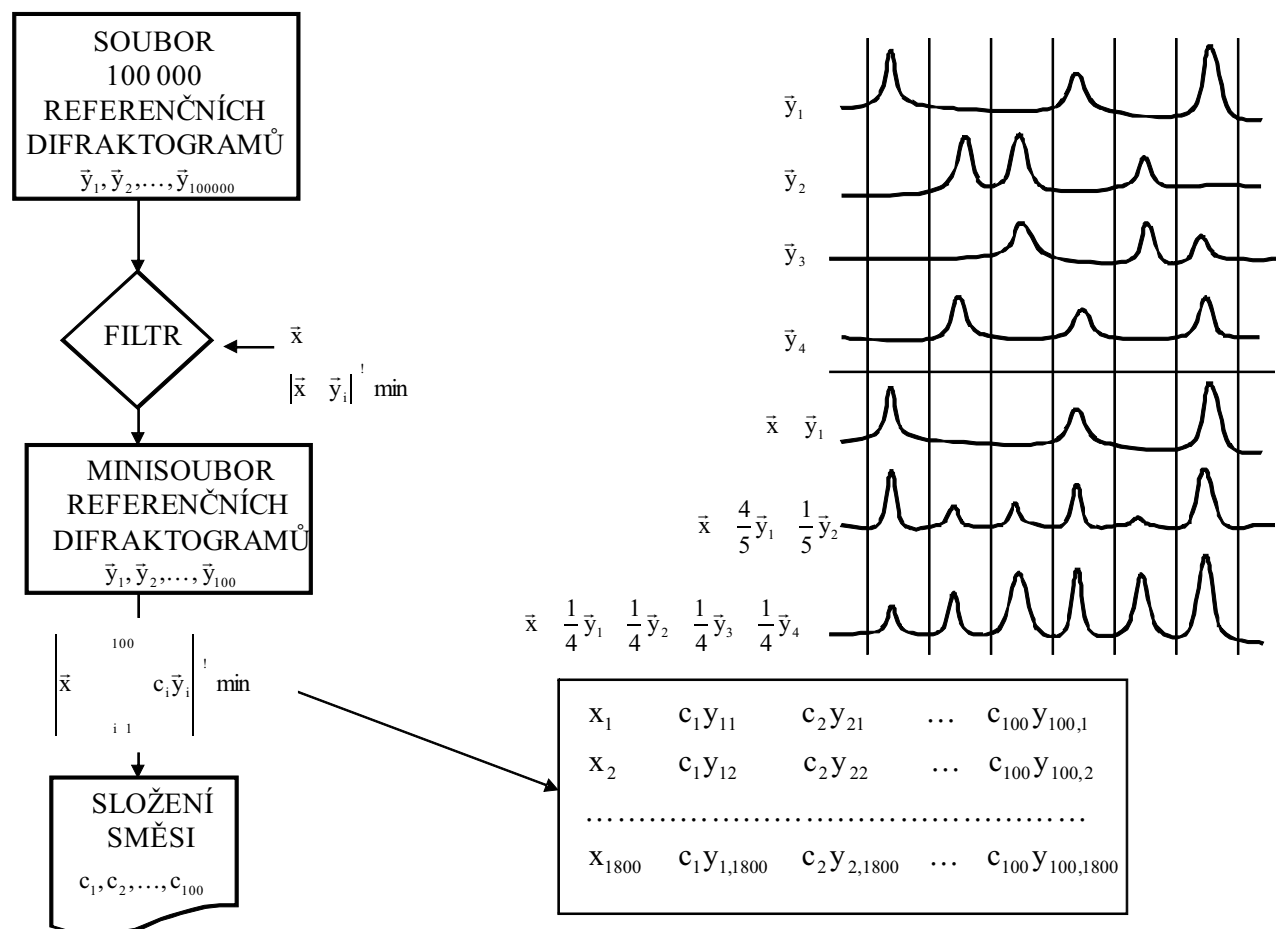
Někdy se „konec“ identifikace zjednodušuje dokonce natolik, že se končí určením pořadí (10) referenčních standardů, jež byly nalezeny procedurou SEARCH. K řešení se pak třeba ještě přiloží hodnoty

$$R(\bar{x}, \bar{y}_1), R(\bar{x}, \bar{y}_2), \dots, R(\bar{x}, \bar{y}_m) \dots \quad (13)$$

kteří mají ilustrovat o kolik je pravděpodobnější, že v neznámé směsi je složka již přísluší referenční difraktogram \bar{y}_1 , než že tam jsou komponenty, jímž přísluší referenční difraktogramy \bar{y}_2 nebo \bar{y}_3 atd. Jedním z filtrů $R(\bar{x}, \bar{y})$, který se takovému „konečnému vypořádání se“ s identifikačním problémem leckdy používá, a který se někdy tradičně označuje jako FOM (figure of merit) [3], jest

$$\text{FOM}(\bar{x}, \bar{y}) = P^2 I^2 (MNY^2)^{-1} \quad (14)$$

kde M = počet linií na difraktogramu \bar{x} ; N = počet linií na difraktogramu \bar{y} ; P = počet linií na difraktogramu \bar{y} , které odpovídají nějaké linii na difraktogramu \bar{x} ; Y = součet intenzit všech linií na difraktogramu \bar{y} ; I = součet intenzit těch linií na difraktogramu \bar{y} , které odpovídají nějaké linii na difraktogramu \bar{x} .



Obrázek 17. Nejslabším, neuralgickým bodem filosofie rtg difrakční kvalitativní fázové analýzy je koncepce, podle které vybíráme referenční difraktogramy ze vztahné databáze při sestavování minisouboru referenčních difraktogramů. Do minisouboru zařazujeme referenční difraktogramy, které se co možná nejvíce podobají difraktogramu analysované látky. To je ovšem kontraproduktivní v případě, že analysovaná směs obsahuje více než několik velmi málo složek, v důsledku čehož se její difraktogram nepodobá difraktogramu žádné z jejích komponent.

Tím se však dostáváme ke koncepci „pravděpodobnosti výskytu“ té které látky ve zkoumané směsi, jež je nepřijatelná: určitá látka se s hlediska uživatele v analysované směsi prostě buď vyskytuje nebo nevyskytuje. Přesto se analytici nezřídka v těchto kategoriích vyjadřují. Nenáležitost takových praktik je dána tím, že k identifikaci směsi se metoda „otisku palce“ vyjádřená sekvencí (13), nehodí.

5. Dimensionální bariera

Identifikace směsi by se neměla řešit metodou „otisku palce“ na základě odděleného porovnávání difraktogramu určované látky \bar{x} s jednotlivými difraktogramy referenční databáze $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_m$ a hledáním takových \bar{y}_i , aby pro ně

$$\|\bar{x} - \bar{y}_i\|$$

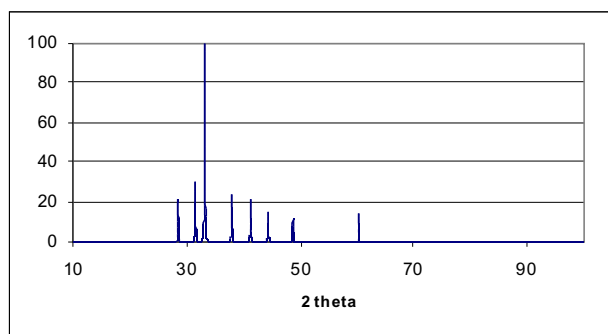
bylo malé, ale regresní analysou, to jest simultánním rozkladem na superposici příspěvků od všech referenčních difraktogramů

$$g(c_1, c_2, \dots, c_m) = \left\| \bar{x} - \sum_{i=1}^m c_i \bar{y}_i \right\|^2 \min \quad (15)$$

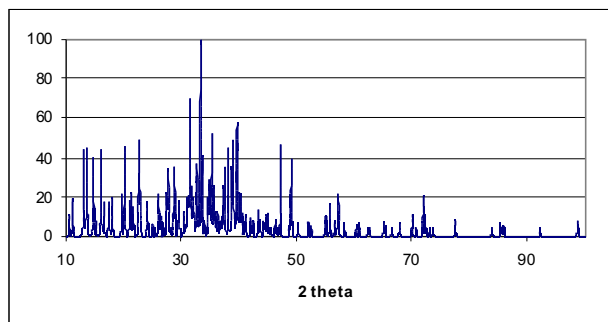
To je korektní matematický model, jenž odpovídá fyzikální realitě kvalitativní fázové analýzy spočívající v rozkladu difraktogramu směsi na superposici difraktogramů ji tvořících komponent. Obtížnost identifikace směsi spočívá v tom, že vyžaduje mnohem více informací než identifikace jednoduchých (jednofázových) látek. Identifikovat neznámou látku pomocí rtg difrakce znamená najít známou látku („standard“), jejíž difraktogram je totožný (identický) s difraktogramem neznámé látky. V případě, že ta neznámá látka je jednoduchá, jedná se o jediný standard (jeden referenční difraktogram). Jestliže však ta neznámá látka je směs, je třeba najít referenční difraktogramy všech jejích komponent. To znamená, že musíme vypočítat hodnoty parametrů c_1, c_2, \dots, c_m pro které funkce $g(c_1, c_2, \dots, c_m)$ z výrazu (15) nabývá minimum: standardy, kterým odpovídají čísla $c_i = 0$ se v neznámé směsi nevyskytují, standardy, kterým odpovídají čísla $c_i > 0$ se v neznámé směsi vyskytují (s koncentrací c_i). Má-li identifikace být proveditelná, je třeba aby databáze, se kterou pracujeme obsahovala referenční difraktogramy všech komponent analysované směsi. V řadě případů nevíme, zdali tomu tak je. A pak to nejlepší, co můžeme udělat, jest použít v dobré víře tu nejobjemnější databázi, která existuje – poslední verzi databáze referenčních difraktogramů ICDD. Tato databáze v současnosti obsahuje kolem 350.000 referenčních difraktogramů, což nás staví před gigantický úkol najít minimum funkce $g(c_1, c_2, \dots, c_m)$, která figuruje ve vztahu (15) a má $m = 350.000$ proměnných c_1, c_2, \dots, c_m . Ta čísla c_i jsou koeficienty lineární superposice

$$\sum_{i=1}^m c_i \bar{y}_i$$

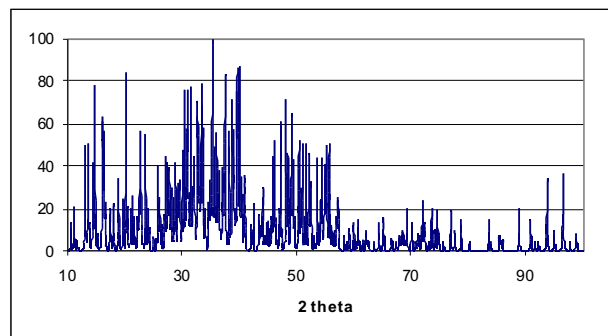
$m = 350.000$ vektorů $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_m$ referenčních difraktogramů, pomocí které máme ve smyslu vztahu (13) difraktogram analysované látky \bar{x} aproximovat



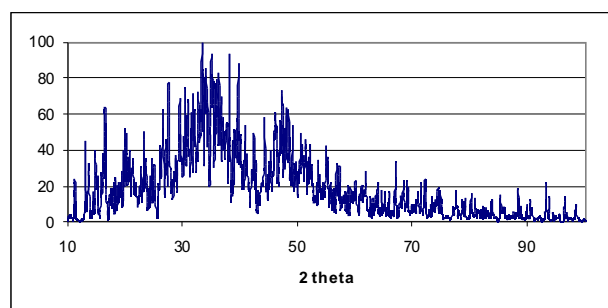
a)



b)



c)



d)

Obrázek 18. Fenomén dimensionální bariery, která se nám postaví do cesty když se pokoušíme identifikovat složitější směsi (směsi většího počtu komponent), ilustrují čtyři syntetické (vypočtené) difraktogramy ekviprocentních směsí n standardů z 30. série databáze Powder Diffraction File obhospodařované ICDD: (a) $n = 1$ – jedná se o referenční difraktogram vedený v databázi pod pořadovým číslem (Powder Diffraction File Number) 30-1; (b) $n = 10$; (c) $n = 20$; (d) $n = 100$.

$$\bar{x} = \sum_{i=1}^m c_i \bar{y}_i; m=350.000 \quad (16)$$

Úloha (14) však má jednoznačné řešení pouze v případě, že vektory

$$\{\bar{y}_1, \bar{y}_2, \dots, \bar{y}_{350.000}\} \quad (17)$$

jsou nezávislé. Existuje však $m = 350.000$ lineárně nezávislých vektorů $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_{350.000}$? Jednoduchá úvaha nám ukáže, že nikoli! Vektorovou reprezentaci difraktogramů jsme zavedli podle koncepce ilustrované obr. 12: používány obor mezirovinných vzdáleností (Braggových úhlů) se rozdělí na dílčí subintervaly, kanály, jimž pak přiřadíme čísla rovná intenzitám difrakčních čar, které svou polohou do příslušného subintervalu padnou. Šířka kanálu nesmí být užší než neurčitost polohy difrakčních čar různých preparátů jedné a téže látky, abychom je mylně nepokládali za látky různé. Kanálů nemůže být tedy nekonečně mnoho: zkušenost ukazuje, že kanály by neměly být užší než $0,1^\circ$, z čehož vychází maximální počet kanálů $n = 1800$. Vektory $\bar{y} [y_1, y_2, \dots, y_n] [y_1, y_2, \dots, y_{1800}]$ jsou tedy řekněme 1800-rozměrné. Mezi vektory $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_{350.000}$ je jich tudíž nezávislých nejvýše 1800. Z toho plyne, že úloha (16) nemá pro $m = 350.000$ jednoznačné řešení. Tato skutečnost je principiální příčinou potíží a nedostatečnosti rtg difrakční identifikace směsí. A v tomto smyslu představuje to číslo 1800 dimensionální bariéru možnostem rtg difrakční kvalitativní fázové analýsy.

Identifikační algoritmy se s tím vypořádají již zmíněnou koncepcí „SEARCH/MATCH“ (obr. 17). Z velkého souboru řekněme 350.000 referenčních difraktogramů nejdříve vybereme „minisoubor“ $r = 100-500$ referenčních difraktogramů

$$\{\bar{y}_1, \bar{y}_2, \dots, \bar{y}_r\} \quad (18)$$

a difraktogram určované látky \bar{x} se pak pokoušíme syntetizovat jako lineární superposici těchto difraktogramů:

$$\bar{x} = \sum_{i=1}^r c_i \bar{y}_i; r = 500 \quad (19)$$

Taková úloha jednoznačně řešitelná je, protože vektory vyskytující se v (19) mají větší počet dimensí než 500 a vzhledem k tomu, že vznikly zpracováním reálných měření, nemohou být prakticky závislé. Řešení úlohy (19) bude však představovat skutečnou identifikaci látky, jejíž difraktogram je zakódován jako \bar{x} pouze v případě, že difraktogramy (18) se podařilo „vyfiltrvat“ z (17) tak, aby se mezi nimi vyskytovaly difraktogramy všech komponent určované látky (směsi) \bar{x} . Na to však dodnes neznáme žádný spolehlivý postup. Filtrování (SEARCH, tedy hledání vhodných kandidátů do (18)) se provádí tak, aby se jejich difraktogramy \bar{y}_i co nejvíce podobaly difraktogramu \bar{x} určované látky

$$\|\bar{x} - \bar{y}_i\| \rightarrow \min \quad (20)$$

Což bude fungovat spolehlivě jenom v případě, že určovaná látka je jednoduchá, tedy tvořená jedinou „komponentou“. Je-li určovaná látka složena z více komponent, bude se difraktogram \bar{y}_i kterékoliv z nich lišit od difraktogramu určované látky \bar{x} , a to tím více, čím více těch komponent je (viz obr. 17). Je-li počet komponent větší než 4-5, začne kritérium (20) fungovat jako kontraproduktivní. Složitější směsi se pomocí rtg difrakce daří identifikovat jenom vyjíměčně.

Jiný aspekt obtížnosti a ošidnosti identifikace složitých směsí ilustruje obr. 18, na němž jsou znázorněny počítačem generované difraktogramy ekviprocentních směsí 10, 20 a 100 komponent. Difraktogramy směsí velkého množství složek vyjadřují rozložení difrakčních linií v souboru difraktogramů těch látek, jež se vzhledem difraktogramu kteréhokoli z nich nemá v zásadě nic společného. Pokud tedy budeme hledat referenční difraktogram podobný difraktogramu tak složité směsi, může snadno dojít k nedorozumění v tom smyslu, že najdeme látku, jejíž difraktogram má s difraktogramem té složité směsi mnoho společných linií přesto, že se ta látka v oné směsi vůbec nevyskytuje. Při rtg difrakční identifikaci složitějších směsí nám nebude příliš užitečná ani globální prvková analýza, protože nám neřekne ve které složce se vyskytuje ten který prvek nalezený ve směsi jako celku. Mnohem lepší je elektronová mikroanalýza prvkového složení jednotlivých částic (strukturních složek). Velice pomůže také separace analysované směsi na jednotlivé složky, jejichž identifikace pomocí rtg difrakce pak už nečiní žádnou potíž.

Možnosti lokální mikroanalýsy jakož i separace směsí jsou však omezené a nemůžeme se na ně spoléhat úplně a vždycky. Z nejrůznějších důvodů se pomocí těchto technik někdy nepodaří rozlišit (oddělit) jednotlivé komponenty dokonale. Tak se například může stát, že úzký elektronový svazek mikrosondy, budící charakteristické rtg záření, jež nese zprávu o prvkovém složení, zasáhne nejenom jednu částici, která nás zajímá, ale pronikne ještě i pod ni. Zde pak může vybudit záření nějaké jiné částice, „schované“ pod tou částicí na kterou jsme svazek zaměřili. A pokud je shodou okolností tato částice tvořena nějakou jinou složkou (fází) analysované směsi než ta terčová (na povrchu situovaná) částice, nese emitované rtg záření informaci o prvkovém složení obou složek a nikoliv jenom o prvkovém složení složky jediné. A nejhorší na té skutečnosti jest, že o tom nevíme, takže naše uvažování o komponentách, jež by zkoumaná složitá směs mohla obsahovat, to bude mást.

Jediný postup, na který se můžeme při překonávání dimensionální bariéry rtg difrakční kvalitativní fázové analýsy obecně spolehnout, spočívá v kombinaci této metody se separačními postupy a statistickým zpracování výsledků, jež touto komplexní procedurou získáme pomocí faktorové analýsy.

6. Faktorová analýza

Dá se také říci, že příčina všech zásadních potíží, na které při kvalitativním fázovém rozboru směsí narážíme jest, že nevíme které difrakční linie analysované směsi odpovídají jejím jednotlivým složkám. Kdybychom to věděli, identifikovali bychom každou složku zvlášť, s čímž žádné problémy nejsou: identifikace jednoduchých (jednofázových) látek na základě porovnání jejich difraktogramu s databází referenčních difraktogramů je účinná a spolehlivá. Abychom poznali která difrakční čára odpovídá jaké složce, musíme difraktogram „rozhrbat“. Nemůžeme zůstat u analýzy dané neznámé směsi, ale musíme tuto směs zpracovat, abychom z ní dostali několik preparátů, jež se od sebe budou lišit kvantitativním zastoupením jednotlivých komponent. Takto vyrobené preparáty jsou vlastně frakce původní směsi, obohacené resp. ochuzené o její jednotlivé komponenty. Jestliže například v nějaké frakci bude určité komponenty méně, všechny její difrakční linie se současně a ve stejném poměru zeslabí, čímž se „prozradí“, že patří k jedné komponentě.

Situaci však komplikuje dále ještě druhá zásadní potíž, se kterou se při identifikaci směsi setkáváme, totiž to, že na jejich difraktogramu je velké množství linií: každá složka má určitý počet difrakčních linií a čím více složek ta směs má, tím více linií bude na jejím difraktogramu. A tím více těch linií se bude překrývat. Překryté linie přinášejí však méně informace než difrakční linie nepřekryté, solitérní. Jestliže např. do nějaké difrakční linie přispějí třeba dvě komponenty, pak ochuzení směsi o jednu z nich nezpůsobí poměrné zeslabení oné difrakční linie: podle toho jak se ve směsi změnil podíl druhé komponenty, může se intenzita té difrakční linie dokonce i zvýšit. Bude-li většina difrakčních linií překrytých, může se stát, že o tom, která difrakční linie patří té či oné komponentě, se nedozvíme téměř nic. Bohudíky má to, že na difraktogramu složité směsi se vyskytuje hodně difrakčních linií, nejen zmíněné nepříznivé důsledky, ale i jeden důležitý důsledek příznivý: možnost využití statistických metod vyhodnocování. Statistické metody jsou účinné pouze v případě, že zpracovávané soubory jsou velké; pro malý počet položek nefungují. Je to jako s přímými metodami strukturní analýzy [64-66]: tyto statistické metody jsou účinné jenom proto, že se uplatňují na stovky a tisíce reflexí. Abychom však mohli kvantitativně vyhodnotit příspěvky jednotlivých složek, jejichž difrakční linie se překrývají, musíme vědět kolik je které komponenty ve směsi (nebo její frakci). Předpokladem pro provedení kvalitativní fázové analýzy je tedy provést nejdříve fázovou analýsu kvantitativní. Při konvenčním pojetí fázové analýzy je však pro to, abychom mohli stanovit kolik je které fáze ve vzorku, nezbytné předem vědět o jaké fáze se jedná, tedy vědět jaké fáze ve vzorku jsou. Čímž se kruh uzavírá. A to je příčinou toho, proč identifikace složitějších směsí selhává. Dostat se z tohoto začarovaného kruhu ven je možno jediným způsobem: tak, že se kvalitativní a kvantitativní fázové složení určí současně. Říkáme tomu bilineární programování [67]: difraktogram zkoumané směsi analyzujeme, tedy rozložíme na příspěvky $c_i \bar{y}_i$ jeho komponent (řekněme, že je jich k)

$$\bar{x} = \sum_{i=1}^k c_i \bar{y}_i \quad (21)$$

a tyto komponenty pak identifikujeme na základě porovnání $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_k$, jež jsme získali řešením (21), s databází referenčních difraktogramů. Výraz na pravé straně rovnice (21) je lineární v c_i i v \bar{y}_i (odtud název „bilineární programování“) a všechny tyto veličiny, difraktogramy $\bar{y}_1, \bar{y}_2, \dots, \bar{y}_k$ komponent analysované směsi i jejich koncentrace c_1, c_2, \dots, c_k musíme vypočítat z difraktogramu té analysované směsi \bar{x} . To arci není jednoduchý úkol a pro jeho řešení byla vyvinuta řada zvláštních postupů [68]. Standardní a ve statistice dobře zpracovanou metodou řešení této úlohy jest faktorová analýza [69-74]. „Faktory“ (tedy tím, co se faktorovou analýzou snažíme najít) se v tomto případě rozumí vektory

$$\{\bar{y}_1, \bar{y}_2, \dots, \bar{y}_k\} \quad (22)$$

Faktory hledáme rozбором difraktogramů

$$\{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_k\} \quad (23)$$

p různých frakcí, které se musí ze zadaného preparátu (jehož difraktogram \bar{x} je jedním ze souboru difraktogramů (23)) vyrobit separací. Ty difraktogramy (22) jsou faktory difraktogramů všech frakcí (23):

$$\bar{x}_l = \sum_{i=1}^k c_{li} \bar{y}_i, \quad l = 1, 2, \dots, p \quad (24)$$

Pokud jsou tyto difraktogramy

$$\bar{x}_i = [x_{i1}, x_{i2}, \dots, x_{in}], \quad i = 1, 2, \dots, p$$

nezávislé, bude jejich Gramova matice

$$\hat{G} = \begin{pmatrix} (\bar{x}_1, \bar{x}'_1) & (\bar{x}_1, \bar{x}'_2) & \dots & (\bar{x}_1, \bar{x}'_p) \\ (\bar{x}_2, \bar{x}'_1) & (\bar{x}_2, \bar{x}'_2) & \dots & (\bar{x}_2, \bar{x}'_p) \\ \dots & \dots & \dots & \dots \\ (\bar{x}_p, \bar{x}'_1) & (\bar{x}_p, \bar{x}'_2) & \dots & (\bar{x}_p, \bar{x}'_p) \end{pmatrix}$$

regulární a její vlastní vektory

$$\{\bar{q}_1, \bar{q}_2, \dots, \bar{q}_n\} \quad (25)$$

vytvoří úplný ortogonální systém. Koefficienty Fourierovy řady (rozkladu) [40] difraktogramů frakcí analysované směsi (23) vzhledem k soustavě vektorů (25) jsou úměrné druhé odmocnině příslušného vlastního čísla

$$\bar{x}_l = \sum_{i=1}^n v_{li} \bar{q}_i = \sum_{i=1}^n u_{li} \sqrt{\lambda_i} \bar{q}_i; \quad l = 1, 2, \dots, p \quad (26)$$

kde

$$\hat{G} \bar{q}_i = \lambda_i \bar{q}_i; \quad i = 1, 2, \dots, n.$$

Uspořádáme-li ta vlastní čísla sestupně

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k > \lambda_{k+1} \geq \dots \geq \lambda_n$$

bude prvních k vlastních čísel odpovídat „chemické“ variabilitě vektorů (23), podmíněné k faktory (22). Zbylých $n-k$ vlastních čísel odpovídá pak chybám měření a nedokonalosti separačního procesu, jímž byly frakce (23) izolovány a v tomto smyslu i co do své velikosti budou vlastní čísla k_1, k_2, \dots, k_n insignifikantní. Je tedy

$$c_{li} \bar{y}_i \quad \bar{x}_l \quad c_{lj} \bar{q}'_j, \quad l = 1, 2, \dots, p$$

a tudíž

$$\bar{y}_i = \sum_{j=1}^k T_{ij} \bar{q}'_j; \quad i = 1, 2, \dots, k \quad (27)$$

Pro určení difraktogramů (22) komponent analysované směsi zbývá vypočítat prvky matice

$$\hat{T} = \begin{pmatrix} T_{11} & T_{12} & \dots & T_{1k} \\ T_{21} & T_{22} & \dots & T_{2k} \\ \dots & \dots & \dots & \dots \\ T_{k1} & T_{k2} & \dots & T_{kk} \end{pmatrix} \quad (28)$$

jež transformuje signifikantní vlastní vektory $\{\bar{q}'_1, \bar{q}'_2, \dots, \bar{q}'_k\}$ na difraktogramy $\{\bar{y}_1, \bar{y}_2, \dots, \bar{y}_k\}$. K tomu stačí využít toho, že intenzity difrakčních čar komponent analysované směsi ani jejich koncentrace nemohou být záporné. Jinými slovy, že musí být

$$\hat{Y} \hat{T} \hat{W} \hat{O} \quad (29)$$

$$\hat{C} \hat{X} \hat{W}' \hat{T}' (\hat{T} \hat{W} \hat{W}' \hat{T}')^{-1} \hat{O} \quad (30)$$

kde čárkou vyznačujeme transpozici příslušné matice,

$$\hat{Y} = \begin{pmatrix} \bar{y}_1 & y_{11} & y_{12} & \dots & y_{1n} & \bar{q}'_1 \\ \bar{y}_2 & y_{21} & y_{22} & \dots & y_{2n} & \hat{T} \bar{q}'_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \bar{y}_k & y_{k1} & y_{k2} & \dots & y_{kn} & \bar{q}'_k \end{pmatrix} \quad \hat{T} \hat{W}$$

a

$$\hat{X} = \hat{C} \hat{Y} \quad \bar{x}_l = \sum_{i=1}^k c_{li} \bar{y}_i; \quad l = 1, 2, \dots, p$$

(tzv. prokrustovské výpočetní metody).

Tento zcela obecně použitelný postup při kvalitativní fázové analýze (jakkoli složitých) směsí je zároveň nejdokonalější technikou fázové analýzy kvantitativní [75, 76]. Hlavním zdrojem neurčitosti výsledků kvantitativní fázové analýzy je totiž reálná struktura materiálu (velikost, tvar a orientace krystalků jakož i jejich nejrušnější vnitřní defekty), jež nezanedbatelnou měrou ovlivňuje intenzity difrakčních linií a je u každého preparátu obecně jiná. Jestliže z analysované směsi vyseparujeme řadu frakcí tak, aby se přitom nezměnila reálná struktura, můžeme faktorovou analýzou difraktogramů těchto frakcí vypočítat intenzity difrakčních linií jednotlivých složek směsi tak, že se přitom vliv reálné struktury vezme v úvahu a neurčitost kvantitativního stanovení fázového složení té směsi se

sníží. Aplikovatelnost jedné a téže statistické metody vyhodnocování (faktorové analýzy) jak pro kvalitativní tak i pro kvantitativní fázovou analýzu odráží skutečnost, že kvalitativní a kvantitativní stanovení jsou jenom dva různé a v podstatě neoddělitelné aspekty fázové analýzy, jež dlužno pojímat v jednotě. A že konservativní lpění na wagnerovském schématu „nejdříve určit kvalitu a teprve potom kvantitu“ je pro fázovou analýzu neštěstím a skutečnou příčinou jejího současného neutěšeného stavu (neurčitost kvantitativního fázového složení se nedaří zredukovat pod onu magickou hranici, o níž psal André Guinier už v roce 1938 že se pohybuje kolem 1-10% absolutních a spolehlivá a bezproblémová identifikace kvalitativního složení směsí obsahujících více než 5 komponent zůstává stále ještě nesplněným snem).

Literatura

- [1] H.P. Peiser, H.P. Rooksby, A.J.C. Wilson (eds): Diffraction by Polycrystalline Materials, London 1960, Chapman and Hall.
- [2] D.L. Bish, J.E. Post (eds): Modern Powder Diffraction, Washington D.C. 1989, The Mineralogical Society of America.
- [3] R. Jenkins, R.L. Snyder: Introduction to X-ray Powder Diffractometry, New York 1996, John Wiley and Sons.
- [4] J. Fiala, Československý časopis pro fyziku A24 (1974) 237-245.
- [5] J. Zýka (ed): Analytická příručka, Praha 1975, SNTL.
- [6] J. Fiala, Slévárenství 32 (1984) 35-39.
- [7] V. Kalous (ed): Metody chemického výzkumu, Praha 1987, SNTL.
- [8] J. Fiala, Vesmír 67(1988)337-344.
- [9] W. Ch. Kidd, P. Varlashkin, Ch. Li: Powder Diffraction 8(1993) 180-187.
- [10] L.K. Frevel, Industrial and Engineering Chemistry Analytical Edition 16(1944) 209-218.
- [11] A.W. Waldo: American Mineralogist 20(1935) 575.
- [12] A.K. Boldyrev, V.I. Michejev, G.A. Kovalev, V.N. Dubinina: Zapiski Leningradskogo Gornogo Instituta imeni G.V. Plechanova 13(1939), 1.
- [13] L.K. Frevel, C.E. Adams, L.R. Ruhberg, Journal of Applied Crystallography 9(1976), 199-204.
- [14] J.D. Hanawalt, H.W. Rinn, L.K. Frevel: Industrial and Engineering Chemistry Analytical Edition 10(1938)457-512.
- [15] W.J. Moore: Physical Chemistry, Englewood Cliffs, New Jersey 1972, Prentice-Hall, Inc.
- [16] F.W. Matthews, Analytical Chemistry 21(1949), 1172-1175.
- [17] R. Jenkins, M. Holomany: Powder Diffraction 2(1987), 215-219.
- [18] R. Jenkins, M. Holomany, W. Wong-Ng: Powder Diffraction 2(1987), 84-87.
- [19] J. Fiala, Archiwum Nauki o Materialach 12(1991) 85-99.
- [20] J. Fiala, J. Šesták: Journal of Thermal Analysis and Calorimetry 60(2000), 1101-1110.



- [21] V.I.Michejev: Rentgenometričeskij opredělitel' mineralov, Moskva 1965, Gosgeol'tečizdat.
- [22] J.D.Hanawalt, H.W.Rinn, Industrial and Engineering Chemistry Analytical Edition 8(1936)244-250.
- [23] R.Jenkins,R.N.Rose, Powder diffraction 5(1990)70-75.
- [24] L.K.Frevel, Analytical Chemistry 37(1965)471-481.
- [25] G.G.Johnson, V.Vand, Industrial and Engineering Chemistry 59(1967)no.8,19-31.
- [26] E.K.Vasiljev, M.S.Nachmanson: Kačestvennyj rentgenofazovyj analiz, Novosibirsk 1986, Nauka.
- [27] D.K.Smith, S.Gorter, Journal of Applied Crystallography 24(1991)369-402.
- [28] T.L.Isenhour, Analytical Chemistry 45(1973)2153-2154.
- [29] B.H.O'Connor, F. Bagliani, Journal of Applied Crystallography 9(1976)419-423.
- [30] W.N.Schreiner, C.Surdukowski, R.Jenkins, Journal of Applied Crystallography 15(1982)513-523.
- [31] B.H.Toby, R.L.Harlow, M.A.Holomany, Powder Diffraction 5(1990)2-7.
- [32] M.Otto, Analytica Chimica Acta 283(1993)500-507.
- [33] B.Adler, P.S.Schütze, J.Will, Analytica Chimica Acta 271(1993)287-291.
- [34] D.Wright, C.L.Lin, D. Stanley, H.C.Chen, J.H.Fang, Computers and Geosciences 19(1993)1429-1443.
- [35] W.N.Schreiner, C.Surdukowski, J.Jenkins, Journal of Applied Crystallography 15(1982)605-610.
- [36] L.K.Frevel, Analytical Chemistry 54(1982)691-697.
- [37] D.H.Filsinger, L.K.Frevel, Powder Diffraction 1(1986),no.1,22-25.
- [38] S.D.Kirik, S.A.Kovyazin, A.M.Fedotov, Powder Diffraction 5(1990)181-185.
- [39] J.Fiala, Silikáty 31(1987)173-191.
- [40] A.Kufner: Geometrie Hilbertova prostoru, Praha 1973, SNTL.
- [41] A.N.Kolmogorov, S.V.Fomin: Základy teorie funkcí a funkcionální analýzy, Praha 1975, SNTL.
- [42] H.T.Clifford, W.Stephenson: An Introduction to Numerical Classification, New York 1975, Academic Press.
- [43] P.R.Sokal, P.H.A.Sneath: Principles of Numerical Taxonomy, London 1963, Freeman.
- [44] N.Jardine, R.Sibson: Mathematical Taxonomy, New York 1971, John Wiley and Sons.
- [45] N.J.Nilsson: Learning Machines, New York 1965, McGraw-Hill Book Company.
- [46] V.I.Vasiljev: Razpoznajuščiye sistěmy, Kiev 1969, Naukova Dumka.
- [47] P.C.Jurs, T.L.Isenhour: Chemical Applications of Pattern Recognition, New York 1975, John Wiley and Sons.
- [48] K.S.Fu, ed.: Digital Pattern Recognition, Berlin 1976, Springer-Verlag.
- [49] R.C.Tryon, D.E.Bailey: Cluster Analysis, New York 1970, McGraw-Hill Book Company.
- [50] M.R.Anderberg: Cluster Analysis for Applications, New York 1973, Academic Press.
- [51] J.A. Hartigan: Clustering Algorithms, New York 1975, John Wiley and Sons.
- [52] T. Cacoullos, ed.: Discriminant Analysis and Applications, New York 1973, Academic Press.
- [53] L.Collatz: Funktionalanalysis und numerische Mathematik, Berlin 1964, Springer-Verlag.
- [54] J.Fiala, M.Kolega, Materials Structure in Chemistry, Biology, Physics and Technology 7(2000)58-65.
- [55] M.Maňas: Optimalizační metody, Praha 1979, SNTL.
- [56] J.Fiala, Journal of Physics D: Applied Physics 5(1972) 1874-1876.
- [57] J.Fiala, Journal of Applied Crystallography 9(1976) 429-432.
- [58] J.Fiala, Analytical Chemistry 52(1980) 1300-1304.
- [59] B.L.Davis, R.Kath, M.Spilde, Powder Diffraction 5(1990)76-78.
- [60] R.L.Snyder, Powder Diffraction 7(1992) 186-192.
- [61] J.Fiala, Zavodskaja laboratorija 54(1988),no.3,32-36.
- [62] K.S.Fu, ed.: Syntactic Pattern Recognition Applications, Berlin 1977, Springer-Verlag.
- [63] M.S.Nachmanson, V.K.Fekličev: Diagnostika sostava materialov rentgenodifrakcionnymi i spektralnymi metodami, Leningrad 1990, Mašinstrojenije.
- [64] A.I.Kitajgorodskij: Teorija strukturnogo analiza. Moskva 1957. Izdatel'stvo akademii nauk SSSR.
- [65] H.A.Hauptman: Crystal Structure Determination. New York 1972. Plenum Press.
- [66] M.F.C.Ladd & R.A.Palmer: Theory and Practice of Direct Methods in Crystallography. New York 1980. Plenum Press.
- [67] H.Martens & T.Naes: Multivariate Calibration, Chichester 1989. John Wiley.
- [68] J.Fiala: Identification of x-ray diffraction patterns of multicomponent mixtures. In X-ray and Neutron Structure Analysis in Materials Science, J.Hašek (ed.), New York 1989, Plenum Press, 3-13.
- [69] E.R.Malinowski & D.G.Howery: Factor Analysis in Chemistry. New York 1980. John Wiley.
- [70] R.P.McDonald: Faktorová analýza a příbuzné metody v psychologii. Praha 1991. Academia.
- [71] P.Blahuš: Faktorová analýza a její zobecnění. Praha 1985. SNTL.
- [72] K. Überla: Faktorová analýza. Bratislava 1976. Alfa.
- [73] H.H.Harman: Modern Factor Analysis. Chicago 1967. The University of Chicago Press.
- [74] R.J.Rummel: Applied Factor Analysis. Evanston 1970. Northwestern University Press.
- [75] J.Fiala, Materials Science Forum 79-82(1991) 27-34.
- [76] L.S.Zevin & G.Kimmel: Quantitative X-ray Diffractometry. New York 1995. Springer-Verlag.

The first version of the paper was submitted for the planned textbook on X-ray analysis already in 2003. Then revised version was written in 2009 (eds.)