

# Sorting the colors

dimensionsbezogene  
generalisierung



vergleichsbasierter  
sortierung



leo  
blume  
15.J.  
Gymnasium  
Essen-  
Werden

$$d(\vec{a}, \vec{b}) = \|\vec{a} - \vec{b}\|_2$$

$$d_E(l) = \sqrt{\sum_{i=1}^n (l_i - l_{i+1})^2}$$

$$d_E((e)) = \sqrt{\sum_{i=1}^n (e_i - e_{i+1})^2}$$

$$\text{Viele Linien: } l_{i+1} > l_i$$

$$d_E((e)) = \sqrt{\sum_{i=1}^n (e_i - e_{i+1})^2}$$

$$\Delta z \approx 0.05$$

$$(l_1, l_2) \in E$$

$$\begin{cases} 0 & (a, b) \in E \\ 1 & (a, b) \notin E \end{cases}$$

$$\Delta p \approx 0.05$$

# Inhaltsverzeichnis

Inhaltsverzeichnis .....	I
1. Einleitung, Fragestellung und Vorgehensweise .....	II
2. Beweisführung .....	1
2.1. Definitionen .....	1
2.2. Beweis: Jede sortierte Liste ist kettensortiert .....	1
2.3. Generalisierung auf n-dimensionale Vektorräume .....	4
3. Graphentheoretische Grundlagen .....	4
3.1. Definitionen .....	4
3.2. Anwendungen auf die Fragestellung .....	4
4. Algorithmen .....	5
4.1. Komplexität .....	5
4.1.1. Handhabbarkeit .....	6
4.1.2. Beweis der Unhandhabbarkeit durch Reduktion .....	6
4.2. Heuristik .....	7
4.3. Pfadkonstruktion .....	7
4.3.1. Triviale Pfadkonstruktion .....	8
4.3.2. Brute Force .....	8
4.3.3. Nächster Nachbar .....	8
4.3.4. Greedy .....	9
4.4. Pfadverbesserung .....	9
4.4.1. Rotation .....	9
4.4.2. Swap .....	10
4.4.3. 2-opt .....	10
4.4.4. 3-opt und k-opt .....	11
4.4.5. Simulated Annealing .....	11
5. Webanwendung .....	11
5.1. Grundlagen und Open-Source-Implementierung .....	11
5.1.1. Frontend .....	12
5.1.2. Backend .....	12
5.2. Server-Client-Kommunikation .....	12
5.3. Zahlen sortieren .....	13
5.4. Vektoren sortieren .....	13
5.5. Orte sortieren .....	14
5.6. Farben sortieren .....	14
5.6.1. Farbräume .....	14
5.6.2. Farbauswahl .....	15
5.6.3. Visualisierung .....	15
6. Fazit und Ausblick .....	i
Literaturverzeichnis .....	ii

# 1. Einleitung, Fragestellung und Vorgehensweise

„Die Mathematik, richtig verstanden, besitzt nicht allein Wahrheit, sondern auch höchste Schönheit.“

— Bertrand Russell (1872-1970), Philosoph und Mathematiker

Wie kann man Bücher nach Farben sortieren? Diese Frage stellte ich mir am 27. August 2023, einem regnerischen Spätsommertag, während ich mein Zimmer aufräumte und feststellte, dass meine Mathematikbücher kaum nach einer anderen Kategorie einzuteilen waren.

Die eindimensionale Sortierung, die auf einem Vergleich von Elementen basiert, funktioniert nicht – die Frage, ob Olivgrün größer als Karmesinrot ist, ergibt keinen Sinn. Zunächst versuchte ich, die Farben numerisch einzuteilen: nach ihrer Helligkeit oder ihrem Farbwert. Beide Methoden erzielten nicht das gewünschte Ergebnis: so könnte im ersten Fall ein Rot direkt zwischen zwei subtil unterschiedlichen Grüntönen stehen, während im zweiten Fall ein Pastellblau zwischen Laubgrün und Bordeauxviolett (zwei sehr dunkle Farben) eingeordnet wurde.

Mit dem Ziel, das Regal dennoch farblich ästhetisch zu sortieren, was (für mich) bedeutet, Farbkontraste zwischen nebeneinanderstehenden Büchern zu minimieren, war die Idee für mein diesjähriges Jugend forscht-Projekt geboren – die vergleichsbasierte Sortierung auf mehrdimensionale Objekte wie Farben zu erweitern und zu generalisieren. In diesem Projekt gehe ich den Fragestellungen nach, inwiefern mehrdimensionale Sortierung möglich und handhabbar ist, welche Rolle dabei die Graphentheorie spielt und welche Anwendungen neben dem (unter dem Gesichtspunkt der Kontrastminimierung) ästhetischsten Bücherregal sie hat.

Dabei gehe ich zu Beginn meiner Arbeit zunächst auf die theoretischen Grundlagen der mehrdimensionalen Sortierung ein (2.1), beweise, dass es sich bei der angestrebten um eine Generalisierung der bekannten eindimensionalen handelt (2.2) und erläutere die äquivalente graphentheoretische Darstellung des Problems (3). Nach der Präsentation der algorithmischen Komplexität (4.1) und der implementierten Algorithmen stelle ich im Anschluss meine Entwicklung einer interaktiven Webanwendung vor (5), in der man selbst verschiedene n-dimensionale Objekte sortieren kann – von abstrakten Vektoren (5.4) über geografische Punkte (5.5) bis hin zu visuellen Farben (5.6). Diese Webanwendung visualisiert zudem die Funktionsweise der implementierten Algorithmen anschaulich und ermöglicht die praktische Sortierung auch größerer Datenmengen.

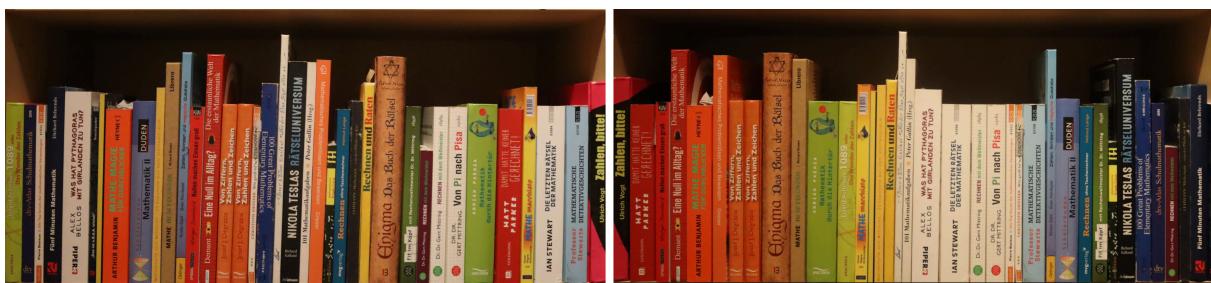


Abbildung 1: Meine Mathematikbücher: links lexikographisch nach Nachname des Autors, rechts nach perzeptueller Farbe mittels Simulated Annealing (Abschnitt 4.4.5) sortiert.

## 2. Beweisführung

In diesem Abschnitt soll bewiesen werden, dass es sich bei der im Folgenden definierten Ketten sortierung um eine Generalisierung der vergleichsbasierten Sortierung auf mehrere Dimensionen handelt. Dabei werden zunächst auf der Zermelo-Fraenkel-Mengenlehre[86, 37] beruhende Definitionen formuliert und im zweiten Teil mittels vollständiger Induktion[70] ein Beweis aufgestellt. Dieser Beweis liefert die Grundlage für die anschließende Ausarbeitung, da es sich ohne ihn nicht um eine n-dimensionale Sortierung, sondern nur um einen beliebigen anderen Algorithmus auf Listen von Vektoren handelte.

### 2.1. Definitionen

Eine **Liste** der Länge  $n \in \mathbb{N}$  ist zum Zwecke dieser Arbeit eine injektive Abbildung  $l$ , die als Eingabe eine natürliche Zahl  $i$  (den **Index**) im Definitionsbereich  $[1, n] \cap \mathbb{N}$ , folglich als  $\mathbb{D}$  bezeichnet, erhält und ein Element des Eingabealphabets  $\Sigma$  zurückgibt. Der Wert von  $l$  an der Stelle  $i$  wird als  $l_i$  notiert. Der **Teilabschnitt** von  $p$  bis  $q$ ,  $p, q \in \mathbb{D}, p \leq q$  beschreibt die  $q - p + 1$ -lange Liste  $[l_p, l_{p+1}, \dots, l_{q-1}, l_q]$  und wird als  $l_{p:q}$  notiert. Die Bildmenge  $Y$  meint die ungeordnete Menge aller in  $l$  vorkommenden Elemente, durch Injektivität gilt  $|Y| = |l|$ .

Das **Vertauschen** Swap meint:  $\text{Swap}(l, m, n)_i := \begin{cases} l_n & \text{falls } i=m \\ l_m & \text{falls } i=n \\ l_i & \text{sonst} \end{cases}$

Die Menge der **Permutationen** Perm einer Liste  $l$  wird definiert als:

$$\text{Perm}(l) := \{l' \mid l'_i = l_{f(i)}, f : \mathbb{D} \rightarrow \mathbb{D} \text{ bijektiv}\} \quad (1)$$

Eine Liste  $l$ , deren Zielmenge Teil einer strikten Totalordnung  $(\mathbb{T}, >)$  mit Ordnungsrelation  $(>)$  ist, heißt genau dann **sortiert**, wenn gilt:

$$\forall i \in \mathbb{D} \setminus \{n\} : l_{i+1} > l_i \quad (2)$$

Die **Kettenlänge**  $d_{\mathbb{L}}(l)$  einer Liste  $l$ , deren Zielmenge Teil eines metrischen Raums  $(\mathbb{T}, d)$  mit Distanzfunktion  $d$  ist, wird definiert durch:

$$d_{\mathbb{L}}(l) := \sum_{i=1}^{n-1} d(l_i, l_{i+1}) \quad (3)$$

Eine solche Liste wird als **kettensortiert** bezeichnet, wenn gilt:

$$d_{\mathbb{L}}(l) = \min_{l' \in \text{Perm}(l)} d_{\mathbb{L}}(l') \quad (4)$$

### 2.2. Beweis: Jede sortierte Liste ist kettensortiert

Neben der abbildenden Definition der Liste kann eine Liste reeller Zahlen (also  $\Sigma \subseteq \mathbb{R}$ ) äquivalent auch induktiv definiert werden. Hierbei repräsentiere  $\mathbb{L}$  die Menge aller solcher Listen,  $\varepsilon$  die leere Liste mit Länge 0.  $\oplus$  bezeichne die strukturelle Konkatenation zweier Listen bzw. einer Liste und einem Element, welches als unitäre Liste interpretiert wird. Die Definition erfolgt wie folgt:

$$\varepsilon \in \mathbb{L} \quad (5)$$

$$\begin{aligned} l \in \mathbb{L} \wedge n = |l| \wedge e \in \Sigma \wedge \forall i \in [1, n] : e > l_i \\ \Leftrightarrow \forall l^e \in \text{Ins}(l, e) : l^e \in \mathbb{L} \end{aligned} \quad (6)$$

Dabei wird die Einfügemenge  $\text{Ins}$  definiert als:

$$\text{Ins}(l, e) := \{l^e \mid n = |l|, i \in [1, n+1], l^e = l_{1:i-1} \oplus e \oplus l_{i:n}\} \quad (7)$$

Die Menge der sortierten Listen  $\mathbb{L}_{\text{sort}}$  bezeichnet dabei:

$$\varepsilon \in \mathbb{L}_{\text{sort}} \quad (8)$$

$$l \in \mathbb{L}_{\text{sort}} \wedge n = |l| \wedge e \in \Sigma \wedge \forall i \in [1, n] : e > l_i \stackrel{[1]}{\Leftrightarrow} l \oplus e \in \mathbb{L}_{\text{sort}} \quad (9)$$

Dabei gilt:  $\mathbb{L}_{\text{sort}} \subset \mathbb{L}$ , da es sich bei der Konstruktion von  $\mathbb{L}_{\text{sort}}$  um einen Spezialfall der von  $\mathbb{L}$  handelt, bei der  $\text{Ins}(l, e) := \{l \oplus e\}$ .

**Nebensatz.** Jede Liste mit total geordneter Zielmenge nach Abschnitt 2.1 ist Element von  $\mathbb{L}$ .

**Beweis.** Man betrachte eine abbildende Liste  $l$ . Aus dieser konstruiere man nun die  $\mathbb{L}$ -Liste,  $w$  genannt, iterativ. So beginne man mit der leeren Liste  $\varepsilon$  (nach (5)  $\in \mathbb{L}$ ) und betrachte stets das kleinste nicht betrachtete Element  $e$  zusammen mit seinem Index  $i$ . Existiert in  $l$  ein  $j < i$ , sodass  $l_j < l_i$ , so ist  $l_j$  bereits in  $w$  und  $e$  wird am darauffolgenden Index eingesetzt, ansonsten am Index 0. Da stets  $\forall i \in \mathbb{D}_w : e > w_i$  (sofern  $w$  noch nicht  $i$  enthält) und  $e$  an einer Stelle eingefügt ( $\text{Ins}$ ) wird, ist (6) erfüllt und  $w$  eine Liste; da durch eindeutige Zuordnung alle Elemente enthalten sind und die Reihenfolge beibehalten wurde, teilen  $w$  und  $l$  alle Eigenschaften und damit identisch. ■

Somit ist die Konstruierbarkeit aus (5) und (6) als Eigenschaft der Listen festzuhalten.

**Satz.** Jede sortierte Liste reeller Zahlen ist unter der Betragsmetrik kettensortiert.

**Beweis.** Die Definition anwendend, bedeutet dies:

$$\forall l \in \mathbb{L}_{\text{sort}} : d_{\mathbb{L}}(l) = \min_{l' \in \text{Perm}(l)} d_{\mathbb{L}}(l') \quad (10)$$

Da das Minimum einer Menge das Element bedeutet, für das kein kleineres Element existiert, ist eine gleichwertige Formulierung, dass für keine sortierte Liste eine Permutation dieser Liste existiert, die eine kleinere Kettenlänge hat. Dabei ist die Metrik  $d(a, b) = |a - b|$ .

Induktionsbeginn:

Länge 0: Die einzige Liste der Länge 0 ist  $\varepsilon$ , somit ist der Definitionsbereich  $\{\}$ . Nach (8) ist  $\varepsilon \in \mathbb{L}_{\text{sort}}$ . Somit kann keine Funktion eine Änderung der Elemente vornehmen (da keine solchen existieren), und jede sortierte Liste der Länge 0 ist sortiert und kettensortiert.

Länge 1: Eine sortierte Liste  $l$  der Länge 1 besteht aus einem Element, also  $l = l_1$ . Der Definitionsbereich ist  $\{1\}$ , die einzige Permutationsfunktion  $f$  ist  $\{(1, 1)\}^{[2]}$ . Da  $\forall i : f(i) = i$  (Identität), wird keine Änderung der Elemente vorgenommen. Somit existiert keine Liste mit kürzerer Kettenlänge und jede sortierte Liste der Länge 1 ist kettensortiert.

Länge 2: Jede sortierte Liste  $l$  der Länge 2 erfüllt  $l = l_1 \oplus l_2$ , wobei  $l_2 > l_1$ . Die möglichen Permutationsfunktionen sind  $\{(1, 1), (2, 2)\}, \{(1, 2), (2, 1)\}\}$ . Beide ändern die Kettendistanz nicht, da die erste hier aufgeführte die Identität ist und die zweite die Liste umkehrt, was durch Kommutativität der Addition sowie Symmetrie der Metrik die Kettendistanz nicht ändert. Also ist auch jede sortierte Liste der Länge 2 kettensortiert.

Induktionsschritt:

---

<sup>[1]</sup> $e > l_n$  ist an dieser Stelle äquivalent, jedoch wird hier, um die Definitionen analog zu halten, die längere Variante gewählt.

<sup>[2]</sup>Der Kürze halber werden Funktionen in diesem Abschnitt ihrer Definition zufolge als Mengen notiert.

$l$  sei eine sortierte und kettensortierte Liste der Länge  $n \in \mathbb{N}, n \geq 2$ . Gemäß der induktiven Definition einer Liste wird nun ein neues Element  $e \in \Sigma, e > l_n$  an einer beliebigen Position  $i$  in die Liste, welche fortan  $l^e$  genannt werde, eingefügt. Definitionsgemäß bleibt die Liste nur dann sortiert, wenn  $i = n + 1$  – andernfalls wäre  $l_{i+1} < l_i$  und die Liste unsortiert. Es wird nun bewiesen, dass beim Anfügen eines neuen Elements an genau dieser Stelle die Liste kettensoriert bleibt.

Zunächst wird die Kettenlänge betrachtet, die sich ergibt, falls  $i = n + 1$ . In diesem Fall gilt:  $d_{\mathbb{L}}(l^e) = d_{\mathbb{L}}(l) + d(l_n, e)$ . Die Differenz zwischen bisheriger und neuer Kettenlänge wird als  $d_{\mathbb{L}}(l^e) - d_{\mathbb{L}}(l) = d(l_n, e)$  als  $\Delta_{\text{opt}}$  bezeichnet.

$K$  sei die Menge aller Distanzen zwischen  $e$  und einem Element  $l_i$  an Index  $i$  der Liste.  $e > l_i$  ist (wie auch  $e > l_n$ ) gegeben. Da  $i < n$ , folgt (durch Definition von  $\mathbb{L}_{\text{sort}}$  und Transitivität von  $<$ ), dass  $l_i < l_n < e$ . Deshalb gilt  $d(l_n, e) = e - l_n$  und  $d(l_i, e) = e - l_i$ , und somit  $l_i < l_n \Leftrightarrow e - l_i > e - l_n$ . Da  $i$  beliebig gewählt wurde, ist die Distanz zwischen jedem Element und  $e$  größer als die zwischen  $l_n$  und  $e$ . Diese Distanz  $d(l_n, e) = \Delta_{\text{opt}}$  ist folglich von allen Distanzen, die  $e$  involvieren, minimal.

Nun wird jede Permutation von  $l$  betrachtet und  $l'$  genannt. Da  $l$  kettensoriert ist, gilt:  $d_{\mathbb{L}}(l) \leq d_{\mathbb{L}}(l')$ . Auf Basis dieser Permutation wird eine neue Liste  $l^{e'}$  konstruiert, in die das Element  $e$  an einer Stelle  $i$  eingefügt wurde, also  $l^{e'} := l'_{1:i-1} \oplus e \oplus l'_{i:n}$ .

$$\Delta := d_{\mathbb{L}}(l^{e'}) - d_{\mathbb{L}}(l') \quad (11)$$

Es soll gezeigt werden, dass für jede Permutation  $l'$  gilt:

$$\Delta \geq \Delta_{\text{opt}} \quad (12)$$

sodass durch  $d_{\mathbb{L}}(l') \geq d_{\mathbb{L}}(l)$  die Kettendistanz  $d_{\mathbb{L}}(l^{e'}) = d_{\mathbb{L}}(l') + \Delta \geq d_{\mathbb{L}}(l) + \Delta_{\text{opt}} = d_{\mathbb{L}}(l^e)$  ist und somit  $d_{\mathbb{L}}(l^e)$  tatsächlich die optimale Kettendistanz ist, wodurch  $l^e$  kettensoriert wäre.

Es wird eine Fallunterscheidung zwischen solchen Permutationen gemacht, bei denen  $e$  am Rand (am Index  $i = 1 \vee i = n + 1$ ) zu finden ist, und jenen, bei denen es zwischen zwei anderen Elementen ( $i \in [2, n]$ ) vorliegt.

Fall 1.  $i = 1 \vee i = n + 1$

Da  $e$  am Rand eingefügt wird, ist  $d_{\mathbb{L}}(l^{e'}) = d_{\mathbb{L}}(l') + d(l_i, e)$  für ein beliebiges  $i$  aus der Indexmenge von  $l$ . Die Distanz ist dabei ein Element von  $K$ , da es eine Distanz zwischen  $e$  und einem Element von  $l$  ist. Da  $\Delta_{\text{opt}}$  das minimale Element aus  $K$  ist, kann  $\Delta = d(l_i, e)$  nicht geringer sein, sodass Ungleichung (12) zutrifft.

Fall 2.  $i \in [2, n]$

Hierbei liegt  $e$  zwischen zwei Elementen,  $l'_{i-1}$  und  $l'_i$ . Die neue Kettenlänge setzt sich nun zusammen aus der alten Kettenlänge minus der Distanz dieser beiden Elemente plus der Distanz jedes dieser Elemente mit dem neu eingefügten:

$$\Delta = d(l'_{i-1}, e) + d(l'_i, e) - d(l'_{i-1}, l'_i) \quad (13)$$

$$d_{\mathbb{L}}(l^{e'}) = d_{\mathbb{L}}(l') + \Delta \quad (14)$$

Gegeben sind  $l'_{i-1} < e$  und  $l'_i < e$ ; o. B. d. A. wird nun von  $l'_{i-1} < l'_i < e$  ausgegangen, der Beweis kann analog durch Tauschen der beiden Elemente in der Ungleichung geführt werden. Explizit wird hier auf Eigenschaften der Betragsfunktion zurückgegriffen:

$$\begin{aligned}
\Delta &= |l'_{i-1} - e| + |l_i - e| - |l'_{i-1} - l'_i| \\
&= (e - l'_{i-1}) + (e - l'_i) - (l'_i - l'_{i-1}) \quad (\text{da } l'_{i-1} < l'_i < e) \\
&= e - l'_{i-1} + e - l'_i - l'_i + l'_{i-1} \\
&= e + e - l'_i - l'_i - l'_{i-1} + l'_{i-1} \\
&= 2 \cdot e - 2 \cdot l'_i = 2 \cdot (e - l'_i) \\
&= 2 \cdot |l'_i - e| = 2 \cdot d(l'_i, e)
\end{aligned} \tag{15}$$

$\Delta$  ist folglich das Doppelte von  $d(l'_i, e)$ . Dabei handelt es sich um ein Element aus  $K$ , sodass Ungleichung (12) erfüllt ist.

Da keine Permutation der finalen Liste eine höhere Kettendistanz als die von  $l^e$  erzielen kann, ist  $l^e$  kettensortiert. ■

### 2.3. Generalisierung auf n-dimensionale Vektorräume

Nun wurde bewiesen, dass die Sortierung einer Liste eindimensionaler Objekte, in diesem Fall repräsentiert durch reelle Zahlen, äquivalent zur Kettensortierung dieser ist<sup>[3]</sup>. Da die vergleichsbasierte Sortierung nur auf Listen von Elementen einer geordneten Menge ausführbar ist, die distanzbasierte Sortierung jedoch in jedem metrischen Raum angewendet werden kann und die Betragstmetrik der reellen Zahlen nur ein Sonderfall jeder durch eine  $p$ -Norm induzierten Metrik[34] n-dimensionalen Vektorräume ist[51], handelt es sich bei der Kettensortierung um eine Generalisierung der vergleichsbasierten Sortierung im Bezug auf die Dimensionalität der Eingaben.

Somit kann nun im Folgenden die Eigenschaft ‚sortiert‘ auch auf Listen mehrdimensionaler Objekte angewandt werden, da sie bis auf Umkehr der Liste die gleiche Bedeutung wie ‚ketten-sortiert‘ hat. Als Metrik wird sich aufgrund ihrer Generalisierungsfähigkeit von nun an auf eine der durch eine  $p$ -Norm induzierte Metrik beschränkt, o.B.d.A. wird für die folgenden Beispiele die euklidische Metrik (gegeben durch  $p = 2$ ) gewählt.

## 3. Graphentheoretische Grundlagen

### 3.1. Definitionen

Ein **Graph**  $G$  ist zum Zwecke dieser Arbeit<sup>[4]</sup> ein Paar  $(V, E)$ ,  $E \subseteq V^2$  mit **Knotenmenge**  $V$  und **Kantenmenge**  $E$ . Ist  $E = V^2$ , so wird er als **vollständig** bezeichnet. Zusammen mit einer **Kantengewichtsfunktion**  $d : E \rightarrow \mathbb{R}$  gilt der Graph als kantengewichtet (kurz **gewichtet**), der Wert dieser Funktion für eine Kante ist ihr **Gewicht**.

Ein **Weg** ist eine Sequenz  $p = v_1 v_2 \dots v_n$  paarweise verschiedener Knoten. Ein Weg ist ein **Pfad**, wenn gilt:  $\forall i \in [1, n-1] : (p_i, p_{i+1}) \in E$ . Das Gewicht eines solchen Pfades bezeichnet die Summe der Gewichte aller verbindenden Kanten:  $d_p = \sum_{i=0}^{n-1} d(p_i, p_{i+1})$ .

Ein Pfad wird **Hamilton-Pfad** genannt, sofern  $n = |V|$  zutrifft, der Pfad also alle Knoten erreicht. Ein Hamilton-Pfad ist minimal, sofern kein Hamilton-Pfad des Graphen mit niedrigrem Gewicht existiert.

### 3.2. Anwendungen auf die Fragestellung

Um die Kettensortierung einer Liste  $l$  mit Eingabealphabet  $\Sigma \subseteq \mathbb{R}^n$ ,  $n \in \mathbb{N}$  als graphentheoretisches Problem aufzufassen, betrachte man zunächst deren Bildmenge  $Y$ . Für diese konstruiere

---

<sup>[3]</sup>Genauer: jede sortierte Liste ist kettensortiert, jedoch nicht zwangsläufig umgekehrt (aufgrund von Symmetrie der Kettenlänge im Gegensatz zur Sortierung).

<sup>[4]</sup>Multigraphen werden nicht berücksichtigt, da sie für die Fragestellung keine Rolle spielen. In der Fachliteratur wird die hier verwendete Art von Graph oft als „schlichter Graph“ bezeichnet [54, 18], da diese Bezeichnung teilweise jedoch Kantengewichte ausschließt, wird obige Formulierung gewählt.

man nun den gewichteten Distanzgraphen  $G = (V, E) = (Y, Y^2)$  mit Kantengewichtsfunktion  $d(\vec{a}, \vec{b}) = \|\vec{a} - \vec{b}\|$ . Der Graph ist vollständig.

Innerhalb dieses Graphen ist die Liste ein Hamilton-Pfad, denn zwischen jedem Paar aufeinanderfolgender Elemente existiert eine Kante (durch Vollständigkeit) und die Liste enthält alle Knoten (durch Konstruktion aus Bildmenge und Eindeutigkeit der Elemente der Liste). Die Kettendistanz der Liste ist gleich dem Gewicht dieses Pfads.

Da sich die Bildmenge durch Vertauschen von Elementen nicht ändert, ist auch jede Permutation der Liste ein valider Hamilton-Pfad, dessen Gewicht gleich der Kettenlänge der Liste ist ( $d_{\mathbb{L}}(p) \triangleq d_p$ ). Aus der Definition der Ketten sortiertheit folgt, dass eine Liste genau dann ketten sortiert ist, wenn ihr zugehöriger Hamilton-Pfad im Graphen der Bildmenge minimal ist. Somit kann mithilfe eines Algorithmus, der für einen Graphen dessen minimalen Hamilton-Pfad ermitteln kann, eine mehrdimensionale Liste sortiert werden.

## 4. Algorithmen

Für den Fall der eindimensionalen Sortierung existieren zahlreiche vergleichsbasierte Sortieralgorithmen, die sich in Eigenschaften wie asymptotischen Komplexitäten von Raum und Zeit, Stabilität<sup>[5]</sup> und Vorgehensweise unterscheiden [3]. Bubble Sort, Insertion Sort und Selection Sort gehören zu den simpleren Verfahren, die jedoch aufgrund ihrer höheren Komplexität ineffizienter arbeiten; Merge Sort, Quick Sort[42] und Heap Sort dagegen werden auch in der Praxis<sup>[6]</sup> genutzt [58]. Zwei Eigenschaften, die diese Sortieralgorithmen definieren, sind Monotonizität der Ausgabe (im eindimensionalen Fall gleich der in Abschnitt 2.1 definierten Sortiertheit) und Invarianz der Zielmenge.

Nun könnte ein mehrdimensionaler Sortieralgorithmus definiert werden als einer, der eine Liste als Eingabe erhält und eine ketten sortierte Liste gleicher Elemente zurückgibt.

In diesem Abschnitt wird zunächst erläutert, warum eine derartige Definition sich für das Problem dieser Arbeit nicht eignet und stattdessen in zwei das Problem im allgemeinen Fall nicht lösende, aber dennoch in der Praxis sehr nützliche Arten von Algorithmen aufgeteilt werden muss. Im Anschluss werden Beispiele genannt, die ich zum Zweck dieser Arbeit auch in Rust[43] implementiert habe.

### 4.1. Komplexität

Da die tatsächlich messbare Laufzeit eines Algorithmus von zu vielen algorithmisch irrelevanten Faktoren wie Hardware, Eingabestruktur und Ressourcenverfügbarkeit abhängt, wird in der theoretischen Informatik eine andere Methode gewählt, um Algorithmen bezüglich ihres Zeitverbrauchs nur abhängig von der Eingabelänge vergleichen zu können: die Landau-Symbole[71, 83, 10] oder auch (nach dem angloamerikanischen Begriff „big O notation“) O-Notation[81, 27, 57]. Diese bezeichnen das asymptotische Verhalten einer Funktion, d.h., wie sich diese Funktion für beliebig größer werdende Eingaben verhält. Wenn die Anzahl der Schritte, die ein Algorithmus für eine Liste der Eingabelänge  $n$  ausführt<sup>[7]</sup>, nun durch eine Funktion  $f$  beschrieben werden kann, so liegt der Algorithmus in  $\mathcal{O}(g)$ , falls  $f$  asymptotisch nicht schneller wächst als  $g$ , und in  $\Theta(g)$ , falls  $f$  asymptotisch genauso schnell wächst wie  $g$ .

Die zu Beginn dieses Abschnitts genannten vergleichsbasierten Sortieralgorithmen wurden dabei nach ihrer Zeitkomplexität sortiert: die als ‚simpel‘ bezeichneten Algorithmen liegen in

---

<sup>[5]</sup>Da die Elemente in einer Liste nicht mehrmals vorkommen können, spielt Stabilität für die Listen dieser Arbeit keine Rolle. Um Listen mit doppelten Werten dennoch sortieren zu können, sortiere man die deduplizierte Liste und füge die entfernten Werte an den Stellen nach dem gleichwertigen Element ein.

<sup>[6]</sup>Meist werden die Algorithmen in Standardbibliotheken an in der Praxis häufig vorkommende Daten angepasst, wie bei Timsort [8] (Python) oder Pattern-Defeating Quicksort [66] (Rust).

<sup>[7]</sup>Auf die komplexere, aber genauere Definition der Komplexitätsklassen durch Turing-Maschinen wird der Kürze halber nur summarisch<sup>[8]</sup> eingegangen.

$\Theta(n^2)$ , während die ‚praktischeren‘ in  $\Theta(n \cdot \log n)$  liegen. Erstere werden daher als quadratisch, letztere als linearithmisch (ein Kofferwort aus ‚linear‘ und ‚logarithmisch‘) bezeichnet. Entscheidend ist im Folgenden allerdings eine andere Dichotomie, die sich nur darauf bezieht, ob  $g$  durch ein Polynom beschreibbar ist oder nicht: polynomiell oder nicht-polynomiell. Erstere werden als in P und **effizient** beschrieben, während letztere NP-schwer und **ineffizient** sind.<sup>[8]</sup>

#### 4.1.1. Handhabbarkeit

Probleme, die in P liegen, werden als **handhabbar** [65], solche, die mindestens in NP liegen, als **unhandhabbar** bezeichnet. Der Grund liegt in der unterschiedlichen Steigung der Terme für wachsende  $n$ : bei einem Polynom wird stets ein von  $n$  abhängiger Term addiert, bei einer Exponentialfunktion (die häufigste Form eines nicht-polynomiellen Terms) multipliziert. Der Unterschied wird deutlich, wenn man das bekannte Beispiel der Weizenkornlegende[45, 30] betrachtet: hierbei sollte auf das erste Feld eines Schachbretts ein Weizenkorn gelegt werden, und auf jedes folgende der doppelte Betrag des vorherigen[67]. Bei gerade einmal 64 Feldern beträgt die Summe der benötigten Weizenkörner bei Anwendung der Exponentialfunktion  $2^{n-1}$ :  $\sum_{n=1}^{64} 2^{n-1} = 2^{64} - 1 = 18'446'744'073'709'551'615 \approx 18$  Trillionen, während beim hoch gewählten Polynom  $n^5$  die Summe  $\sum_{n=1}^{64} n^5 = 11'997'107'200 \approx 12$  Milliarden nicht übersteigt. Könnte ein landwirtschaftlicher Betrieb in jeder Sekunde eine Million Weizenkörner produzieren, so bräuchte er für die letztere Summe etwa dreieinhalb Stunden, für die erstere dagegen mehr als 580'000 Jahre.

#### 4.1.2. Beweis der Unhandhabbarkeit durch Reduktion

Um zu beweisen, dass ein Problem unhandhabbar ist, wird in der theoretischen Informatik eine **Reduktion** ausgeführt, bei der ein Problem auf ein anderes, für welches diese Eigenschaft bereits bekannt bzw. bewiesen ist, zurückgeführt („reduziert“) wird [52](S. 452-454). Im Folgenden wird das Problem des minimalen Hamilton-Pfads im Fall vollständiger Graphen auf das bewiesenermaßen NP-schwere Hamilton-Pfad-Problem[52](S. 474-479)[41, 36] reduziert, welches aus der Frage besteht, ob zu einem beliebigen gegebenen Graphen ein Hamilton-Pfad *existiert*.

Dazu wird die Hypothese aufgestellt, es gäbe einen Algorithmus  $A$ , der das Problem für jeden vollständigen Graphen in polynomieller Zeit abhängig von der Anzahl der Knoten lösen könnte. Man betrachte nun jeden möglichen (insbesondere unvollständigen) Graphen  $G = (V, E)$ . Man konstruiere nun einen neuen vollständigen gewichteten Graphen  $G' = (V, E', d)$ , wobei  $E' = V^2$  und

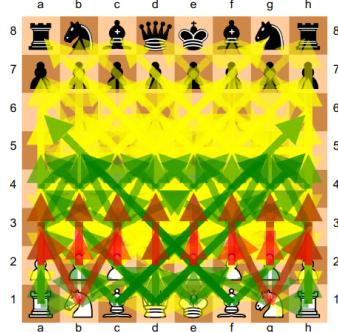


Abbildung 2: Auch die vollständige Spielbaumtraversierung ist im Schach nicht handhabbar, da exponentiell. [13](S. 6-8)

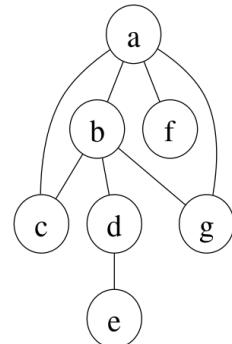


Abbildung 3: Ein Beispiel für den Graphen  $G$ .

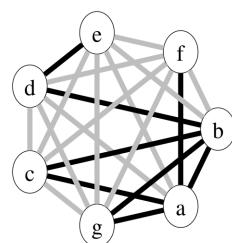


Abbildung 4: Der für  $G$  neu erstellte Graph  $G'$ .

<sup>[8]</sup>In der theoretischen Informatik werden die Komplexitätsklassen als Mengen dargestellt, die Entscheidungsprobleme beinhalten. P ist die Menge aller Probleme, die sich durch eine deterministische Turingmaschine in einer polynomiellen Anzahl von Schritten („effizient“) lösen lässt, während NP die umfasst, die eine nichtdeterministische Turingmaschine effizient lösen kann. Da jedes in P lösbare Problem auch in NP lösbar ist, gilt  $P \subseteq NP$ . (Obwohl in der praktischen Anwendung davon ausgegangen wird, dass  $P \neq NP$ [52](S. 465), bleibt dies ein ungelöstes Problem der Informatik[80]). Dagegen müssen NP-schwere Probleme nicht in NP liegen, sondern die ‚Schwierigkeit‘ dieser Klasse nur übertreffen.

$d(a, b) = \begin{cases} 0 & \text{falls } (a, b) \in E \\ 1 & \text{falls } (a, b) \notin E \end{cases}$ . Nun führe man  $A$  mit der Eingabe  $G'$  aus Schwarze Kanten sind in  $E$ , graue nicht. und erhalte den minimalen Hamilton-Pfad  $p_{\text{opt}}$ .

Anhand dieser Ausgabe kann das Hamilton-Pfad-Problem für den Graphen  $G$  gelöst werden: ist das Gewicht von  $p_{\text{opt}}$  0, so existiert für  $G$  ein Hamilton-Pfad, andernfalls nicht. Dies folgt daraus, dass  $A$  stets den minimalen Hamilton-Pfad auswählt – existiert einer im originalen Graphen, so kann dieser nur aus Kanten mit Gewicht 0 gebildet werden, sonst nicht.<sup>[9]</sup>

Somit entsteht ein Widerspruch: die Aussagen, dass das Hamilton-Pfad-Problem NP-schwer ist, und, dass der in P liegende Algorithmus  $A$  es lösen kann, widersprechen sich. Somit ist die Hypothese falsch,  $A$  notwendigerweise ineffizient und das Problem NP-schwer<sup>[10]</sup>.

## 4.2. Heuristik

Wie bereits in Abschnitt 4.1.1 gezeigt wurde, kann die Kettenortierung im Allgemeinfall nicht in einer sinnvollen Zeitspanne gelöst werden. Aus diesem Grund werden von nun an andere Algorithmen zum Einsatz kommen: **Heuristiken**. Eine Heuristik ist dabei ein effizienter Algorithmus, der ein (zu komplexes) Problem nicht vollständig löst, sondern nur eine ungefähre Lösung liefert.

Im Fall der Sortierung mehrdimensionaler Objekte bedeutet dies, dass ein derartiger Algorithmus strategisch versucht, die Kettenlänge zu reduzieren, jedoch nicht zwangsläufig das globale Minimum (welches der Sortierung der Liste entspräche) findet.

Zudem ist eine weitere auf graphentheoretischer Grundlage beruhende Dichotomie der im Folgenden erläuterten Algorithmen sinnvoll: **Konstruktionsalgorithmen** auf der einen und **Verbesserungsalgorithmen** auf der anderen Seite. Dabei erhalten erstere als Eingabe eine Liste und geben einen Pfad zurück, während letztere versuchen, die Kettenlänge eines bestehenden Pfades zu verringern. Im Bezug auf die verwendeten Datenstrukturen ist dies unerheblich, da auch Pfade als Listen repräsentiert werden; die Nutzerfreundlichkeit der Oberfläche nimmt allerdings zu, da Kombinationen des Ausführens von Algorithmen, die zu einer Annäherung an die Kettenortierung nicht beitragen, grundlegend verhindert werden.<sup>[11]</sup> Grundlegend betrachten Konstruktionsalgorithmen die Liste also eher als ungeordnete Menge, während sie bei Verbesserungsalgorithmen als geordneter Pfad interpretiert wird.

## 4.3. Pfadkonstruktion

Um die folgenden Konstruktionsalgorithmen anwenden zu können, wird eine Beispielliste  $l$  gewählt, welche sich aus sechs zweidimensionalen Vektoren zusammensetzt.

Zur besseren Veranschaulichung werden die Elemente der Liste folglich auch durch Farben im RGB-Farbraum dargestellt. Dabei repräsentieren die Komponenten des Vektors jeweils den Rot- und Blauwert der Farbe im Intervall  $[0, 1]$ , der Grünwert wird auf 0 festgelegt:

$$l := \left[ \begin{pmatrix} 0.6 \\ 0.2 \end{pmatrix}, \begin{pmatrix} 0.1 \\ 0.7 \end{pmatrix}, \begin{pmatrix} 0.9 \\ 0.4 \end{pmatrix}, \begin{pmatrix} 0.2 \\ 0.2 \end{pmatrix}, \begin{pmatrix} 0.8 \\ 0.7 \end{pmatrix}, \begin{pmatrix} 0.5 \\ 0.5 \end{pmatrix} \right] \triangleq \bullet \bullet \bullet \bullet \bullet \bullet \quad (16)$$

---

<sup>[9]</sup>Der Beweis basiert auf [52](S. 479) und [50] und wurde hier statt auf das Travelling Salesman Problem auf das gegebene Problem des minimalen Hamilton-Pfads angewandt. Dies bestätigt erneut die Ähnlichkeit der beiden Probleme.

<sup>[10]</sup>Wichtig zu beachten ist dabei, dass nur das Problem des minimalen Hamilton-Pfad auf beliebigen Graphen NP-schwer sein muss; die im weiteren Verlauf dieser Arbeit betrachteten Distanzgraphen könnten immer noch in P liegen.

<sup>[11]</sup>Ein Beispiel dafür ist, dass es keinen Sinn ergibt, nach dem Ausführen des Greedy-Algorithmus einen anderen Konstruktionsalgorithmus wie Nearest Neighbor zu nutzen, da dieser die vorherigen Ergebnisse nicht berücksichtigt.

Um einen Pfad sowie die darin enthaltenen Vektoren zu visualisieren, wurde ein Programm mittels der JavaScript-Bibliothek p5.js[61] entwickelt, welches die Listenelemente und ihre Farben sowie die dazwischenliegenden Kanten in einem zweidimensionalen kartesischen Koordinatensystem darstellt. Die folgenden derartigen Abbildungen entspringen diesem.

### 4.3.1. Triviale Pfadkonstruktion

Die triviale Pfadkonstruktion gibt die Eingabeliste unverändert zurück – in diesem Fall entsteht der in Abbildung 5 erkennbare Pfad. Dieser Algorithmus ermöglicht es, beliebige Pfade selbst zu erstellen und zu verändern, ohne dabei an eine bestimmte Heuristik gebunden zu sein. Zudem können so die Verbesserungsalgorithmen teils besser dargestellt werden. Eine Alternative dieses Verfahrens ist, die Liste vorher zufällig zu mischen.

### 4.3.2. Brute Force

Der Brute-Force-Algorithmus ist der einzige hier aufgelistete nicht heuristische Algorithmus, der tatsächlich eine Liste mehrdimensional sortieren kann, also das Minimum aller Permutationen findet. Dazu wird jede einzelne dieser Permutationen auf ihre Kettenlänge überprüft und jene mit der minimalen zurückgegeben (siehe Tabelle 1).

Die Zeitkomplexität dieses Algorithmus liegt in  $\Theta(n!)$ , wächst also proportional zur Fakultät der Eingabelänge, da diese zugleich die Anzahl der Permutationen einer derartigen Liste beschreibt. In der Praxis zeigt sich, dass das Verfahren für Listen mit zehn oder weniger Elementen durchaus nutzbar ist, jedoch ab einer Länge von zwölf mit einem geschätzten Zeitaufwand von einer Stunde keine Option mehr darstellt (siehe auch Abschnitt 4.1.1).

Um den NN-Algorithmus unabhängig vom ersten Element der Liste zu gestalten, kann – auf Kosten der Laufzeitkomplexität, die in diesem Fall kubisch wird – der bisherige Algorithmus für alle Rotationen der Originalliste ausgeführt werden und das Ergebnis minimaler Kettenlänge ausgewählt werden.

### 4.3.3. Nächster Nachbar

Beim Nächster-Nachbar-Algorithmus (engl. *nearest neighbor*, kurz NN) handelt es sich um ein sog. naives gieriges Verfahren. Dabei wird vom ersten Punkt der Liste aus begonnen und stets der Punkt ausgewählt und folglich betrachtet, dessen Distanz zum aktuell betrachteten minimal ist und noch nicht im zu erstellenden Pfad enthalten ist, bis alle Punkte im Pfad enthalten sind.

Da es stets einen konkreten Bezugspunkt gibt, von dem aus vorgegangen wird, ist der Algorithmus einfach

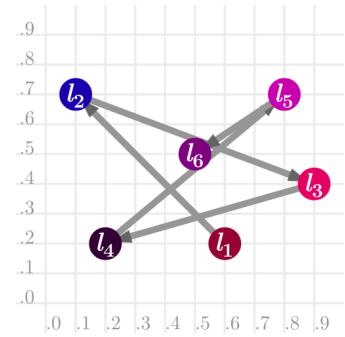


Abbildung 5: Der durch  $l$  gegebene triviale Pfad.

#	$p$	$d_{\mathbb{L}}(p)$	min
1	●●●●●●	3.43	3.43
2	●●●●●●	3.07	3.07
3	●●●●●●	3.18	3.07
4	●●●●●●	2.66	2.66
5	●●●●●●	3.12	2.66
6	●●●●●●	2.62	2.62
...	(229 weitere)		
236	●●●●●●	1.88	1.88
...	(482 weitere)		
718	●●●●●●	3.08	1.88
719	●●●●●●	3.07	1.88
720	●●●●●●	3.43	1.88

Tabelle 1: Brute Force überprüft jede Permutation.

Pfad $p$ ( $d_{\mathbb{L}}(p)$ )	Nachbarn von $p_{ p }$
● (0)	● 0.32 ● 0.36 ● 0.40 ● 0.54 ● 0.71
●● (0.32)	● 0.36 ● 0.41 ● 0.42 ● 0.45
●●● (0.68)	● 0.32 ● 0.70 ● 0.78
●●●● (0.99)	● 0.73 ● 0.85
●●●●● (1.72)	● 0.51
●●●●●● (2.23)	

zu verstehen und zu visualisieren. Für die praktische Anwendung reicht er nicht, da stets nur das nächste Element berücksichtigt und der restliche Kontext vernachlässigt wird – so kommt es insbesondere am Ende des Pfades meist zu besonders langen Kanten hin zu Knoten, die unter Reduzierung der Kettenlänge bereits vorher hätten besucht werden können, es allerdings nicht wurden, da andere Knoten unmittelbar näher lagen. Die Komplexität ist quadratisch, da für jeden hinzuzufügenden Knoten jeder andere Knoten überprüft wird.

Es handelt sich dabei um einen Algorithmus, der von der ursprünglichen Anordnung der Liste abhängig ist, da stets beim ersten Punkt dieser Liste begonnen wird. In Tabelle 3 können die verschiedenen von NN erzeugten Pfade im Fall der Beispieldatei betrachtet werden – aufgrund des unidirektionalen Auswahlverfahrens weisen sie alle eine unterschiedliche Kettenlänge auf.

#### 4.3.4. Greedy

Der Greedy-Algorithmus (auch Multiple Fragments) enumeriert alle Kanten und sortiert sie nach ihrer Länge aufsteigend. Diese werden in einem Stapel abgelegt und stets wird die erste Kante ausgewählt (und entfernt sowie dem Pfad angefügt), die die Validität des Pfads nicht verletzt, also keine bereits zweifach verbundenen Knoten besucht oder Zyklen erstellt.

Sobald die Länge dieser Liste von Kanten gleich der der Eingabe minus eins ist, ist ein validierender Pfad gefunden, da die Validität nicht verletzt wurde und alle Knoten besucht werden. Zudem existiert stets ein derartiger Pfad, da von der Vollständigkeit des Graphen ausgegangen wird. Bei Wahl von zweckmäßigen Datenstrukturen ist die Zeitkomplexität quadratisch ( $O(n^2 \cdot \log n)$ ).

Die Fragmentierung kann analog zu Abschnitt 4.3.3 zum Schluss zu suboptimalen Kanten führen, jedoch wird das Problem der Nichtberücksichtigung von „Außenseitern“ zumeist umgangen, da auch der Weg von einem solchen zu einem beliebigen anderen Knoten zumeist kürzer ist als ein anderer, der zwischen „Clustern“ wechselt.

### 4.4. Pfadverbesserung

Die Pfadverbesserungsalgorithmen werden im Folgenden am Beispiel des eben durch NN (bei Start am ersten Element der Liste) erstellten Pfades,  $p = \text{●●●●●●}$ , erklärt (siehe Abbildung 6).

#### 4.4.1. Rotation

Mit einer linearen Laufzeit ist die Rotation das algorithmisch einfachste Verfahren aus dieser Liste. Dabei wird das Maximum der Kantengewichte aller Kanten des Pfades sowie der

Tabelle 2: Stets wird vom aktuellen Punkt aus der nächste Nachbar besucht.

$p_0$	$p$	$d_L(p)$
●	●●●●●●	2.23
●	●●●●●●	2.22
●	●●●●●●	1.90
●	●●●●●●	2.25
●	●●●●●●	1.93
●	●●●●●●	2.20

Tabelle 3: Die NN-Pfade abhängig vom Anfangspunkt.

	●	●	●	●	●	●
●	(0)	0.71	0.36	0.40	0.54	0.32
●	0.71	(0)	0.85	0.51	0.70	0.45
●	0.36	0.85	(0)	0.73	0.32	0.41
●	0.40	0.51	0.73	(0)	0.78	0.42
●	0.54	0.70	0.32	0.78	(0)	0.36
●	0.32	0.45	0.41	0.42	0.36	(0)

Tabelle 4: Adjazenzmatrix  $A \in \mathbb{R}^{|p| \times |p|}$ :  $A_{i,j} = d(p_i, p_j)$  des Graphen der Bildmenge von  $l$ .

nicht berücksichtigten Kante zwischen Start und Ziel des Pfads ermittelt. Falls diese Kante ein höheres Gewicht als alle im Pfad enthaltenen aufweist, so ist der Pfad bereits rotationsoptimal und kann nicht optimiert werden. Ansonsten wird die maximale Kante aus dem Pfad entfernt und die neue Kante hinzugefügt.

Tabelle 5 zeigt die Anwendung der Rotation auf den Beispelpfad. Hierbei zeigt sich, dass eine im Pfad vorkommende Kante ( $\text{●} \rightarrow \text{●}$ ) ein höheres Gewicht als die ausgelassene ( $\text{●} \rightarrow \text{●}$ ) aufweist. Somit kann durch eine Rotation der den Pfad repräsentierenden Liste die Kettenlänge (wenn auch in diesem Fall nur um  $0.73 - 0.71 = 0.02$  Einheiten) reduziert werden.

#### 4.4.2. Swap

Swap betrachtet den Pfad als Liste und überprüft, ob das Tauschen (Swap nach Abschnitt 2.1) zweier Elemente zu einem Pfad mit geringerer Kettenlänge führt. Bei Konstruktion mittels einer nicht-trivialen Heuristik (wie auch hier) ist dies allerdings selten der Fall, weshalb das Verfahren eher nur als Grundlage für andere wie Abschnitt 4.4.3 und 4.4.5 dient.

#### 4.4.3. 2-opt

Das 2-opt-Verfahren beruht auf der folgenden geometrischen Erkenntnis: sobald sich in der zweidimensionalen geometrischen Repräsentation eines Pfads zwei als Strecken repräsentierte Kanten schneiden, so kann die Kettenlänge des Pfades optimiert werden, indem der Schnittpunkt wie in Abbildung 7 durch Tauschen zweier Kanten entfernt wird. Diese Pfadmodifikation wird als 2-opt-Tausch bezeichnet.

Dabei wird die Kettenlänge stets reduziert, da die neue Strecke zwei Punkte direkt verbindet, statt einen ‚Umweg‘ zu enthalten. Im Bezug auf einen Pfad als Liste bedeutet ein 2-opt-Tausch dabei, eine Teilliste dieser umzukehren. Durch Ungerichtetetheit des Graphen ändert sich die Kettenlänge der Teilliste nicht.

Während die Existenz eines solchen Schnittpunkts im zweidimensionalen kartesischen Koordinatensystem mithilfe eines Sweep-line-Algorithmus in linearithmischer Laufzeit ermittelbar ist[9], existiert ein solches Verfahren für höhere Dimensionen nicht. Aus diesem Grund ist die Laufzeit im allgemeinen Fall pro Optimierungsschritt quadratisch, da jede Kante mit jeder weiteren Kante auf Tauschbarkeit überprüft wird, die auch nicht mehr nur auf Schnittpunkten basiert, sondern die Kettendistanzen vergleicht.

Es handelt sich nur um einen Optimierungsschritt, da häufig weitere Kanten dem Verfahren gemäß getauscht werden können – während das Auflösen aller Überschneidungen im Pessimalfall

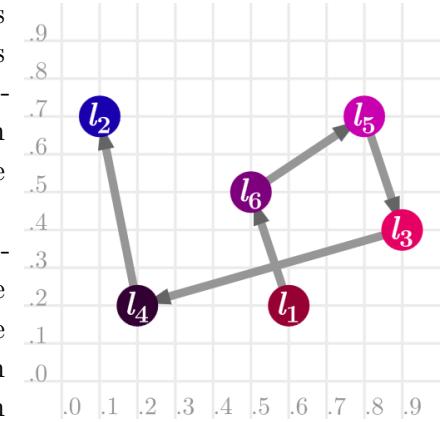


Abbildung 6: Beispelpfad  $p$ .

$i$	$e \in E$	$d(e_0, e_1)$	$e \in p?$
0	● → ●	0.32	ja
1	● → ●	0.36	ja
2	● → ●	0.32	ja
3	● → ●	0.73	ja
4	● → ●	0.51	ja
5	● → ●	0.71	nein

Tabelle 5: Alle im Pfad vorkommenden Distanzen sowie die ausgelassene.

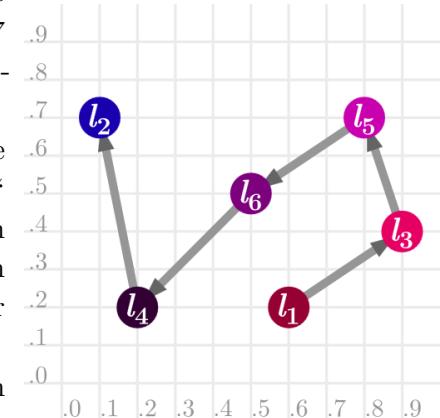


Abbildung 7: Der 2-opt-optimale Pfad  $p'$  nach 2-opt-Tausch von  $l_3$  und  $l_6$ .

eine Laufzeitkomplexität von  $\tilde{O}(n^{10})$ <sup>[12]</sup> [6] benötigt, konvergiert der Algorithmus meist schneller. Ist kein 2-opt-Tausch mehr ausführbar, wird der Pfad 2-opt-optimal genannt.

#### 4.4.4. 3-opt und k-opt

Ähnlich funktioniert das 3-opt-Verfahren: hierbei werden allerdings zwei statt drei Kanten getauscht. Da kein exaktes geometrisches Äquivalent existiert, wird die Liste betrachtet. Für einen 3-opt-Swap des Pfads  $p$  an den paarweise verschiedenen Indizes  $i, j, k \in \mathbb{D}$  existieren dabei vier Möglichkeiten, die verschiedene Teillisten dabei umzukehren oder nicht, wenn die Permutationen der Identität und des 2-opt-Tausches vernachlässigt werden, sonst acht. 3-opt generalisiert im letzteren Fall 2-opt, auch eine Generalisierung auf beliebige  $k \in \mathbb{N}, k \geq 2$  ist möglich. Effizient gelöst werden kann das Problem jedoch durch beliebig ansteigende  $k$  nicht, denn die Auswahl aller möglichen Kanten liegt in  $O(k!)$ .

#### 4.4.5. Simulated Annealing

Simulated Annealing ist ein Verfahren, welches auf der physikalischen Kristallisierung von Materialien beruht[53]. Dabei wird das Swap-Verfahren so erweitert, dass nicht nur solche Kommutationen ausgeführt werden, die kürzere Kettenlängen erzeugen, sondern zu Beginn auf Basis einer Zufallsvariable auch solche, die es nicht tun. Im Laufe der Zeit wird die Wahrscheinlichkeit dafür kontinuierlich reduziert, bis zum Schluss ein swap-optimaler Pfad gefunden wurde – da jedoch ein größerer Teil der möglichen Pfade abgedeckt werden kann, untertrifft er den des Swap-Verfahrens zumeist stark. Im Fall von  $p$  (siehe Abbildung 8) konnte sogar der optimale Pfad gefunden werden.

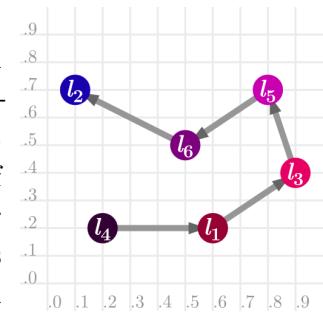


Abbildung 8:  
Simulated Annealing.

Algorithmisch wird dabei eine Starttemperatur  $t \in [0, 1]$  festgelegt und ein Eingabepfad  $p$  erhalten. Man betrachte nun in einem Graphen mit  $G = (\text{Perm}(p), \{p' \mid i, j \in \mathbb{D}, p' := \text{Swap}(p, i, j)\})$  die Nachbarn von  $p$  und wähle einen zufälligen  $p'$  aus, dies geschieht durch Wahl von  $i$  und  $j \neq i$  und Tauschen der Elemente. Falls nun  $d_{\mathbb{L}}(p') < d_{\mathbb{L}}(p)$ , wird mit  $p'$  in jedem Fall weiterverfahren, ansonsten beträgt die Wahrscheinlichkeit dafür  $\exp\left(-\frac{d_{\mathbb{L}}(p') - d_{\mathbb{L}}(p)}{t}\right)$ , sinkt also kontinuierlich mit der Temperatur.  $t$  wird nun arithmetisch um eine Konstante  $k$  verringert, je nach gewünschter Konvergenzgeschwindigkeit – für den Fall der Farbsortierung von hunderten Farben hat sich der Bereich um  $k \approx 10^{-10}$  bewährt – mit  $t \rightarrow 0$  findet kein Tausch mehr statt und der Algorithmus wird abgebrochen.

### 5. Webanwendung

Was nützen die besten Algorithmen, wenn sie nicht auf Daten aus der realen Welt angewandt werden können? Um dies und vieles mehr zu ermöglichen, habe ich eine interaktive Webanwendung entwickelt, die genutzt werden kann, um mehrdimensionale Daten verschiedener Kategorien zu sortieren. Im folgenden Abschnitt werden dabei zunächst die verwendeten Technologien grundlegend erklärt, bevor auf die einzelnen Unterseiten eingegangen wird. Dabei sticht der Abschnitt zur Sortierung von Farben hervor, da dieser meine ursprüngliche Fragestellung beantwortet: Wie können algorithmisch Bücher ästhetisch nach ihrer Farbe sortiert werden?

#### 5.1. Grundlagen und Open-Source-Implementierung

Der Quelltext der gesamten Anwendung ist offen und unter der GNU GPL v3[79], einer freien Copyleft-Lizenz, veröffentlicht. Über GitHub ist der Quelltext des Frontends unter <https://github.com/leo848/jufo2024-frontend>; der des Backends unter <https://github.com/leo848/jufo2024-backend> zu erreichen.

<sup>[12]</sup>Die Tilde signalisiert, dass hier logarithmische Faktoren ignoriert werden.

### 5.1.1. Frontend

Das Frontend der Anwendung habe ich in Svelte[38] und TypeScript[40, 12] geschrieben und verwaltet mittels SvelteKit[39] die Unterseiten. Dabei erfolgt die Modellierung nach dem Component-Prinzip – individuelle Components verwalten dabei sowohl einen Zustand als auch die dazugehörigen (reaktiven) DOM-Elemente. So existiert etwa ein **PathProperties**-Component, welches die Eigenschaften eines Pfades (wie Länge, Kettenlänge und Kettensortiertheit) anzeigt und selbst akquiriert, und analog ein **PathAlgorithms**-Component, das das Ausführen von Konstruktions- und Verbesserungsalgorithmen über die Serververbindung ermöglicht.

### 5.1.2. Backend

Um maximale (zeitliche) Effizienz der Algorithmen mit Speicher- und Typensicherheit sowie der Nutzung eines algebraischen Typensystems zu ermöglichen, habe ich mich dafür entschieden, für das Backend der Anwendung die Programmiersprache Rust[60, 43] zu nutzen. Ein Pfadkonstruktionsalgorithmus ist dabei beispielsweise eine Funktion **fn(PathCreateContext) -> Path**, wobei **Path** ein Typ ist, der eine Liste als Pfad repräsentiert und **PathCreateContext** als **struct PathCreateContext { action: ActionContext, dim: u8, points: Points, norm: Norm }** definiert ist.

Code 1 zeigt die Implementierung von NN in Rust. Dabei wird an einen Pfad (wie in Abschnitt 4.3.3 beschrieben) stets der nächste nicht besuchte Nachbar angehängt. **Point::comparable\_dist** erstellt eine vergleichbare Distanz – im euklidischen Fall wird hier das Quadrat der Distanz zurückgegeben, da es einfacher zu berechnen ist und dennoch (durch Monotonie von  $\sqrt{x}$ ) zur Ermittlung des Minimums genutzt werden kann.

```
pub fn nearest_neighbor(ctx: PathCreateContext) -> Path {
    let PathCreateContext { action, dim, points: values, norm, } = ctx;

    let mut visited = HashSet::new();
    let mut path = Path::try_new(
        vec![values[0].clone()],
        dim).expect("invalid dimension");
    while path.len() != values.len() {
        let last = &path[path.len() - 1];
        visited.insert(last.clone());

        let min = values
            .iter()
            .filter(|&point|
                Not::not(visited.contains(point)))
            .min_by_key(|point|
                point.comparable_dist(last, norm))
            .unwrap();

        path.push(min.clone());
        action.send(
            PathCreation::from_path(path.clone())
                .progress(path.len() as f32
                    / values.len() as f32),
        );
    }
    path
}
```

Code 1: NN als Rust-Programm. Alle anderen aus Abschnitt 4 implementierten Verfahren sind über GitHub (s. Abschnitt 5.1) ersichtlich.

## 5.2. Server-Client-Kommunikation

In bisherigen Projekten (wie [14]) nutzte ich das HTTP-Protokoll, um mit dem Server zu kommunizieren. Dieses zeichnet sich durch ein Request-Response-Schema aus, bei dem auf eine Anfrage (Request) des Clients genau eine Antwort (Response) des Servers erfolgen soll, die idealerweise nur auf den Daten der Anfrage beruht.

Während einige Methoden entwickelt wurden, um diese Einschränkungen aufzuheben oder ihre negativen Auswirkungen zu mindern<sup>[13]</sup>, eignet sich für diese derartig dynamische Anwendung, bei der eine einzige Anfrage hunderte Status-Antworten zur Folge haben sollte, um Responsivität zu gewährleisten, eher ein anderes in Webbrowsersn mittlerweile universell implementiertes Protokoll: das **WebSocket**-Protokoll[29]. Dieses baut auf einem dauerhaft aktiven TCP-Server auf und ermöglicht so bidirektionale zustandsbehaftete Kommunikation.

Server und Client tauschen JSON-Dokumente aus, die auf der Backend-Seite in Rust mithilfe von **serde**[82] und auf der

```
{"type": "action", "latency": 100,
"action": {"type": "createPath", "method": {"type": "nearestNeighbor"}, "dimensions": 3, "values": [[0.6, 0.2], [0.1, 0.7], [0.9, 0.4], [0.2, 0.2], [0.5, 0.5], [0.8, 0.7]]}},

{"type": "pathCreation", "currentEdges": [[[0.6, 0.2], [0.5, 0.5]]], "progress": 0.33333334},
{"type": "pathCreation", "currentEdges": [[[0.6, 0.2], [0.5, 0.5]], [[0.5, 0.5], [0.8, 0.7]]], "progress": 0.5},
 {"type": "pathCreation", "currentEdges": [[[0.6, 0.2], [0.5, 0.5]], [[0.5, 0.5], [0.8, 0.7]], [[0.8, 0.7], [0.8, 0.7]]], "progress": 0.66666667},
 {"type": "pathCreation", "currentEdges": [[[0.6, 0.2], [0.5, 0.5]], [[0.5, 0.5], [0.8, 0.7]], [[0.8, 0.7], [0.9, 0.4]]], "progress": 0.83333334}
```

Frontend-Seite in TypeScript mithilfe von `zod`[62] typensicher deserialisiert werden. In Code 2 wird gezeigt, wie eine solche Kommunikation aussehen kann: zunächst fragt der Client die Erstellung eines Pfads an und übergibt die gewünschte Konstruktionsmethode und Latenz<sup>[14]</sup>, daraufhin schickt der Server für jeden relevanten Schritt eine Antwort zurück, bis der vollständige Pfad ausgegeben wird. Im Fall von NN ist dies für die Responsivität noch nicht entscheidend; bei länger andauernden Prozessen wie mehrschrittigen Verbesserungsalgorithmen dagegen signifikant.

### 5.3. Zahlen sortieren

Auf der Seite `/sort-colors` können ganze Zahlen eingegeben und nach einem Sortierungsalgorithmus der Wahl (aktuell implementiert sind Bubble Sort, Insertion Sort, Selection Sort, Quick Sort und Merge Sort) aufsteigend sortiert werden. Dabei wird bei jedem durch den Server ausgeführten algorithmischen Schritt (wie Vergleich zweier Werte, Vertauschen oder Einstufen als bereits sortiert) die Liste samt der aktuell ausgeführten Schritte zurückgegeben. Dabei kann die Liste neben der üblichen Darstellung ihrer Elemente auch in einem Balkendiagramm dargestellt werden. Abbildung 9 zeigt ein solches Diagramm, das Quick Sort angewandt auf die ersten sechzehn Zahlen der OEIS-Sequenz A107833[73, 76] darstellt.

### 5.4. Vektoren sortieren

Die Seite `/sort-vectors` kann dazu genutzt werden, um eine Liste von Vektoren beliebiger Dimension zu sortieren. Diese Seite erlaubt es, nach Festlegen einer Dimension Vektoren hinzuzufügen, ihre Komponenten zu modifizieren und anschließend mittels `PathAlgorithms` einen Path zu erzeugen. Da jede andere Seite ebenfalls n-dimensionale Objekte sortiert, die als Vektoren repräsentiert werden können, enthalten sie eine Weiterleitung auf diese Seite, welche die Vektoren als URL-Parameter übergibt.

Um die Daten zu visualisieren, wurde ein kraftgerichteter (*force-directed*) Graph-Layout-Algorithmus implementiert. Dieser enthält ein Partikelsystem, sodass jeder Vektor durch ein Partikel repräsentiert wird, auf welches physikalische Kräfte wirken. Dabei existiert zwischen jedem Paar von Vektoren eine Feder mit einer angestrebten Länge der durch die gewählte Metrik gegebenen Distanz, sodass analog zum Hookeschen Gesetz[21] Kräfte auf die betroffenen Partikel wirken. So approximiert die euklidische Distanz zwischen den sehbaren Punkten die genannte Distanz zwischen den Vektoren und skaliert damit mehrdimensional.<sup>[15]</sup> Durch einen einstellbaren Parameter des Strömungswiderstands kann zwischen einer zeitintensiveren, genaueren Darstellung und einer schnelleren und damit ungenauerer Konvergenz gewählt werden.

```
[0.8,0.7],[[0.8,0.7],[0.9,0.4]],  
[[0.9,0.4],[0.2,0.2]]],"progress":  
0.8333333  
{"type":"pathCreation","donePath":  
[[0.6,0.2],[0.5,0.5],[0.8,0.7],  
[0.9,0.4],[0.2,0.2],  
[0.1,0.7]],"currentEdges":[[[0.6,0.2],  
[0.5,0.5]],[[0.5,0.5],[0.8,0.7]],  
[[0.8,0.7],[0.9,0.4]],[[0.9,0.4],  
[0.2,0.2]],[[0.2,0.2],  
[0.1,0.7]]],"progress":1.0}
```

Code 2: Eine Anfrage (nach Abschnitt 4.3.3) und die fünf darauffolgenden Antworten.

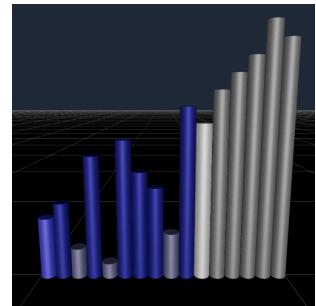


Abbildung 9: Quick Sort partitioniert die Teilliste.

<sup>[13]</sup>Dazu zählt das wiederholte Anfragen einer Ressource vom Server oder das in HTTP/2 implementierte Push-Modell. Während erstere negative Auswirkungen auf die Performanz der Anwendung hat, ist zweitere nicht universell nutzbar und keine Alternative für den gewählten Servertyp.

<sup>[14]</sup>Falls der Server zu viele Antworten in zu kurzer Zeit verschickt, kann dies zu Unresponsivität des Clients und einem potenziellen Speicherleck führen – aus diesem Grund kann die Latenz, die der Server vor Abschicken einer jeden Antwort wartet, hier konfiguriert werden.

<sup>[15]</sup>Im zweidimensionalen euklidischen Fall konvergiert der Graph dabei zu einer Rotation der tatsächlichen Punkte.

## 5.5. Orte sortieren

Die Probleme des bilokal sesshaften, polylokal handelnden Händlers, der drohnenbasierten Paketzustellung sowie der optimalen U-Bahn-Strecke werden durch die Seite [/sort-places](#) lösbar. Dabei zeigt eine interaktive auf der Basis von Leaflet[2] und über OpenStreetMap[23] auch der Mercator-Projektion entwickelte Karte die Punkte sowie den Pfad an.

Als Beispiel wurden für Abbildung 10 zehn Kirchen auf Marler Stadtgebiet ausgewählt, zwischen denen eine U-Bahn verkehren soll. Dazu wurden die geographischen Positionen dieser auf der Website eingegeben und im Anschluss mittels Brute Force (Abschnitt 4.3.2) der kürzeste Pfad konstruiert. In diesem Fall beginnt die neue Linie an der katholischen Kirche St. Georg und endet bei Herz Jesu am Pastoratsberg in Hüls. Ein stets all diese Punkte besuchen müssender Händler<sup>[16]</sup> sollte ebenfalls am genannten Start- und Endpunkt seine Wagenlager aufbauen und den gezeigten Weg nutzen.

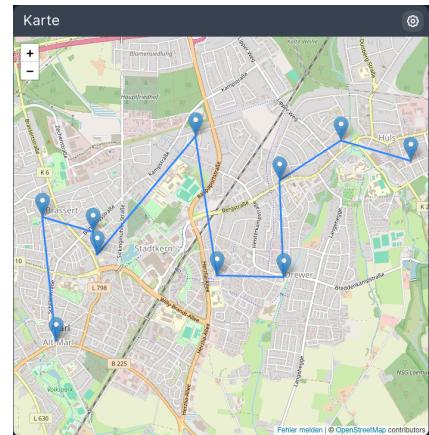


Abbildung 10: Optimale U-Bahn-Linie entlang 10 Kirchen Marls als Kettensortierung einer Liste ihrer Orte.

## 5.6. Farben sortieren

Im Alltag gibt es viele Dinge, die nach Farben sortiert werden können – während Bücher zumeist alphabetisch oder nach Kategorien sortiert werden sollten, ist dies bei Malstiften jeglicher Art, farblicher Dekoration und womöglich auch Kleidung anders. Hier ergibt eine Farbsortierung Sinn, und die meisten Menschen haben eine intuitive Vorstellung davon, was das bedeutet: ähnliche Farben gehören nah zueinander und unterschiedliche auseinander – es scheint also eine quantifizierbare intuitive Distanz zwischen zwei Farben zu geben. Während Ästhetik subjektiv bleibt und daher nicht die ästhetischste Liste für jeden existieren kann, treffe ich die Annahme, dass die Kettensortierung einer Liste am ästhetischsten ist, da sie Farbunterschiede minimiert.

Menschen sind Trichromaten[44, 16], was bedeutet, dass sie drei verschiedene Arten von Zapfen besitzen, die jeweils für eine bestimmte Wellenlängenreichweite des sichtbaren Lichts empfindlich sind, dessen Intensität messen und die Information ans Gehirn weiterleiten. Abbildung 11 zeigt ein Diagramm der Empfindlichkeit abhängig von der Wellenlänge. Im Gegensatz zu Fischen mit vier[17] und Hunden mit zwei[63] braucht es beim Menschen folglich drei Dimensionen, um jede Farbe repräsentieren zu können.

### 5.6.1. Farbräume

Dabei gibt es eine Vielzahl an Farbräumen[78, 49], die allesamt Farben als dreidimensionale Vektoren[24] encodieren. Der womöglich bekannteste[5] ist **sRGB**, der eine Rot-, eine Grün- und eine Blau-Komponente enthält (und der in Abschnitt 4 bereits genutzt wurde, um Vektoren zu illustrieren). Daraus setzt sich jede auf einem Computerbildschirm darstellbare Farbe zusammen: jeder Pixel besteht aus drei Subpixeln, die jeweils R, G oder B in einer bestimmten Intensität anzeigen. Zur Auswahl von Farben eignet sich **HSV**[77] dagegen besser[26] – hier repräsentieren die Komponenten eines Vektors den Farbton (**Hue**), die Sättigung der Farbe sowie die Helligkeit (**Value**). Dieses Farbmodell deckt ebenfalls alle sRGB-Farben ab.

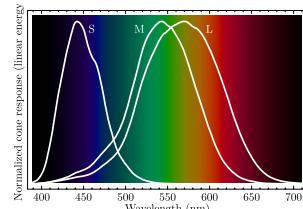


Abbildung 11:  
Empfindlichkeit  
menschlicher Zapfen.  
[11]

<sup>[16]</sup>In diesem Fall wäre ‚permanent kircheninteressierter Tourist‘ wohl zutreffender.

Beide können jedoch nicht dazu genutzt werden, um Farben nach ihrem Aussehen in der realen Welt zu vergleichen – in sRGB sind die Farben  $\vec{c}_1 := \textcolor{red}{\bullet}$  und  $\vec{c}_2 := \textcolor{magenta}{\bullet}$  genauso weit entfernt wie  $\vec{c}_3 := \textcolor{purple}{\bullet}$  und  $\vec{c}_4 := \textcolor{green}{\bullet}$ , obwohl die ersten beiden viel ähnlicher erscheinen.

Aus diesem Grund (und weiteren) wurde das perzeptuelle **OKLAB**-Farbsystem[64] entwickelt. Perzeptuell bedeutet zum Zwecke dieser Arbeit, dass die euklidische Distanz zwischen zwei OKLAB-Farbwerten den wahrgenommenen Abstand modelliert, und dass Eigenschaften wie Farbwert, Sättigung und Helligkeit experimentellen Daten eher entsprechen[56]. LAB bezieht sich darauf, dass das Farbsystem Farben als Helligkeit (**Luminosity**) sowie zwei Farbwerten, **a** und **b**, repräsentiert. In diesem Farbsystem beträgt  $|\vec{c}_1 - \vec{c}_2| \approx 0.39$ , während  $|\vec{c}_3 - \vec{c}_4| \approx 0.69$ .

In der Webanwendung wurden diese und noch drei weitere (linear-sRGB, CMY und HSL – weitere folgen) implementiert und können ineinander konvertiert werden.

### 5.6.2. Farbauswahl

Um Farben hinzuzufügen, habe ich einen Farbauswahldialog entworfen (Abbildung 12), der intuitiv nutzbar ist und zugleich alle Möglichkeiten der Farbauswahl abdeckt. Jeder Bestandteil ist interaktiv und reaktiv, passt sich also der ausgewählten Farbe direkt bei Veränderung an und ermöglicht durch Auswahl des Farbraums, beliebige Komponenten der Farbe zu verändern. Zudem können verschiedene Listen zur Benennung der Liste ausgewählt werden, wie etwa HTML-, X11- oder RAL-Farben.

### 5.6.3. Visualisierung

Die Farben werden dem gewählten Farbraum entsprechend in einem dreidimensionalen Koordinatensystem mittels three.js als Kugeln angezeigt, die Kanten des Pfads als zwischen diesen liegende Zylinder. Die Achsen repräsentieren die Bedeutung der Koordinate. Der Projektionstyp ist einstellbar und über die Maus kann die Darstellung skaliert und rotiert werden. In Abbildung 13 wird ein 3-opt-optimaler Pfad zwischen 18 zufällige Farben im OKLAB-Farbraum gezeigt; Abbildung 14 zeigt einen Screenshot der vollständigen Anwendung auf 40 Farben in HSL.



Abbildung 12: Der Farbauswahldialog. Aktuell ist Ockergelb im RGB-Farbraum ausgewählt.

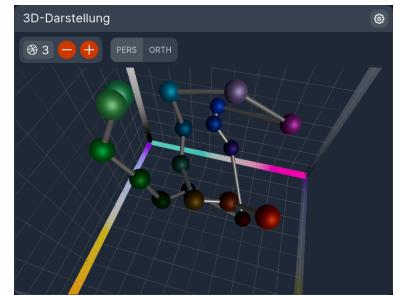


Abbildung 13: 3D-Visualisierung.

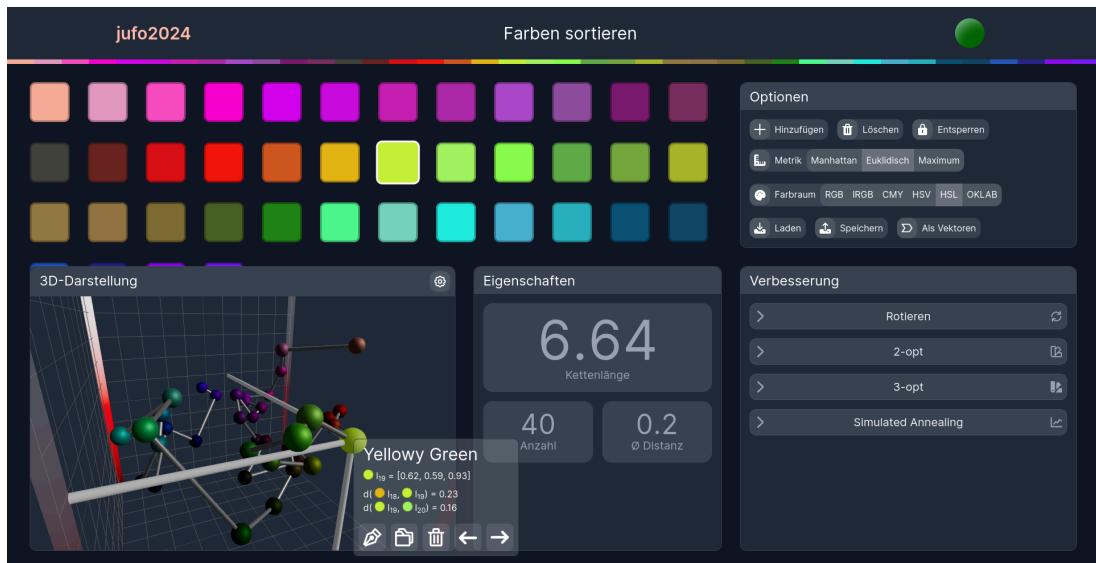


Abbildung 14: Die vollständige Farbsortierungsseite.

## 6. Fazit und Ausblick

In diesem Projekt ist es mir gelungen, die vergleichsbasierte Sortierung auf n-dimensionale Daten zu generalisieren und im Anschluss eine Webanwendung zu entwickeln, mithilfe derer diese Sortierung auch in der Praxis anwendbar wird.

Überrascht hat mich zunächst die Komplexität des Problems. Schließlich werden vergleichsbasierte Sortierungsalgorithmen bereits jetzt universell eingesetzt und die verschiedenen Ansätze sind – auch, wenn immer noch an Mikrooptimierungen gefeilt wird[59] – mittlerweile im algorithmischen Repertoire und der Fachliteratur etabliert[4]. Anders ist das bei der mehrdimensionalen Sortierung, zu der (meiner Kenntnis nach) keine Bibliotheken oder Methoden existieren, die sich explizit zum Ziel setzen, eine Liste mehrdimensionaler Daten zu sortieren. Die NP-Schwere ist nach Erkenntnis der Äquivalenz zum Hamilton-Pfad einleuchtend; ohne das Problem graphentheoretisch zu betrachten, hätte ich sie jedoch nicht erkannt.

Insbesondere im Bezug auf Farben ist die entwickelte Webanwendung mittlerweile ausgereift und kam auch im Alltag bereits zum Einsatz, um Buntstifte, Nagellackfläschchen und Bücher nach Farben zu sortieren. Die Antwort auf die Fragestellung, ob Bücher nach Farben sortiert werden können lautet also: „ja, mit Einschränkungen“. Für die in einem Regal mit unterschiedlichen Farben ist dies (abhängig von der Anzahl der Bücher) zumeist möglich (s. Abbildung 1); für in Bibliotheken vorkommende Mengen dagegen nur approximativ.

Besonders fasziniert und motiviert hat mich die Interdisziplinarität der Thematik bei Entwicklung und Ausarbeitung des Projekts. Während ich mit einem rein mathematischen Problem begann, entwickelte es sich über die Graphentheorie hin zu einem praktisch-algorithmischen der Informatik; in der Visualisierung der Ansätze nutze ich das Hookesche Gesetz aus der Physik (Abschnitt 5.4) und die Mercator-Projektion aus der Kartographie (Abschnitt 5.5). Algorithmen beinhalten neben klassischen Ansätzen der Informatik auch solche mit Bezug zur Materialwissenschaft / Chemie (Simulated Annealing) und in Zukunft womöglich ein Ameisenkolonieverfahren[25] auf Basis biologischer Systeme und Bionik. Nicht zu vernachlässigen ist auch Farbtheorie und -lehre, mit der sich seit der Antike schon Aristoteles[20, 75], da Vinci[1], Newton[74, 19], Werner[85], Goethe[35, 19, 68], Kant[48, 69] und Wittgenstein[72, 84], um nur einige zu nennen, bereits auseinandergesetzt haben.

Allerdings ist die Arbeit bezüglich dieser Thematik noch lange nicht abgeschlossen. Während bereits der Beweis erbracht wurde, dass alle sortierten Listen kettensortiert sind, bedeutet dies noch nicht, dass alle nicht sortierten Listen nicht kettensortiert sind. Dieser mathematische Gegenbeweis muss noch ausgeführt werden, um die Generalisierung zu vervollständigen bzw. weitere Eigenschaften der Kettensortierung zu beweisen.

Während mithilfe von Simulated Annealing schon jetzt sehr gute Pfade heuristisch erzeugt werden können, sollte mein Fokus auch darauf liegen, an einem nicht-heuristischen Algorithmus, der nicht alle Permutationen überprüfen muss, zu arbeiten. Dazu werde ich genauer versuchen, die bereits erfolgreich auf das TSP angewandte Lagrange-Relaxierung[47, 55] anzuwenden und damit auch für Listen größerer Länge mehrdimensional sortieren zu können. Auch eine Implementierung von heuristischen TSP-Pfadverbesserungen wie dem genannten Ameisenkolonieverfahren[25] und komplexeren TSP-Pfadkonstruktionen wie Christofides[22] können in Zukunft die Möglichkeiten der Anwendung noch mehr erweitern.

## Literaturverzeichnis

Sofern nicht anders angegeben, handelt es sich bei allen Abbildungen und Tabellen um eigene Darstellungen von Graphen wurden mit dem Programm Graphviz[33] (und den darin implementierten Algorithmen Dot[31] und Circo[32]) erstellt.

- [1] Ackerman, J. S. 1980. „On early renaissance color theory and practice“. *Memoirs of the American Academy in Rome* 35:11–44.
- [2] Agafonkin, V., I. S. Ortega, D. Leaver, und andere. „Leaflet – a JavaScript library for interactive maps“. [Online]. Verfügbar unter: <https://leafletjs.com/> [zuletzt geprüft: Januar 6, 2024].
- [3] Akhter, N., M. Idrees, und Furqan-ur-Rehman. 2016. „Sorting Algorithms – A Comparative Study“. *International Journal of Computer Science and Information Security*, 14:930–936.
- [4] Al-Kharabsheh, K. S., I. M. AlTurani, A. M. I. AlTurani, und N. I. Zanoon. 2013. „Review on sorting algorithms a comparative study“. *International Journal of Computer Science and Security (IJCSS)* 7(3):120–126.
- [5] Anderson, M., R. Motta, S. Chandrasekar, und M. Stokes. 1996. „Proposal for a standard default color space for the internet—srgb“. In *Color and imaging conference*. S. 238–245.
- [6] Arora, S. 1998. „Polynomial time approximation schemes for Euclidean traveling salesman and other geometric problems“. *Journal of the ACM (JACM)* 45(5):753–782.
- [7] Asteroth, A., und C. Bayer. 2003. *Theoretische Informatik*. Bonn: Pearson Studium.
- [8] Auger, N., V. Jugé, C. Nicaud, und C. Pivoteau. 2019. „On the Worst-Case Complexity of TimSort“.
- [9] Balaban, I. J. 1995. „An optimal algorithm for finding segments intersections“. In *Proceedings of the eleventh annual symposium on Computational geometry*. S. 211–219.
- [10] Beer, S. 2018. *Vergleich und Analyse von Partitionierungsalgorithmen für Quicksort*.
- [11] BenRG. 2009. „Cone fundamentals with srgb spectrum“. [Online]. Verfügbar unter: <https://commons.wikimedia.org/wiki/File:Cone-fundamentals-with-srgb-spectrum.svg> [zuletzt geprüft: Januar 6, 2024].

- [12] Bierman, G., M. Abadi, und M. Torgersen. 2014. „Understanding TypeScript“. In *European Conference on Object-Oriented Programming*. S. 257–281.
- [13] Blume, L. 2021. „Effizienzanalyse des Minimax-Algorithmus im Bezug auf Schach“. [Online]. Verfügbar unter: [https://wv.jugend-forsch.de/media/2021/project\\_91120/description/description\\_2021-02-26\\_00-15-51.pdf](https://wv.jugend-forsch.de/media/2021/project_91120/description/description_2021-02-26_00-15-51.pdf) [zuletzt geprüft: Dezember 14, 2023].
- [14] Blume, L. 2022. „Erweiterung klassischer Unterrichtsmedien durch intuitive Webserviceanwendung“. [Online]. Verfügbar unter: [https://wv.jugend-forsch.de/media/2021/project\\_93889/description/description\\_2022-01-16\\_10-57-20.pdf](https://wv.jugend-forsch.de/media/2021/project_93889/description/description_2022-01-16_10-57-20.pdf) [zuletzt geprüft: Dezember 16, 2023].
- [15] Bohnacker, H., B. Gross, J. Laub, und C. Lazzeroni. 2012. *Generative design: visualize, program, and create with processing*. Princeton Architectural Press.
- [16] Bompas, A., G. Kendall, und P. Sumner. 2013. „Spotting fruit versus picking fruit as the selective advantage of human colour vision“. *i-Perception* 4(2):84–94.
- [17] Bowmaker, J., und Y. Kunz. 1987. „Ultraviolet receptors, tetrachromatic colour vision and retinal mosaics in the brown trout (*Salmo trutta*): age-dependent changes“. *Vision research* 27(12):2101–2108.
- [18] Brandstädt, A. 1994. „Graphen und algorithmische Graphenprobleme“. In *Graphen und Algorithmen*. Wiesbaden: Vieweg+Teubner Verlag, S. 12. [Online]. Verfügbar unter: [https://doi.org/10.1007/978-3-322-94689-8\\_1](https://doi.org/10.1007/978-3-322-94689-8_1).
- [19] Burwick, F. 2012. *The damnation of Newton: Goethe's color theory and romantic perception*. Walter de Gruyter.
- [20] Caston, V. 2018. „Aristotle on the Reality of Colors and Other Perceptible Qualities“. *Res Philosophica* 95(1):35–68.
- [21] Chmelka, F., und E. Melan. 1972. „Spannung und Verformung. Das Hookesche Gesetz“. *Einführung in die Festigkeitslehre für Studierende des Bauwesens*:26–31.
- [22] Christofides, N. 1976. „Worst-case analysis of a new heuristic for the travelling salesman problem“.
- [23] Coast, S. 2004. „OpenStreetMap“. [Online]. Verfügbar unter: <https://www.openstreetmap.org/> [zuletzt geprüft: Januar 3, 2024].

- [24] Cohen, J., und T. P. Friden. 1975. „The Euclidean nature of color space“. *Bulletin of the Psychonomic Society* 5(2):159–161.
- [25] Dorigo, M., M. Birattari, und T. Stutzle. 2006. „Ant colony optimization“. *IEEE computational intelligence magazine* 1(4):28–39.
- [26] Douglas, S. A., und A. E. Kirkpatrick. 1999. „Model and representation: the effect of visual feedback on human performance in a color picker interface“. *ACM Transactions on Graphics (TOG)* 18(2):96–127.
- [27] Dörn, S. 2016. „Entwicklung von Computerprogrammen“. *Programmieren für Ingenieure und Naturwissenschaftler: Grundlagen*:95–115.
- [28] Esponda, M. 2012. „Induktion und Rekursion“. :18–28. [Online]. Verfügbar unter: [http://www.inf.fu-berlin.de/lehre/WS12/ALP1/lectures/V18\\_ALPI\\_Strukturelle\\_Induktion\\_2013.key.pdf](http://www.inf.fu-berlin.de/lehre/WS12/ALP1/lectures/V18_ALPI_Strukturelle_Induktion_2013.key.pdf) [zuletzt geprüft: Dezember 22, 2023].
- [29] Fette, I., und A. Melnikov. 2011. „The websocket protocol“.
- [30] Frater, H., und N. Podbregar. 2005. „Die Legende vom Weizenkorn“. [Online]. Verfügbar unter: <https://www.wissenschaft.de/allgemein/die-legende-vom-weizenkorn/> [zuletzt geprüft: Dezember 23, 2023].
- [31] Gansner, E. R., E. Koutsofios, S. C. North, und K.-P. Yo. „A method for Drawing Directed Graphs“. [Online]. Verfügbar unter: [https://graphviz.org/documentation/TSE\\_93.pdf](https://graphviz.org/documentation/TSE_93.pdf) [zuletzt geprüft: Januar 7, 2024].
- [32] Gansner, E. R., und Y. Koren. „Improved Circular Layouts“. [Online]. Verfügbar unter: <https://graphviz.org/documentation/GK06.pdf> [zuletzt geprüft: Januar 7, 2024].
- [33] Gansner, E. R., und S. C. North. 1999. „An open graph visualization system and its applications to software engineering“. *Software – Practice and Experience*. [Online]. Verfügbar unter: <https://graphviz.org/documentation/GN99.pdf> [zuletzt geprüft: Januar 7, 2024].
- [34] Glaubitz, J., D. Rademacher, und T. Sonar. 2019. „Metrik, Norm, Topologie“. In *Lernbuch Analysis 1: Das Wichtigste ausführlich für Bachelor und Lehramt*. Wiesbaden: Springer Fachmedien Wiesbaden, S. 389–411. [Online]. Verfügbar unter: [https://doi.org/10.1007/978-3-658-26937-1\\_13](https://doi.org/10.1007/978-3-658-26937-1_13).
- [35] Goethe, J. W. von. 1810. „Zur Farbenlehre“. In *Goethe – Die Schriften zur Naturwissenschaft*. Leopoldina-Ausgabe.

- [36] Gurevich, Y., und S. Shelah. 1987. „Expected computation time for Hamiltonian path problem“. *SIAM Journal on Computing* 16(3):486–502.
- [37] Halbeisen, L., und R. Krapf. 2020. „The Axioms of Set Theory (ZFC)“. In *Gödel's Theorems and Zermelo's Axioms: A Firm Foundation of Mathematics*. Cham: Springer International Publishing, S. 153–171. [Online]. Verfügbar unter: [https://doi.org/10.1007/978-3-030-52279-7\\_13](https://doi.org/10.1007/978-3-030-52279-7_13).
- [38] Harris, R., A. Faubert, T. L. Hau, B. McCann, und andere. 2016. „Svelte – cybernetically enhanced web apps“. [Online]. Verfügbar unter: <https://svelte.dev/> [zuletzt geprüft: Januar 3, 2024].
- [39] Harris, R., A. Faubert, T. L. Hau, B. McCann, und andere. „SvelteKit: Web development, streamlined.“. [Online]. Verfügbar unter: <https://kit.svelte.dev/> [zuletzt geprüft: Januar 2, 2024].
- [40] Hejlsberg, A. 2012. „TypeScript: JavaScript with types“. [Online]. Verfügbar unter: <https://www.typescriptlang.org/> [zuletzt geprüft: Januar 2, 2024].
- [41] Held, M., und R. Karp. 1956. „The construction of discrete dynamic programming algorithms“. *IBM Systems Journal* 4(2):136–147.
- [42] Hoare, C. A. R. 1961. „Algorithm 64: Quicksort“. *Communications of the ACM* 4(7): 321.
- [43] Hoare, G., und andere. 2015. „A language empowering everyone to build to build reliable and efficient software.“. [Online]. Verfügbar unter: <https://www.rust-lang.org/> [zuletzt geprüft: Januar 3, 2024].
- [44] Hofer, H., J. Carroll, J. Neitz, M. Neitz, und D. R. Williams. 2005. „Organization of the human trichromatic cone mosaic“. *Journal of Neuroscience* 25(42):9669–9679.
- [45] Hoffbauer, L. 2020. „Die Weizenkornlegende: Schach und exponentielles Wachstum“. [Online]. Verfügbar unter: <https://schachmatt.net/ratgeber/schach-weizenkornlegende/> [zuletzt geprüft: Dezember 23, 2023].
- [46] Hopcraft, J. E., R. Motwani, und J. D. Ullman. 2002. *Einführung in die Automatentheorie, Formale Sprachen und Komplexitätstheorie*. Pearson Studium.
- [47] Huber, S. 2014. *Bündelmethoden für Lagrange-Relaxationen: Anwendung auf ein Problem der ambulanten Krankenpflege*.

- [48] Jahn, T. 2023. „Zwei mögliche Wege mit dem Dilemma umzugehen“. *Die Eigenarten der Farben*:209–233.
- [49] Joblove, G. H., und D. Greenberg. 1978. „Color spaces for computer graphics“. In *Proceedings of the 5th annual conference on Computer graphics and interactive techniques*. S. 20–25.
- [50] Kingsford, C. „CMSC 451: SAT, Coloring, Hamiltonian Cycle, TSP“. [Online]. Verfügbar unter: <https://www.cs.cmu.edu/~ckingsf/bioinfo-lectures/sat.pdf> [zuletzt geprüft: Januar 6, 2024].
- [51] Kivinen, J., M. K. Warmuth, und B. Hassibi. 2006. „The p-norm generalization of the LMS algorithm for adaptive filtering“. *IEEE Transactions on Signal Processing* 54(5): 1782–1793.
- [52] Kleinberg, J., und E. Tardos. 2005. *Algorithm Design* 1. Aufl. Pearson Education, Inc.
- [53] Van Laarhoven, P. J., und E. H. Aarts. 1987. *Simulated Annealing*. Springer.
- [54] Lasch, R. 2020. „Grundlagen der Graphentheorie“. In *Strategisches und operatives Logistikmanagement: Distribution*. Wiesbaden: Springer Fachmedien Wiesbaden, S. 14. [Online]. Verfügbar unter: [https://doi.org/10.1007/978-3-658-31869-7\\_2](https://doi.org/10.1007/978-3-658-31869-7_2).
- [55] Lemaréchal, C. 2001. „Lagrangian relaxation“. *Computational combinatorial optimization: optimal or provably near-optimal solutions*:112–156.
- [56] Lilley, C. 2023. „Color on the Web“. *Fundamentals and Applications of Colour Engineering*:271–291.
- [57] Läuchli, P. 1991. „Komplexität“. *Algorithmische Graphentheorie*:17–24.
- [58] Mahmoud, H. M. 2000. *Sorting: A distribution theory*. John Wiley & Sons.
- [59] Mankowitz, D. J., A. Michi, A. Zhernov, M. Gelmi, M. Selvi, C. Paduraru, E. Leurent, S. Iqbal, J.-B. Lespiau, A. Ahern, und andere. 2023. „Faster sorting algorithms discovered using deep reinforcement learning“. *Nature* 618(7964):257–263.
- [60] Matsakis, N. D., und F. S. Klock II. 2014. „The Rust language“. In *ACM SIGAda Ada Letters*. S. 103–104.
- [61] McCarthy, L. L., Q. Ye, und D. Shiffman. „home | p5js.org“. [Online]. Verfügbar unter: <https://p5js.org/> [zuletzt geprüft: Januar 7, 2024].

- [62] McDonnell, C., und andere. „TypeScript-first schema validation with static type inference“. [Online]. Verfügbar unter: <https://zod.dev/> [zuletzt geprüft: Januar 4, 2024].
- [63] Neitz, J., T. Geist, und G. H. Jacobs. 1989. „Color vision in the dog“. *Visual neuroscience* 3(2):119–125.
- [64] Ottoson, B. 2020. „A perceptual color space for image processing“. [Online]. Verfügbar unter: <https://bottosson.github.io/posts/oklab/> [zuletzt geprüft: Januar 7, 2024].
- [65] O A. „Theoretische Informatik im Rückblick“. [Online]. Verfügbar unter: <https://www.cs.uni-potsdam.de/ti/lehre/03-Theorie-II/slides-chapter9.pdf> [zuletzt geprüft: Januar 5, 2024].
- [66] Peters, O. R. L. 2021. „Pattern-defeating Quicksort“.
- [67] Pickover, C. A. 2009. „The Math Book: From Pythagoras to the 57th Dimension“.: 102.
- [68] Ribe, N., und F. Steinle. 2002. „Exploratory experimentation: Goethe, Land, and color theory“. *Physics today* 55(7):43–49.
- [69] Riley, C. A. 1995. *Color codes: Modern theories of color in philosophy, painting and architecture, literature, music, and psychology*. UPNE.
- [70] Rollnik, S. 2022. „Vollständige Induktion“. *Übungsbuch fürs erfolgreiche Staatsexamen in der Mathematik: Aufgaben und Lösungen für angehende Lehrkräfte der Sekundarstufe I*:45–50.
- [71] Scherer, W. 2016. „Anhang C--Landau-Symbole“. *Mathematik der Quanteninformatik: Eine Einführung*:267–268.
- [72] Schwarte, L. 2015. „Farbliche Evidenzerzeugung“. *Pikturale Evidenz*:131–143.
- [73] Seidov, Z. 2005. „A107833“. [Online]. Verfügbar unter: <https://oeis.org/A107833> [zuletzt geprüft: Januar 6, 2024].
- [74] Shapiro, A. E. 1994. „Artists' colors and Newton's colors“. *Isis* 85(4):600–630.
- [75] Silverman, A. 1989. „Color and color-perception in Aristotle's De Anima“. *Ancient Philosophy* 9(2):271–292.

- [76] Sloane, N. 2024. „The On-Line Encyclopedia Of Integer Sequences“. [Online]. Verfügbar unter: <http://oeis.org/> [zuletzt geprüft: Januar 6, 2024].
- [77] Smith, A. R. 1978. „Color gamut transform pairs“. *ACM Siggraph Computer Graphics* 12(3):12–19.
- [78] Spencer, D. E. 1943. „Adaptation in color space“. *JOSA* 33(1):10–17.
- [79] Stallman, R. 2007. „GNU General Public License“. [Online]. Verfügbar unter: <https://www.gnu.org/licenses/gpl-3.0.en.html> [zuletzt geprüft: Januar 6, 2024].
- [80] Tate, J., und M. Atiyah. 2022. „The Millennium Price Problems“. [Online]. Verfügbar unter: <https://www.claymath.org/millennium-problems/> [zuletzt geprüft: Dezember 30, 2023].
- [81] Thielemann, H. 2004. „Klein, aber O“.
- [82] Tolnay, D., und andere. „Serde: Serialization framework for Rust“. [Online]. Verfügbar unter: <https://serde.rs/> [zuletzt geprüft: Januar 4, 2024].
- [83] Weitz, E. 2021. „Die Landau-Symbole“. *Konkrete Mathematik (nicht nur) für Informatiker: Mit vielen Grafiken und Algorithmen in Python*:479–494.
- [84] Wenning, W., W. Leinfellner, E. Kraemer, und J. Schänk. 1982. „Wittgensteins ‘Logik der Farbbegriffe’ und die Geometrie des Farbraums“. In *Language and Ontology. Proceedings of the 6th International Wittgenstein Symposium*.
- [85] Werner, A. G., und P. Syme. 1814. „Werners Nomenklatur der Farben: angepasst an Zoologie, Botanik, Chemie, Mineralogie, Anatomie und die Kunst“.
- [86] Zermelo, E. 1908. „Untersuchungen über die Grundlage der Mengenlehre“. In *Mathematische Annalen*. Leipzig: Springer, S. 261–281.