

RESEARCH

Open Access



Application of attenuated total reflectance—Fourier transform infrared spectroscopy—(ATR-FTIR) and principal component analysis (PCA) in identification of copying pencils on different paper substrates

Małgorzata Grzelec^{1,2*}

Abstract

Presence of copying pencils in heritage objects poses a significant challenge for conservators due to their proneness to fading, sensitivity to solvents and difficulties in differentiation from regular graphite pencils. In this paper a method of copying pencils identification by means of ATR—FTIR spectroscopy is used. A protocol for spectra processing and dimensionality reduction of spectral data by means of principal component analysis has been developed, allowing for pencil types differentiation and providing an easy to read visual representation of the collected data. The protocol has been developed on mock-up samples and tested on objects from the archives of the State Museum Auschwitz – Birkenau.

Keywords Copying pencils, Aniline dyes, FTIR, PCA, Methyl Violet, Methylene Blue, Graphite

Introduction

The first graphite pencils appeared in the sixteenth century and the first copying pencils emerged at the turn of the nineteenth and twentieth centuries and achieved considerable popularity during that time period. Comprising a graphite lead with an admixture of water-soluble aniline dye, these pencils were invented for the purpose of producing written document copies quickly, hence their most commonly used names of “copying pencils” or “copy pencils”. An alternative name, “indelible pencils”, points to their other common application as precursors to ball-point pens [1]. As their use quickly became widespread,

and the marks they left were often indistinguishable from those of graphite pencils, copying pencils were often used as graphite pencils replacements. Such liberal applications make it difficult to predict whether a pencil of the period is a copying pencil or not, which can have implications during performing conservation treatments and designing preservation policies for an object.

The emergence of copying pencils was made possible by the invention of mauveine dye, which was synthesized from coal tar by William Perkin in 1856. This chemical launched a new era of mass-produced synthetic dyes and the subsequent development of the aniline dye class [2]. Aniline dyes are derivatives of the simplest aromatic amine, the aniline molecule, which consists of an amino group attached to a phenyl ring. Aniline dyes constitute a broad class of dyes that are differentiated into separate classes based on their characteristic structures. These classes include triphenylmethane dyes, xanthene dyes or azo dyes, among

*Correspondence:

Małgorzata Grzelec
malgorzata.grzelec@uni-hamburg.de

¹ Center for the Study of Manuscript Cultures, Hamburg University, Hamburg, Germany

² Auschwitz – Birkenau State Museum, Oświęcim, Poland

others. Aniline dyes impart color due to presence of polycyclic conjugated bond systems and unsaturated groups acting as chromophores, and various functional groups acting as auxochromes. Aniline dyes quickly dominated the colorant market and became omnipresent in various applications, such as textile industry [3, 4], leather dyeing [5], papermaking [6, 7], printing inks [8], lake pigments [9], and production of colored and copying pencils [10].

Technology and chemistry of copying pencils

The first copying pencils emerged in the 1870s [11]. The technology used to manufacture copying pencils was similar to that used in production of graphite pencils. The pencil lead in graphite pencils was formed by mixing graphite and clay in varying proportions, sometimes with small admixtures of other substances, such as waxes. In the nineteenth and early twentieth centuries, usually kaolin was used as a clay extender in pencils [12]. Kaolin refers to a mineral rich in kaolinite clay ($\text{Al}_2\text{Si}_2\text{O}_5(\text{OH})_4$), composed of silica and alumina sheets. The different proportions of clay and graphite are responsible for differences in pencil hardness. The more clay the pencil lead contains, the harder it is, resulting in less graphite being deposited on the paper surface during writing and, as a result, a less intense pencil mark being produced. In copying pencils, clay extender, dyestuff and graphite were mixed in varying proportions (Table 1). If the graphite was not present at all, the formulation would result in an aniline dye-based colored pencil. Colored aniline pencils were primarily based on kaolin clay extender and aniline dye, often with small additions of waxes such as paraffin wax, carnauba wax, or beeswax, and binders such as vegetable gums, resulting in a softer lead [13]. Since the colored pencils did not contain any graphite, they left a colored trace when writing on dry paper.

Aniline dyes used in copying pencils were available in a range of colours. Violet and blue dyes were much more popular than other colours, owing to brilliant, vivid hues of the compounds used in their production

and thermoplastic properties of the dyes in solid form that made them suitable lead binders [10].

Pencil lead in violet copying pencils usually contained Methyl Violet (MV) dye, synthesized in 1862 by Charles Lauth, and sold since under the name 'Violet de Paris'. MV belongs to a class of triarylmethane dyes, characterized by presence of three aryl groups, each connected to the central carbon atom. The name Methyl Violet refers to a group of compounds derived from a pararosanine precursor on the way of nitrogen methylation. The homologues present in the MV mixture are chloride salts of tetra-methyl, penta-methyl, and hexa-methyl pararosanilines (Table 2). The hexa-methyl pararosanine is often singled out and referred to by the name of Crystal Violet [14].

Blue copying pencils most often contained Methylene Blue (MB) dye, but other blue aniline dyes such as Patent Blue V were also used sometimes, as well as blue pigments such as Prussian Blue [10, 12] (Table 2). MB was synthesized in 1876 by Heinrich Caro [15] in the form of methylthionium chloride salt. MB is a polymethine thiazine dye characterized by tricyclic structure with one N-methylated group on each of the two outer rings.

Dyes in the copying pencils formulations could be sometimes mixed as well. In some recipes two aniline dyes or an aniline dye with an inorganic pigment, such as Prussian Blue, could be mixed to obtain new shades of copy pencils [12].

Copying pencils preservation

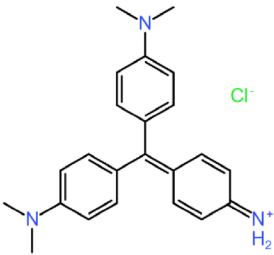
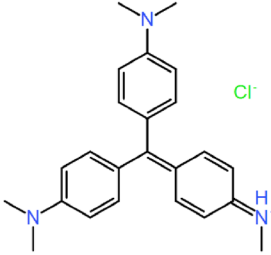
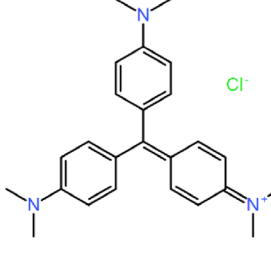
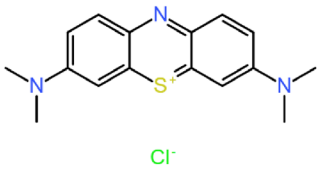
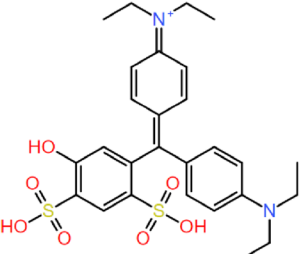
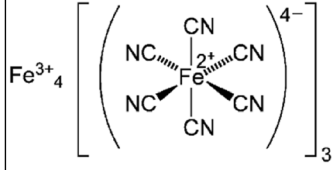
The poor stability of copying pencils is recognized by heritage conservation community and poses a significant challenge when designing conservation treatments and exhibition conditions of objects containing copying pencils. Due to the aniline dye content copy pencils show poor lightfastness with light sensitivity classified as very high, comparable to Blue Wool 1 level on the Blue Wool lightfastness scale [16]. The mechanism of color fading of both MV and MB has been a topic of research. In both cases the process consists in light induced oxidative degradation, where the first step involves N-demethylation, followed by photo-oxidative cleavage of the central C-phenyl bond and photo-reduction of an excited state dye cation to a leuco form for MV [17] and for MB by breaking of the central aromatic ring and the side aromatic rings leading to production of a range of reaction intermediates and their further decomposition [18].

In addition to sensitivity to light, copying pencils show sensitivity to solvents. Although copying pencils have been designed to be readily soluble, their solubility is problematic in the context of conservation treatments, because introducing humidity or solvents to an object containing unidentified copy pencil marks can lead to

Table 1 Recipes for copying pencil lead formulations by Faber, 1899 (Dobruskin 1999)

Pencil type	Graphite	Clay extender	Aniline dye
Soft	37.5%	12.5%	50%
Medium	34%	24%	46%
Hard	30%	40%	30%
Very hard	24%	50%	25%

Table 2 Structural formulas of dyes and pigments commonly used in manufacturing of copying pencils

Violet dyes and pigments		
		
Methyl Violet 2B	Methyl Violet 6B	Crystal Violet
Blue dyes and pigments		
		
Methylene Blue	Patent Blue V	Prussian Blue

bleeding of the dyes. Depending on the functional groups aniline dyes can be sensitive to different range of solvents. The dyes commonly used in copying pencils, MV and MB, are polar in character and thus show more sensitivity to polar solvents, such as water, ethanol or acetone, and are less soluble in nonpolar solvents such as toluene [11]. The property of aniline dye solubility is the basis of one of the most common tests used in identification of copying pencils in heritage objects. The test consists in dripping a drop of solvent onto a pencil marking and subsequent blotting, which in case of presence of an aniline dye, leads to dye transfer onto the blotting paper. The test is commonly used by conservation professionals,

but proves challenging due to the risk of bleeding of the diluted dye into the original paper, resulting in smudging and blurring of the pencil lines. In order to minimize the risks resulting from the test, an alternative method of non-invasive copy pencils identification is proposed.

Documents of the SS - Neubauleitung in the State Museum Auschwitz-Birkenau—history, technology and preservation

The State Museum Auschwitz-Birkenau encompasses the buildings, objects, and documents secured after the liberation of the Nazi extermination camp on January 27, 1945. The majority of the documents are stored in the

Table 3 List of pencils and papers used in the mock-up samples

Symbol	Manufacturer	Main components	Color
Pencils			
1	Omnium Cedrus (V)	MB dye, kaolin clay	Blue
4	ZPMB Pruszków PIAST (V)	MV dye, graphite, kaolin clay	Violet
6	7,9,60 (V)	MV dye, graphite, kaolin clay	Violet
8	Polonia 200 nr 2 (V)	MV dye, graphite	Blue
9	270 DEMON ST MAJEWSKI (V)	MB dye, graphite	Blue
10	2255 TREFLE ST MAJEWSKI (V)	MV dye, graphite	Violet
11	KRAK LAJKONIK (V)	MV dye, graphite	Violet
12	PIAST No 283 (V)	MV dye, graphite	Violet
14	MIKADO (V)	MV dye, graphite	Violet
A	Koh-I-Noor Hardtmuth 1500	Graphite	–
B	Grand 3B	Graphite	–
C	Stabilo Othello 2988	Graphite	–
D	Paperminz	Graphite, kaolin clay	–
E	DERWENT GRAPHIC	Graphite	–
F	CENTRUM HB	Graphite, kaolin clay	–
G	STAEDTLER	Graphite	–
H	LYRA 38	Graphite	–
I	ROCHE	Graphite	–
Papers			
I	Unknown	Softwood and hardwood chemical fibers	Blue
II	Fabriano	Softwood and hardwood chemical fibers	Cream
III	Unknown	Hardwood chemical fibers	Yellow with blue fibers
IV	Unknown	Softwood and hardwood chemical fibers	White
V	TSC	Chemical softwood fibers and mechanical wood fibers	Gray

Vintage materials were indicated with letter “V”

Museum Archives and include original Nazi administrative documentation, source materials created after the war (memoirs, accounts of former prisoners, documentation of Nazi crime trials), photographs, microfilms, negatives, documentary movies, and a variety of historical research publications.

The documents examined within this project constitute a part of the documentation issued by the New Construction Management SS in Auschwitz (SS-Neubauleitung in Auschwitz), an organizational unit that managed the camp’s construction works. This unit was formed in November 1941 after Heinrich Himmler’s decision to launch a concentration camp in the Polish city of Oświęcim [19]. The testing of writing media was performed on eight sheets of correspondence from a folder comprising documents issued between September 1941 and January 1942, concerning demands for prisoner workforce.

Majority of the documents in the folder was produced on thin sheets of typewriter paper, often referred to as

onion skin. The type of paper usually consisted of either cotton fibers or bleached chemical wood pulps, or their combination [20]. The type of paper could be sized with starch or rosin, and oftentimes calendred. While strong, yet thin and translucent to a varying degree, onion skin papers were used together with carbon paper sheets to produce printed text by means of a typewriter. Most of the printed documents in the SS Neubauleitung folder were manually filled in and signed using different writing media, majority of which were various kinds of pencils. The documents did not undergo any previous conservation treatments and their condition can be classified as good.

ATR-FTIR

Attenuated Total Reflection–Fourier Transform Infrared Spectroscopy is a tool well established in non-destructive analysis of materials in heritage objects. In ATR-FTIR the sample is pressed to the surface of a refractive crystal. An IR laser beam is directed through the crystal, undergoing

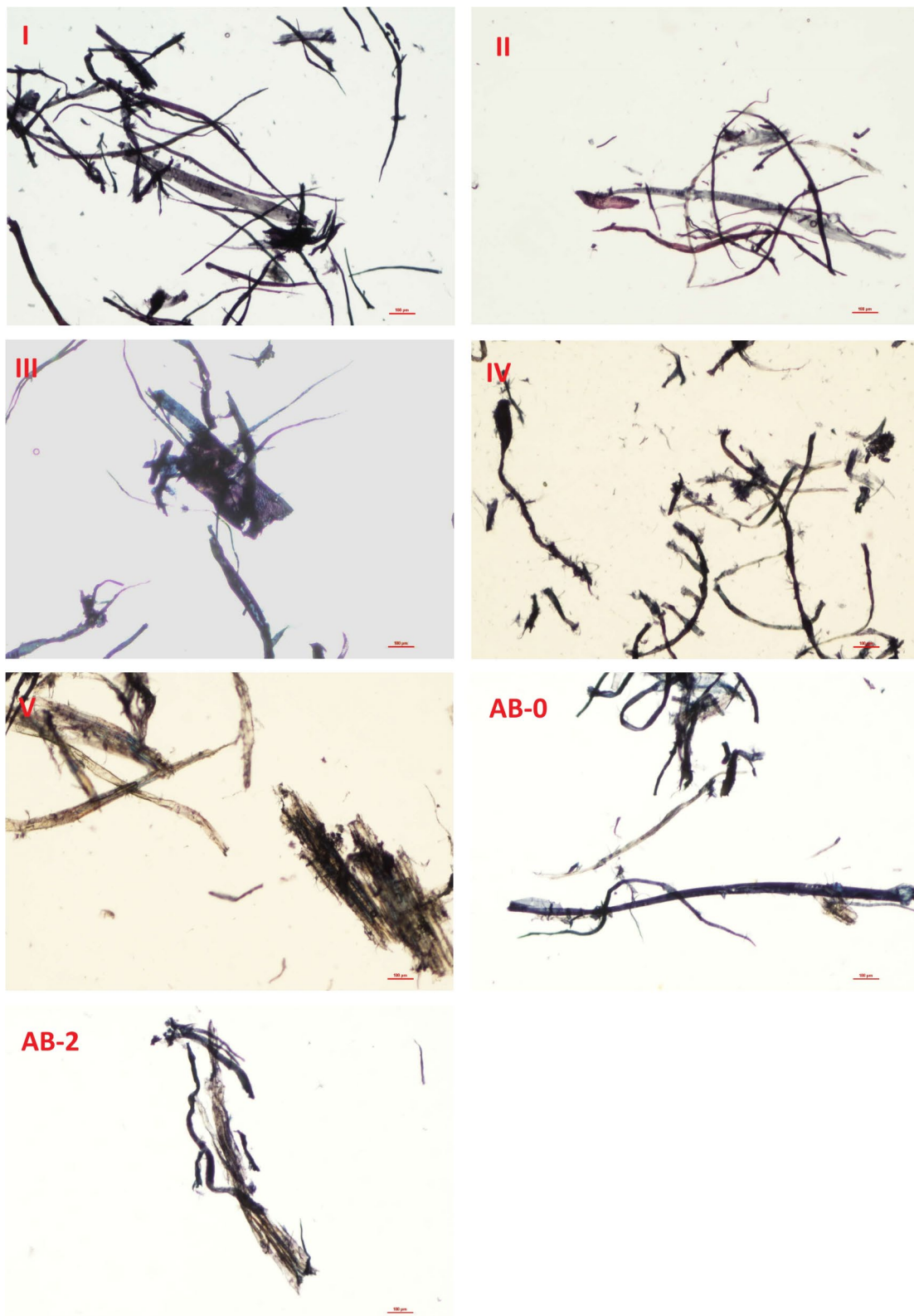


Fig. 1 Microscopic images of stained fibers from mock-up paper samples I, II, III, IV, V and archival paper samples AB-0 and AB-2

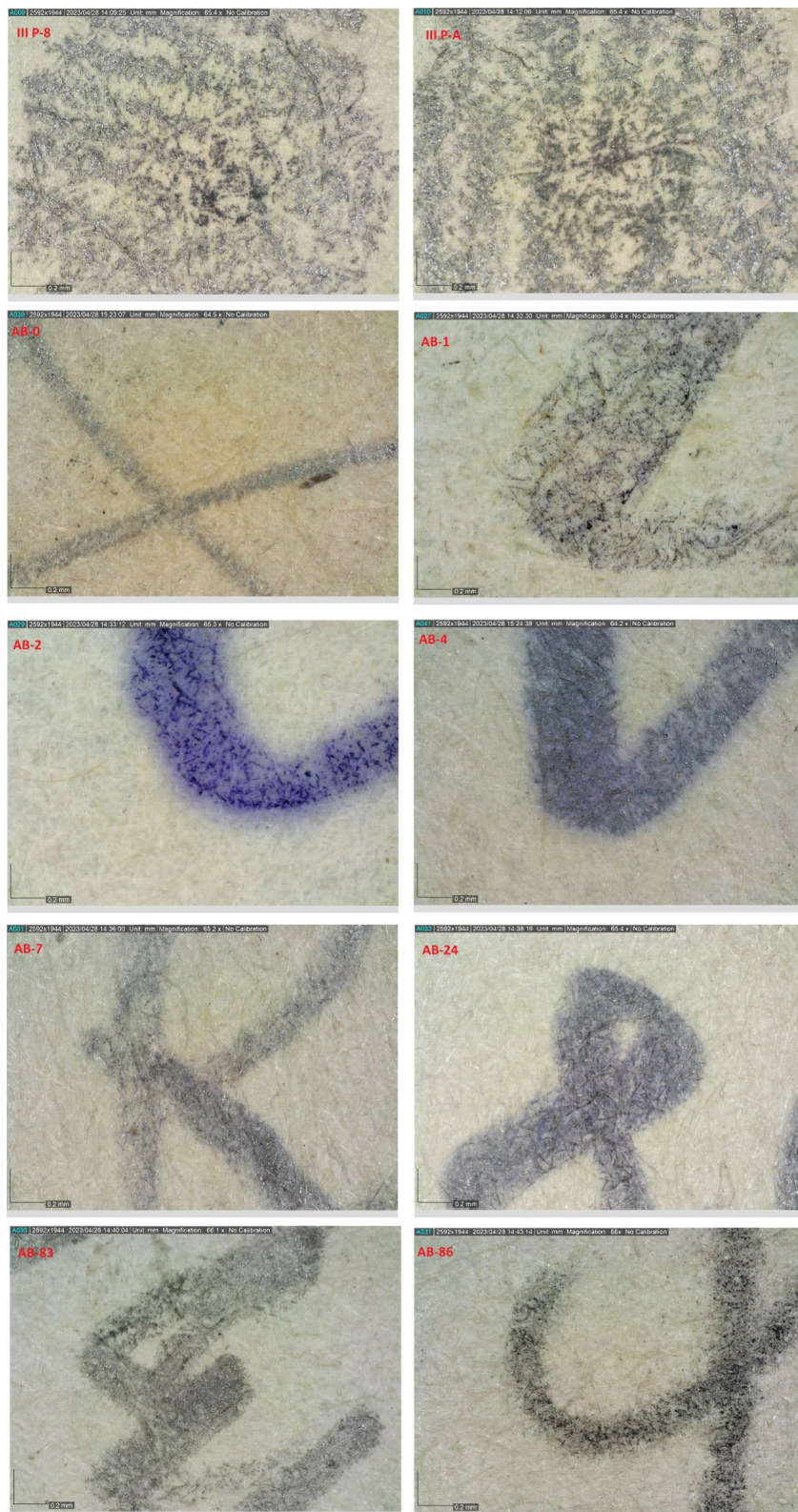


Fig. 2 Microscopic images of the pencils' marks on mock-up samples seen in the top row, with the eight archival documents below

internal reflection where part of the light is absorbed by the sample and the rest is reflected back to the detector. In FTIR spectroscopy the absorption of radiation by the molecules occurs due to presence of bond vibrations involving a change in dipole moment. ATR-FTIR has been successfully used to identify a range of primarily organic and sometimes inorganic materials in heritage objects; the method has proven effective for identification of papers [21], textile fibers [22], colorants [23–25], adhesives and binders [26], or plastics [27].

PCA

Principal component analysis (PCA) is one of the core multivariate statistics tools used in chemometrics, which allows for data compression and data features enhancement. The main principle of PCA is reduction of data dimensionality leading to isolation of the so-called principal components (PC), which represent directions of

greatest variation in the data. Data dimensionality reduction is especially useful when dealing with large datasets, such as in the case of spectroscopic measurements. The method of dimension reduction consists in linear transformation of data points onto a new coordinate system. The first PC is a vector that represents the greatest variability in the data, while each subsequent PC is a vector orthogonal to the previous one and stands for next greatest remaining variability. Due to these transformations PCA analysis allows for spectral data clustering on the basis of differences and similarities between the spectra.

As a result of PCA analysis, so called score plots and loading plots can be generated. Score plots visualize the similarities and differences between variables, resulting in clustering of samples that are similar and separation of samples that show differences with respect to the particular principal components. Loading plots provide a visualization of the degree of contribution of each dependent

Table 4 Absorption bands for paper samples I–V and AB-0–AB-86 within the region of 1300 cm^{-1} to 1750 cm^{-1}

Absorption cm^{-1}													Group
I	II	III	IV	V	AB-0	AB-1	AB-2	AB-4	AB-7	AB-24	AB-83	AB-86	
–	–	–	–	1737	1726	1732	1730	1732	1728	1729	–	1731	C=O stretch in hemicelluloses
1637	1638	1632	1639	1642	1643	1635	1638	1633	1643	1650	1649	1648	Adsorbed water
–	–	–	–	1600	–	–	–	–	–	1596	–	1601	C=O stretch in lignin
–	–	–	–	1508	1509	1511	1508	1507	1509	1507	1505	1509	Lignin aromatic skeletal vibrations
–	–	–	1449	1452	1450	1456	1455	1456	1455	1451	1451	1455	O–H bending in lignin/hemicelluloses/cellulose
1423	1428	1429	1429	1424	1427	1428	1427	1429	1425	1422	1427	1423	C–H bending in lignin and cellulose
1365	1360	1365	1369	1366	1370	1369	1369	1370	1369	1370	1367	1368	C–H bending in cellulose and hemicelluloses
1335	1332	1335	1336	1335	1331	1333	1335	1334	1337	1333	1334	1332	C–H ₂ wagging in cellulose
1316	1313	1315	1315	1314	1314	–	1313	1317	1318	1318	1318	1315	C–H bending in cellulose

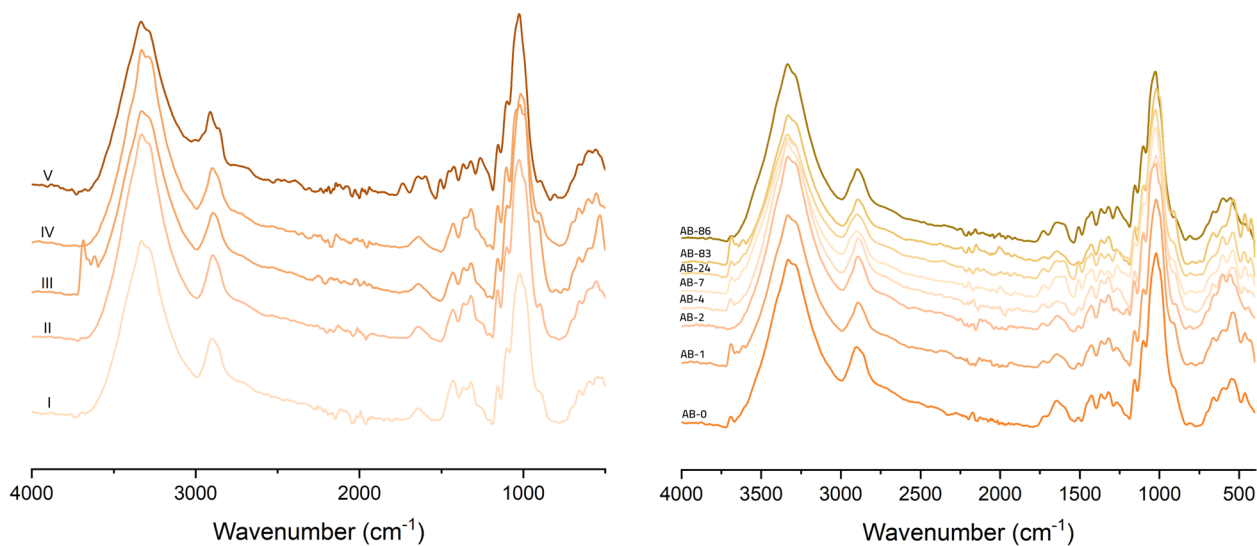


Fig. 3 ATR-FTIR spectra collected for the paper supports in mock-up samples (left) and archival samples (right)

variable, which in the case of spectroscopic data is the absorption intensity per wavelength with respect to the particular PC. Additionally, to allow for better understanding of which Principal Components contribute the most to the collected data variation, a so-called scree plot is generated. It represents the percentage by which each of the consecutive PCs explains the overall variation. PCA analysis is a tool often used in heritage science applications, employed in processing of data collected from spectroscopic analyses [28–30], chemical tests [31], or image analyses [32, 33].

Materials

Analytical grade MV and MB dyes (Warchem) were purchased to collect reference spectra. Four types of different contemporary, machine-made papers, including dyed papers and one type of cardboard were selected as substrates for the mock-up samples (Table 3). The fiber content of the papers was unknown, so the samples needed further profiling by means of microscopic observation. Vintage copying pencils, dating to the 1930s were bought at an online auction. They were tested for presence of aniline dye by applying water. Two out of nine copying pencils were blue, while the rest exhibited various shades of violet. One of the blue pencils turned out to be a colored aniline pencil, leaving a blue mark when writing on dry paper. Altogether, nine types of copying pencils (1–14) and nine types of contemporary graphite pencils (A–I) of various brands (Table 2) were applied to each of the 4 types of paper and onto a cardboard (I–V).

The pencils were applied in fine layer to ensure that the paper support was still visible from underneath (Fig. 1).

Methods

The pencil traces on documents from the Auschwitz—Birkenau State Museum Archives underwent observation under magnification with portable digital microscope Dinolite, equipped with white LED diodes. The photos of the areas of interest were taken at a $\times 65$ magnification and captured using the DinoCapture 2.0 software. Paper fibers sourced from mock-up samples and the selected archival documents were boiled in deionized water to remove adhesives, then fibers were manually separated with a preparation needle. After drying they were placed on microscopic slides and stained with Pol-Aura Herzberg reagent, a mixture of zinc chloride, iodine and potassium iodide solutions [34, 35]. Prepared microscopic slides underwent observation with the Nikon Eclipse 80i microscope, equipped with a digital camera Nikon DS-Fi3, a 10×0.30 objective, and 10×0.30 nose-piece, in transmitted light mode. Images were collected by means of NIS-Elements 5.01 software. During ATR-FTIR measurements paper samples were tested with Thermo Scientific FT-IR Nicolet iS50 spectrometer, equipped with an ATR accessory furnished with a diamond crystal. The measurement spot diameter was equal to 0.5 mm. Spectra were collected at a rate of 64 scans per minute. Background spectrum was recorded prior to the measurements. Collected spectra underwent processing in Omnic 9 software, including ATR correction,

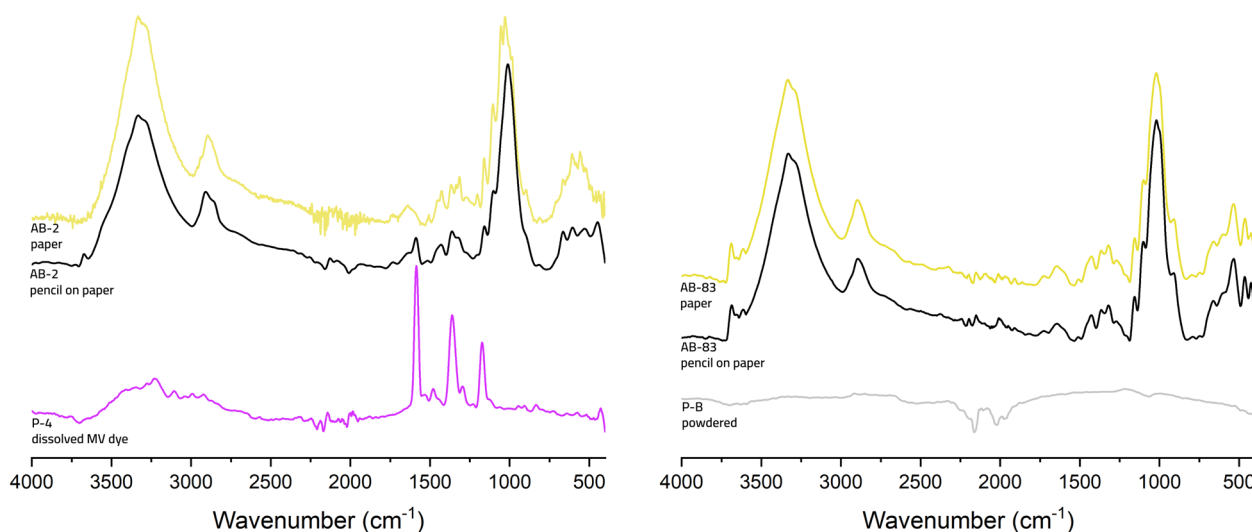


Fig. 4 Spectra of two pencil marks on archival objects; the aniline dye-containing pencil AB-2 and the graphite pencil AB-83 are compared with spectra of, respectively, Methyl Violet dye (P-4) and pure graphite (P-B), as well as the spectra of archival paper substrates

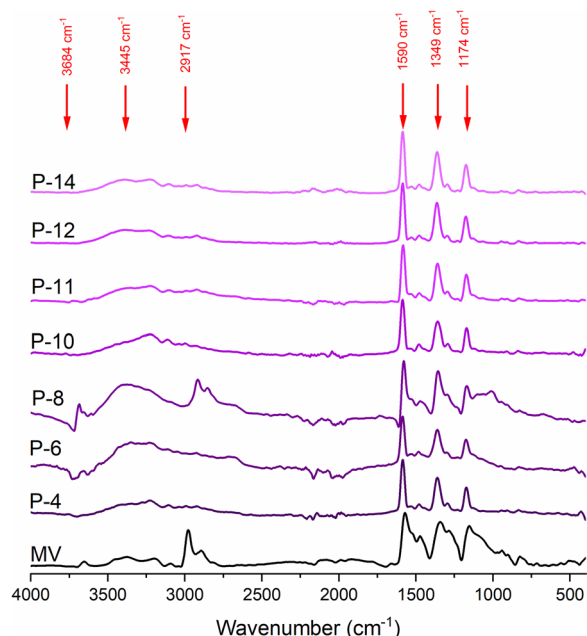


Fig. 5 Spectra of the aniline dyes dissolved from the copying pencils P-4, P-6, P-8, P-10, P-11, P-12 and P-14 show peaks characteristic for the Methyl Violet dye

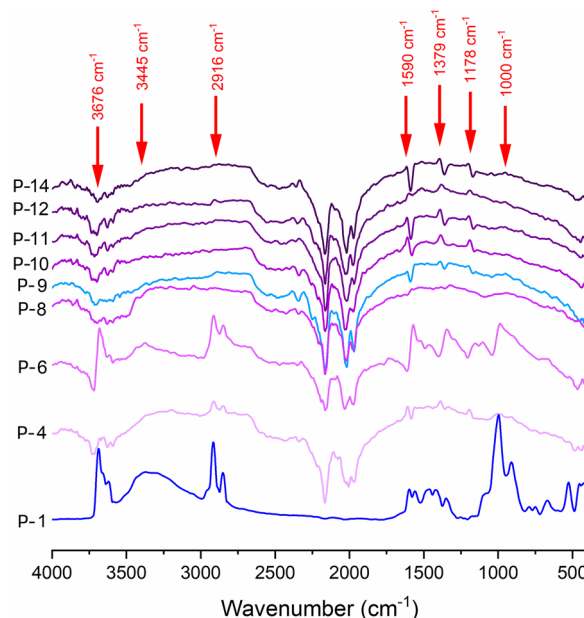


Fig. 7 Spectra of powdered copying pencils showing low intensity absorption bands characteristic for aniline dye admixture

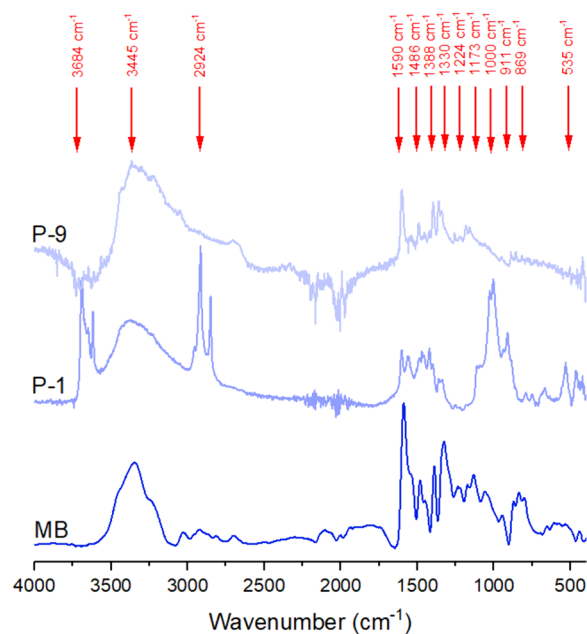


Fig. 6 Spectra of the aniline dyes dissolved from the copying pencils P-1 and P-9 show peaks characteristic for the Methylene Blue dye

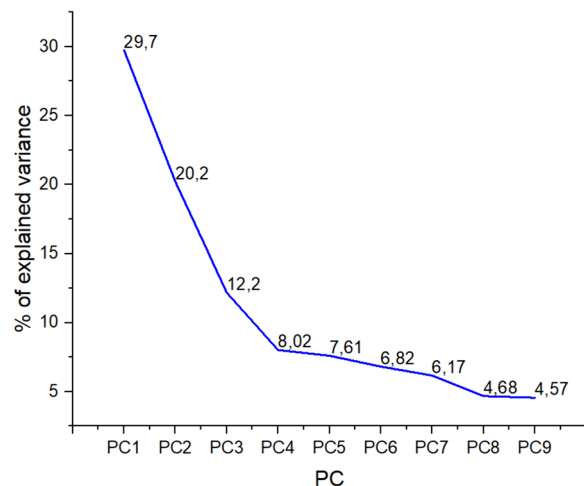


Fig. 8 Scree plot representing the contribution of the particular principal components (PC) to the explained variance of spectral data

background correction, and normalization to the peak at 3445 cm⁻¹. Omnic 9 software was used as well for preliminary comparison to spectral library entries. Selection of the Region of Interest (ROI) and the PCA analysis

of spectral data was performed with Peak Spectroscopy software, and the generated data was used to create score plots, loading plots and scree plots within OriginLab software.

Results

Visual analysis

Fibers in the paper samples were identified by means of optical microscopy and staining with Herzberg stain. The

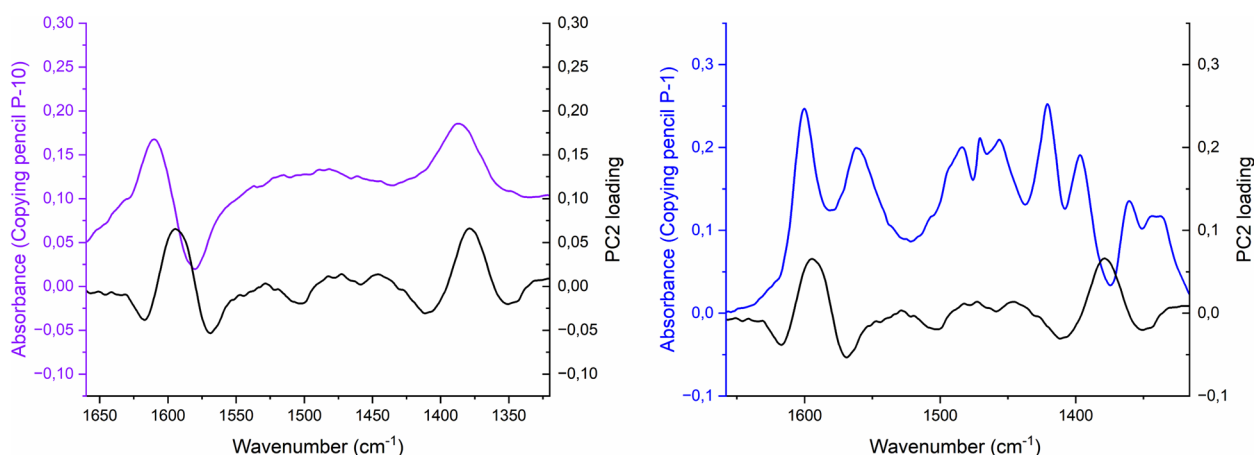


Fig. 9 Loading plots of the PC2 component (black), juxtaposed with the spectra of the Methyl Violet (purple)—containing pencil (P-10) and Methylene Blue (blue) – containing pencil (P-1)

paper no I consisted of softwood and hardwood chemical fibers dyed blue, the cream—colored paper (II) contained softwood and hardwood chemical fibers, yellow paper (III) showed presence of hardwood chemical fibers, while white paper (IV) exhibited content of softwood and hardwood chemical pulp fibers. Cardboard (V) showed presence of chemical softwood pulp and mechanical wood pulp. Similarly, samples from two sheets of archival documents (AB-0, AB-2) were taken in order to perform staining tests. The results indicate presence of both mechanical and chemical bleached wood pulp (Fig. 1).

Dinolite microscope at $\times 65$ magnification was used to inspect the pencil marks on the mock-up samples and eight selected documents (Fig. 2). Parts of the pencil marks where more writing material was deposited were selected for photographing with Dinolite and subsequent ATR-FTIR measurements. Line thickness of pencil marks varied from 0.2 mm to about 0.6 mm for different locations. Microscopic observation revealed that copying pencils are not differentiable from graphite pencils based on difference in color or shine. As pencil trace shine is related to its softness (the softer the lead, the greater the shine) it did not contribute to identification. Additionally, both Methylene Blue and Methyl Violet crystals show metallic luster in their solid form, which can lead to additional confusion when differentiating based on surface reflectivity.

ATR-FTIR

First, spectra for paper substrates in all 4 types of paper and the cardboard, as well as in the original documents were collected. All of the spectra recorded show features characteristic for cellulosic materials; the main bands comprising a broad peak around 3300 cm^{-1} representing

O–H stretching, a peak around $2900\text{--}3000\text{ cm}^{-1}$ representing C–H stretching vibrations, and a large peak around 1060 cm^{-1} that stands for C–O stretching vibrations in glycosidic linkages in the D–glucopyranose ring. The differences between particular paper types that reflect varying content of lignin and hemicelluloses, detectable in the $1300\text{ to }1750\text{ cm}^{-1}$ range are listed in the table (Table 4).

Paper spectra reveal as well an admixture of other ingredients such as fillers, whose content differs for particular paper types, with paper no III showing highest content of kaolinite filler, represented by the characteristic peaks at 3684 cm^{-1} and 535 cm^{-1} . The peaks representing admixture of kaolinite are identified as well in the tested samples of the archival paper substrates (Fig. 3).

For the purpose of ATR-FTIR measurements, MV and MB dyes in their solid forms, as well as powdered samples of graphite from the copying pencils no 1–14 were placed on microscopic slides and dissolved with ethanol. A drop of each dye was placed on the ATR crystal. After solvent evaporation the spectra were collected for each sample. Spectra for dissolved dyes from the copying pencils were compared to the spectra of MV and MB. As a result, dyes in 6 pencils were identified as MV and dyes in 2 pencils were classified as MB. The main bands allowing for differentiation of MV in copy pencil dyes were the 1590 cm^{-1} band representing the stretching of the C=C bond in the aromatic ring, as well as 1349 cm^{-1} band and the 1174 cm^{-1} band characteristic for stretching of the C–N bond in the aromatic amine. The bands in the single bond region, around 3000 cm^{-1} to 3700 cm^{-1} representing the N–H and O–H stretching vibrations and the band at 2917 cm^{-1} characteristic for C–H vibration in the methyl group were observed in

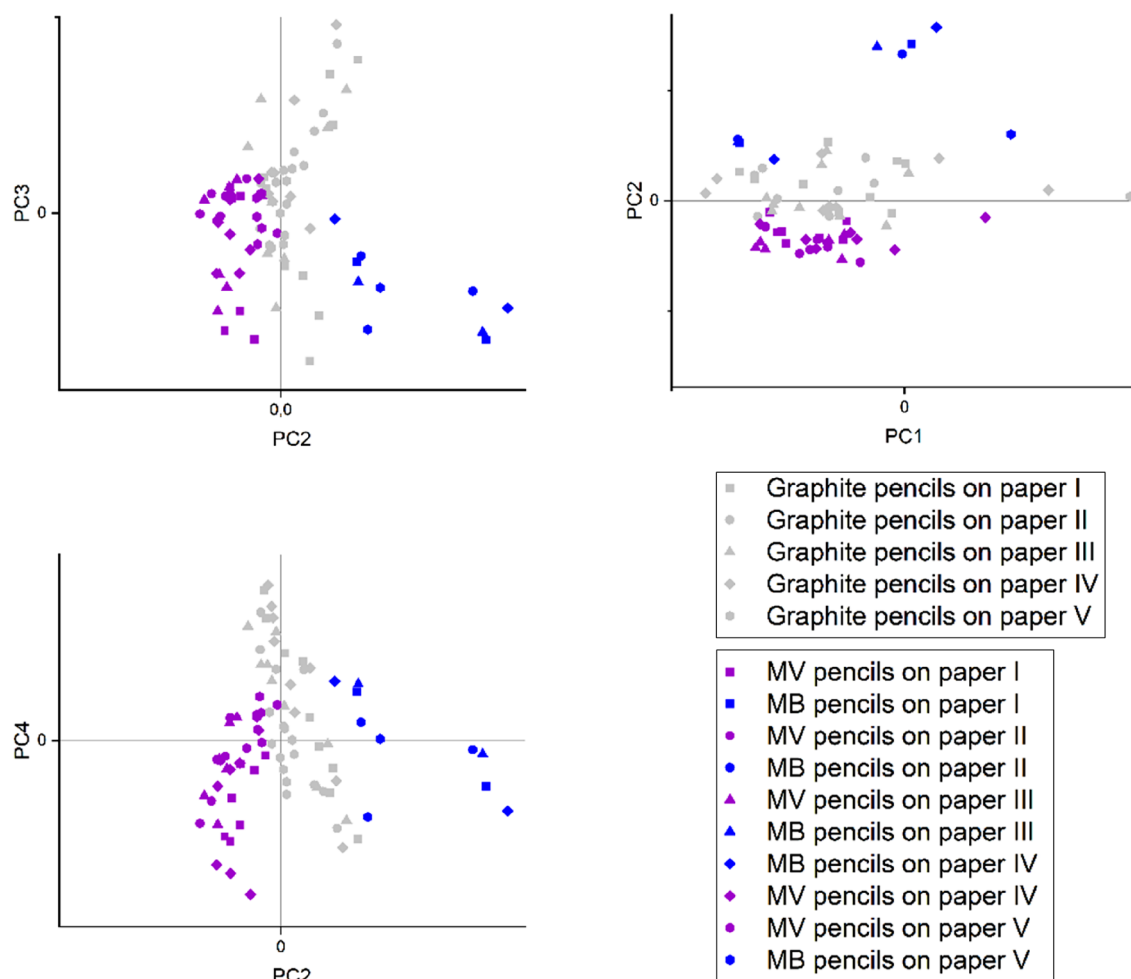


Fig. 10 Score plots representing PCA results of mock-up samples analysis

the MV spectrum, but not identifiable in all of the pencil dyes. For spectra identified as MB the main diagnostic bands were the C=N stretching vibration at 1590 cm^{-1} , C=C and C=S ring stretch bands between 1388 and 1486 cm^{-1} and the C-N stretch at 1330 cm^{-1} . Other bands characteristic for MB, such as C-H bending vibrations at 1224 , 1173 , 869 and 802 cm^{-1} , as well as C-N bending band at 1135 cm^{-1} and the C-S-C vibrations at 1060 cm^{-1} are less pronounced in the dye dissolved from the copying pencils lead, or can partly overlap with bands characteristic for the kaolinite vibrations. Kaolinite bands are clearly visible in the P-1 sample of the colored aniline pencil; the region between 3617 – 3684 cm^{-1} representing the O-H stretching vibrations, the region between 1000 and 1110 cm^{-1} characteristic for Si-O stretching vibrations and the bands at 911 cm^{-1} and 535 cm^{-1} denoting, respectively, the Al-Al-OH bending and the Al-O-Si bond bending.

Because pencil marks on paper and other materials in heritage objects contain dye in its solid, powdered form mixed with graphite, and are deposited on the surface by forces of adhesion, the next step involved ATR-FTIR measurements of dry powdered graphite sampled from the copying and regular pencils. As expected, spectra of regular graphite-based pencil lead do not show strong characteristic features; C-C bonds in graphite do not absorb IR radiation to a significant degree, which is visible on the spectrum of one of the graphite pencils (Fig. 4, right, P-B). In the case of copying pencils, the recorded spectra reveal that, depending on the proportions of the lead ingredients, characteristic bands around 1178 cm^{-1} , 1379 cm^{-1} and 1590 cm^{-1} represent the aniline dye content, while for pencils with higher concentration of kaolin clay, P-1, P-4 and P-6, their spectra show additional absorption bands at 3676 , 3445 , 2916 and 1379 cm^{-1} and in the case of P-1, with high kaolin proportion, also

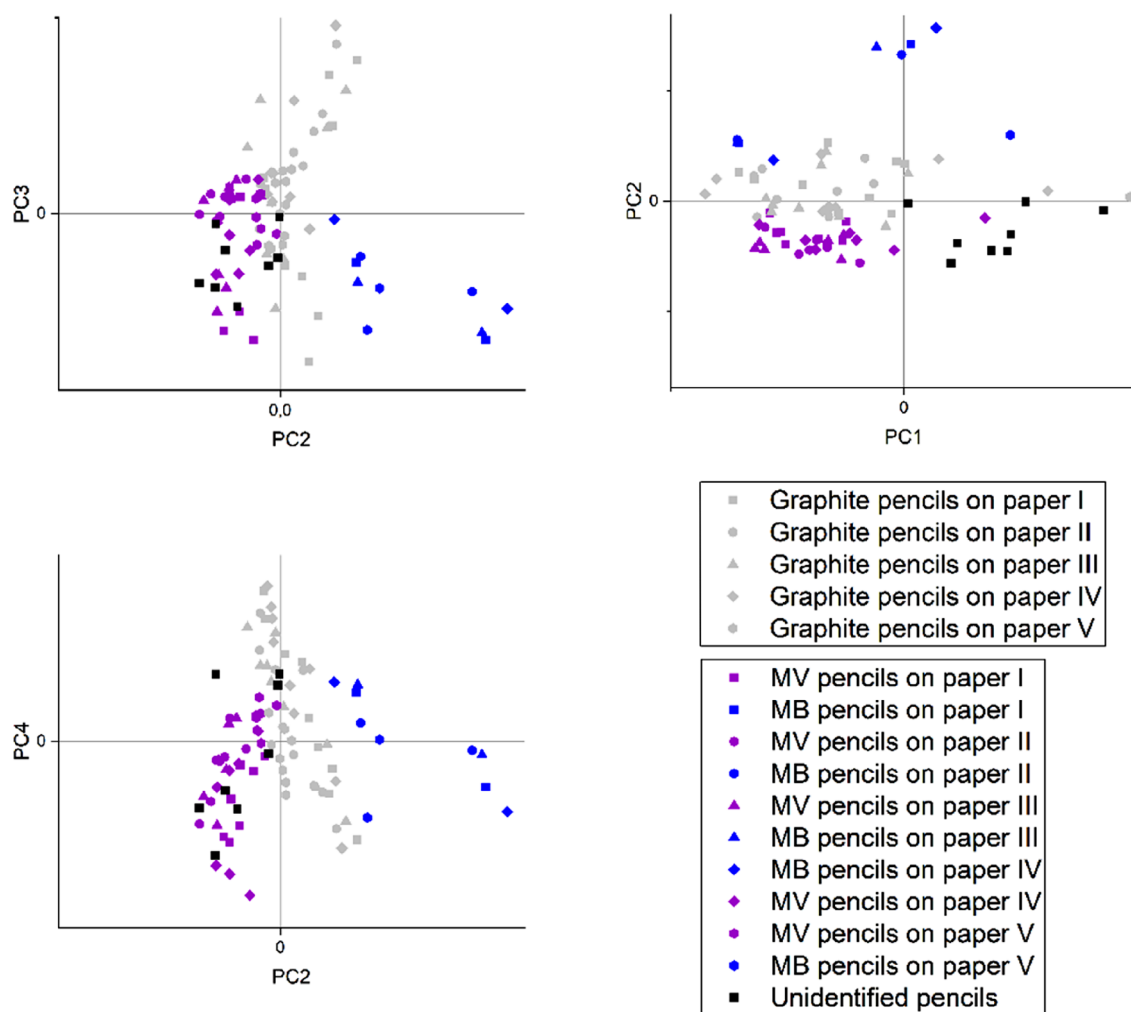


Fig. 11 Score plots representing PCA results of mock-up samples and archival samples analysis

within the range of $900\text{--}1100\text{ cm}^{-1}$, and around 530 cm^{-1} (Figs. 5, 6). Possible kaolin bands have been noted as well in the spectra of the pure graphite pencils; P-D and P-E, their intensities, however, were notably weak. In the next step, pencil marks in the mock-up samples as well as eight signatures and inscriptions executed with various kinds of pencils were selected for ATR-FTIR measurements from the documents in the SS Neubauleitung folder. Spectra were recorded for the pencil samples, in locations photographed previously with Dinolite microscope. Spectra were collected for the paper substrates as well Fig. 7.

PCA analysis

First, the mock-up samples were used as the training set for the PCA algorithm. Different regions of interest were tested to maximize the samples differentiation and

clustering. The region between 1300 and 1700 cm^{-1} was selected to generate the final PCA model. To verify which principal components were statistically significant, the percentage of data variation represented by each PC was calculated and a scree plot was generated. Statistically significant PCs were defined based on visual identification of the breaking point of the plot, a so called “elbow”. In the case of the following analysis these PCs encompass PC1 to PC4 (Fig. 8). PCA loading plots were examined for the first four PCs which allowed to identify the PC2 as the loading containing the most significant variation features (Fig. 9). The score plots for PC1 vs PC2, PC2 vs PC3 and PC2 vs PC4 were subsequently generated. In the final step, the PCA model was tested by adding the spectra recorded for the 8 samples of pencils from the archival documents, results of which were presented in the form of score plots (Figs. 10, 11).

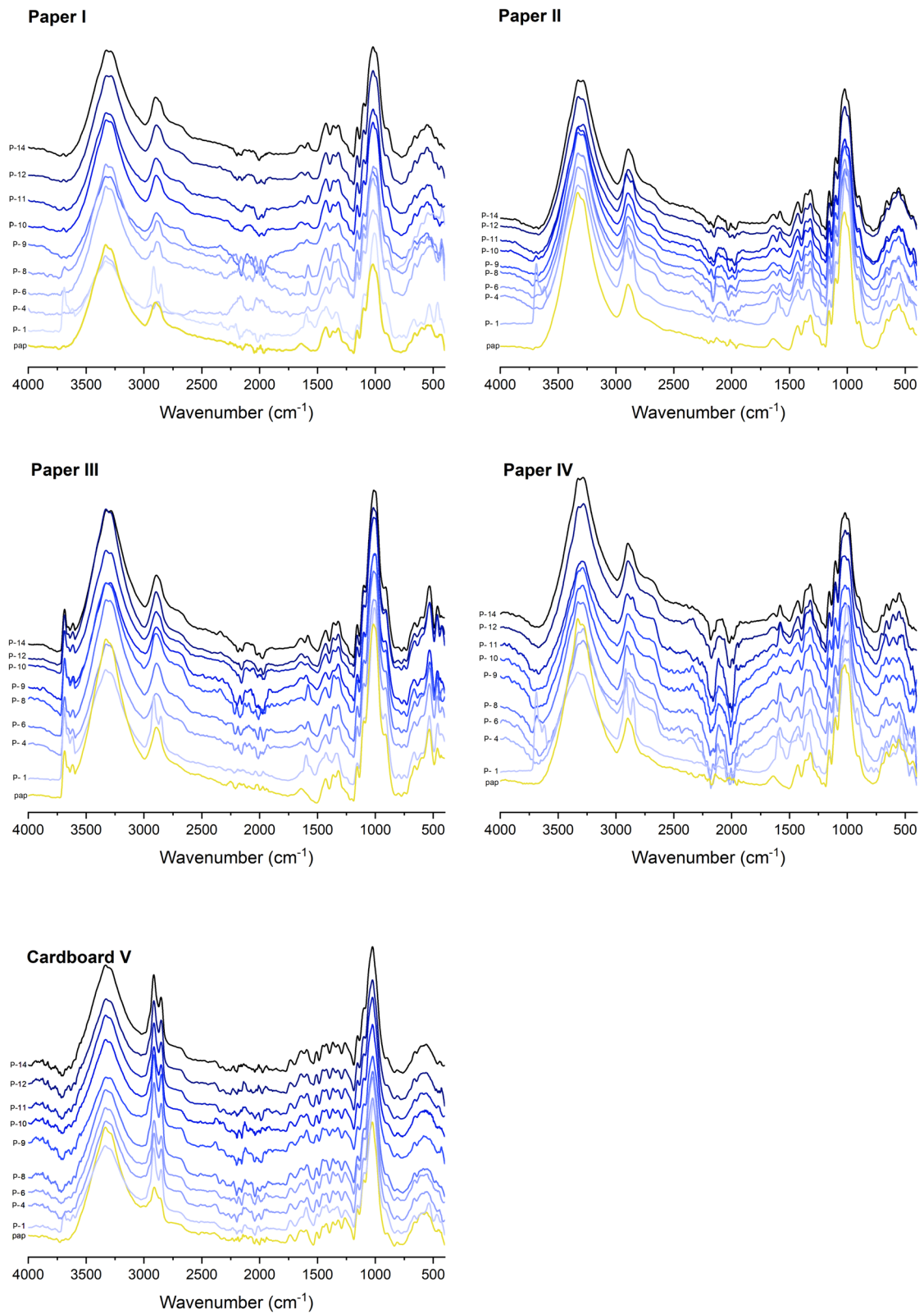


Fig. 12 Spectra of copying pencils on paper substrates from five mock-up samples (I–V)

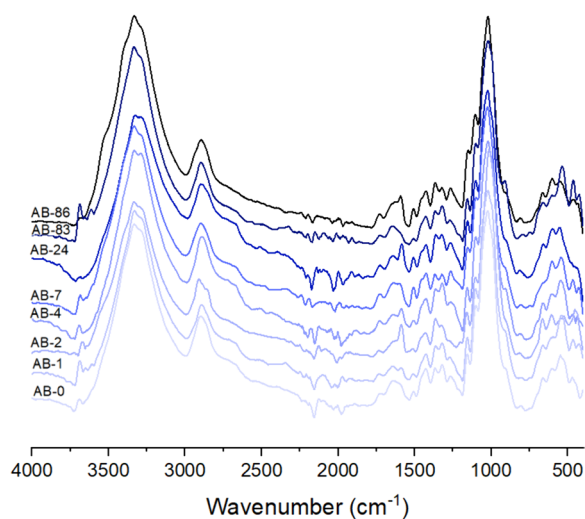


Fig. 13 Spectra of unidentified pencils on different paper substrates (AB-0 to AB-86) sampled from the documents from the archives of the Auschwitz—Birkenau State Museum

Discussion

ATR-FTIR

The ATR-FTIR measurements of complex samples yield spectra with multiple, sometimes overlapping features, which may prove difficult to identify using algorithms applied for spectra comparison within spectral library in the ATR-FTIR software. In the case of measurements performed for pencils on paper substrates, an attempt to identify presence of copying pencils by comparing recorded results to spectra collected in the spectral library did not prove successful. Upon juxtaposition of the collected spectra of all samples and the spectra of the substrates, bands predominantly attributable to the lignocellulosic content of the papers appear most prominent. For the analyzed set of samples, the first interesting wavelength range, containing features characteristic both for different paper types and the aniline dyes, ranges from around 1300–1700 cm^{-1} . The region is populated with peaks characteristic for cellulose, hemicelluloses, lignin and different functional groups in aniline dyes. While in the case of some samples a peak attributable to the C=N stretching vibration at 1590 cm^{-1} can be recognized (e.g. P-1, P-8, P-9 on paper I, Fig. 12), other bands characteristic for the aniline dye are not readily visible. Visibility of the aniline dye contribution seems to depend strongly on the type of paper support; additional peaks are most prominent for copying pencils on paper IV, containing chemical wood fibers, with characteristic high cellulose to lignin ratio, while the lowest aniline dye visibility can be identified for pencil marks on cardboard V, where lignin content is significant. Similarly, for the group of archival materials the prominence of features

attributable to aniline dye content in this range is highly variable. For certain samples the band at 1590 cm^{-1} but also the band at 1349 cm^{-1} are detectable (AB-2, Figs. 4, 8), while for others the bands are not visible (AB-83, Figs. 13, 4). While outside the range of 1300–1700 cm^{-1} some regions with features characteristic for aniline dye presence can also be identified, their significant overlap with peaks from other compounds such as kaolinite, or cellulose does not render them good diagnostic regions. Among the bands detected for FTIR measurements of powdered copy pencil leads, the bands within the 1300–1700 region appear consistently, even for pencil samples with relatively low aniline dye concentration, however for pencils on paper the difficulty arises upon manual comparison due to peak overlap in this range. Nevertheless, this region appears the most promising in an attempt to differentiate copying pencils and graphite pencils on paper with PCA analysis.

PCA

Based on the distribution of the features in the collected FTIR spectra, a PCA analysis protocol has been developed, which included identification of the region of interest as a wavelength range between 1300 and 1700 cm^{-1} . The generated PCA score plots (Fig. 10, 11) allowed for successful differentiation of graphite and copying pencils as well as the colored aniline pencil resulting in three clusters; a cluster encompassing copying pencils containing MV, a group of copying pencils containing MB, and a group of pure graphite pencil samples. The PC2 is negatively correlated for samples containing MV dye and positively correlated for samples containing MB dye. Loading plots analysis suggests major contribution of two peaks; 1585 cm^{-1} and 1378 cm^{-1} to the PC2 loading, whose presence is the basis of the observed clustering in its juxtapositions with PC1, PC3 and PC4. Within the loading spectrum, the first peak at 1585 cm^{-1} represents a feature common for both MV and MB containing pencils, while the second peak at 1378 cm^{-1} is characteristic for MV containing samples. Also, minor contributions from bands characteristic for MB in the range of 1400–1560 cm^{-1} are the basis for sample clustering (Fig. 9). Upon testing of the model with samples collected from the archival documents, the results indicated that majority of the pencil marks was executed by means of copying pencils containing MV dye, including samples AB-1, AB-2, AB-4, AB-7, AB-24 and AB-86. Two remaining samples, AB-0 and AB-83, can be classified as pure graphite pencils.

Based on the results, it is possible to assess the combined application of FTIR spectroscopy and PCA analysis, as a suitable method for effective diagnosis for archival collections containing many objects with various types of graphite-based writing materials. Results of such an

analysis allow for determination of groups of objects containing copy pencils, and thus requiring application of targeted conservation treatments and exhibition illumination policies tailored to objects within the highest light–sensitivity class. The method, however, bears certain limitations which have to be considered. It is plausible that in the case of copying pencils with low concentrations of aniline dye admixture, a sample undergoing PCA analysis may be classified as graphite–based pencil. Interestingly, the results of MV dye dissolution from the powdered pencil leads indicated different shades of purple in each pencil, from light pink to purple–black. The color differences, however, do not seem to correspond to the degree of clustering. It seems, however, that chemical composition of the substrate itself can have greater significance. For example, the copying pencils no P-9 and P-10, applied on a substrate with high content of lignin (cardboard V), show an overlap of the 1590 cm^{-1} peak, diagnostic for aniline dye presence and the 1600 cm^{-1} peak, characteristic for C=O stretch in lignin. As a result, the samples show weaker correlation with the PC2 component. Another important limitation concerns identification of copying pencil marks, that underwent degradation processes, such as photo–oxidation. In the process of degradation of organic dyes, the diagnostic features of their chemical structure, such as functional groups or the conjugated bond system, are affected, which may lead to changes in the FTIR spectrum and resulting difficulty in correct identification. A development of diagnostic method independent of dye degradation processes could constitute the next step within the research on copying pencils identification.

Acknowledgements

We acknowledge financial support from the Open Access Publication Fund of Universität Hamburg.

Author contributions

The author confirms sole responsibility for the following: study conception and design, data collection, analysis and interpretation of results, and manuscript preparation.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Availability of data and materials

The data that support the findings of this study are available from State Museum Auschwitz–Birkenau, but restrictions apply to the availability of these data, which were used under license for the current study, and so are not publicly available. Data are however available from the authors upon reasonable request and with permission of State Museum Auschwitz–Birkenau.

Declarations

Competing interests

The authors declare that they have no competing interests.

Received: 5 March 2024 Accepted: 10 July 2024

Published online: 30 July 2024

References

1. Stroud J., *Ink on Manuscripts and Documents*, Conservation of Archival Materials: Fourth Annual Seminar. The Harry Ransom Humanities Research Center. 1985; 58.
2. Travis AS. Perkin's mauve: ancestor of the organic chemical industry. *Technol Cult.* 1990;31(1):51–82.
3. Barnett JC. Synthetic organic dyes, 1856–1901: an introductory literature review of their use and related issues in textile conservation. *Stud Conserv.* 2007;52(sup1):67–77.
4. Hagan E, Poulin JA. Statistics of the early synthetic dye industry. *Heritage Science.* 2021;9:1.
5. Püntener AG. Colorants for non-textile applications. *Leather Dyes.* 2000. <https://doi.org/10.1016/B978-044482888-0/50042-8>.
6. Tucker EC, et al. *The manufacture of pulp and paper: a textbook of modern pulp and paper mill practice*, vol. 4. New York: McGraw-Hill Book Company; 1924.
7. Kelly S, Walker A, Ellsworth AA, Hedlund JK. Characterization of the aniline dyes in the colored papers of José Posada's prints using time-of-flight secondary ion mass spectrometry to aid in developing a treatment protocol for the removal of pressure-sensitive tapes. *Book Paper Group Annual.* 2017;36:87–100.
8. Cesaretto A, Luo Y, Smith HD, Leona M. A timeline for the introduction of synthetic dyestuffs in Japan during the late Edo and Meiji periods. *Herit Sci.* 2018. <https://doi.org/10.1186/s40494-018-0187-0>.
9. Townsend JH, Carlyle L, Khandekar N, Woodcock S. Later nineteenth century pigments: evidence for additions and substitutions. *Conservator.* 1995;19:65–78.
10. Dobrusskin S. *Frühe, nichtphotographische Kopier- und Vervielfältigungstechniken*, Preprints of the 9th International IADA Congress. Copenhagen. 1999;15–21:195–206.
11. Dube L. The Copying Pencil: Composition History, and Conservation Implications. *Book Paper Group Annual.* 1998;17:45–52.
12. Mitchell CA. Copying-ink pencils and the examination of their pigments in writing. *Analyst.* 1916;42:3–11.
13. Ellis MH, Yeh MB. The history, use and characteristics of wax-based drawing media. *Paper Conserv.* 1998;22(1):48–55.
14. Cooksey CJ. Quirks of dye nomenclature 7. gentian violet and other violets. *Biotechnic Histochem.* 2017;92(2):134–40.
15. Howland RH. Methylene blue: the long and winding road from stain to brain: part 1. *J Psychosoc Nurs Ment Health Serv.* 2016;54(9):21–4.
16. Hagan E, Castro-Soto I, Breault M, Poulin J. The lightfastness of early synthetic organic dyes. *Herit Sci.* 2022;10(50):1–11.
17. Cesaratto A, Lombardi JR, Leona M. Tracking photo-degradation of triarylmethane dyes with surface – enhanced Raman spectroscopy. *J Raman Spectrosc.* 2016;48(3):418–24.
18. Khan I, Saeed K, Zekker I, Zhang B, Hendi A, Ahmad A, Ahmad S, Zada N, Ahmad H, Shah L, Shah T, Khan I. Review on methylene blue: its properties, uses toxicity and photodegradation. *Water.* 2022;14(24):1–30.
19. Długoborski W., Piper F. *Auschwitz 1940 – 1945. Węzłowe zagadnienia z dziejów obozu. Założenie i organizacja obozu.* Wydawnictwo Państwowego Muzeum Oświęcim – Brzezinka; 1995.
20. Hofmann C, Mecklenburg M. Modern transparent papers: materials, degradation, and the effects of some conservation treatments. *J Am Inst Conserv.* 1992;32(2):177–206.
21. Wertz JH, McClelland AA, Mayer DD, Knipe P. Modeling chemical tests and fiber identification of paper materials using principal component analysis and specular reflection FTIR. *Heritage.* 2022;5(3):1960–73.
22. Garside P, Wyeth P. Identification of cellulosic fibers by FTIR spectroscopy. *Stud Conserv.* 2003;48:269–75.
23. Bruni S, De Luca E, Guglielmi V, Pozzi F. Identification of natural dyes on laboratory-dyed wool and ancient wool, silk, and cotton fibers using attenuated total reflection (ATR) Fourier transform infrared (FT-IR) spectroscopy and fourier transform Raman spectroscopy. *Appl Spectrosc.* 2011;65(9):1017–23.
24. Tomasini E, Siracusano G, Maier MS. Spectroscopic, morphological and chemical characterization of historic pigments based on carbon. Paths for the identification of an artistic pigment. *Microchem J.* 2012;102:28–37.
25. Grazenaite E, Kiuberis J, Beganskiene A, Senvaitiene J, Kareiva A. XRD and FTIR characterisation of historical green pigments and their lead-based glazes. *Chemija.* 2014;25(4):199–205.
26. Nodari L, Ricciardi P. Non-invasive identification of paint binders in illuminated manuscripts by ER-FTIR spectroscopy: a systematic study of

- the influence of different pigments on the binders' characteristic spectral features. *Herit Sci.* 2019;7:7.
27. Bell J, Nel P, Stuart B. Non-invasive identification of polymers in cultural heritage collections: evaluation, optimisation and application of portable FTIR (ATR and external reflectance) spectroscopy to three-dimensional polymer-based objects. *Heritage Science.* 2019;7:95.
 28. Mitchell G, France F, Nordon A, et al. Assessment of historical polymers using attenuated total reflectance-Fourier transform infra-red spectroscopy with principal component analysis. *Herit Sci.* 2013;1:28.
 29. Noake E, Deborah L, Nel P. Identification of cellulose nitrate based adhesive repairs in archaeological pottery of the University of Melbourne's Middle Eastern archaeological pottery collection using portable FTIR-ATR spectroscopy and PCA. *Herit Sci.* 2017;5(3):1–15.
 30. Albrecht M, de Noord O, Meloni S, et al. Jan Steen's ground layers analysed with principal component analysis. *Herit Sci.* 2019;7:53.
 31. Cardinali F, Bracciale MP, Santarelli ML, Marrocchi A. Principal component analysis (PCA) combined with naturally occurring crystallization inhibitors: an integrated strategy for a more sustainable control of salt decay in built heritage. *Heritage.* 2021;4:220–9.
 32. Salerno E, Tonazzini A, Grifoni E, Lorenzetti G, Legnaioli S, Lezzerini M, Marras L, Pagnotta S, Palleschi V. Analysis of multispectral images in cultural heritage and archaeology. *J Applied Laser Spectrosc.* 2014;1:1.
 33. Sidiropoulos AA, Konstantinos NL, Vasiliki KM. PCA-based localization of pathology on heritage building masonry. *Int J Const Res Civil Eng.* 2016;2:3.
 34. Shi JL, Li T. Technical investigation of 15th and 19th century Chinese paper currencies: fiber use and pigment identification. *J Raman Spectrosc.* 2013;44:6.
 35. Standard Test Method for Fiber Analysis of Paper and Paperboard, ASTM D 1999;1030 – 95

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.