



Kippfigur

Eine optische Täuschung für das Ohr

Neunzig Sekunden, in denen sich kein einziger Ton ändert — und trotzdem alles

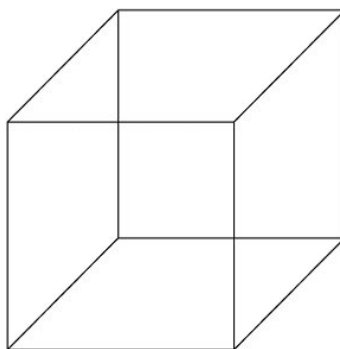
KI-MODIFIZIERT · AI-MODIFIED redaktionell geprüft von Olaf Engel

AUF EINEN BLICK

- **Eine Kippfigur** ist ein Bild mit zwei Lesarten, das sich nicht entscheidet: der Necker-Würfel, die Hase-Ente-Zeichnung. Nicht das Bild wechselt — der Betrachter wechselt.
- **Für das Ohr gibt es dasselbe**, an einer Stelle, die jeder kennt und kaum jemand bemerkt: bei der Frage, wo die Eins ist. Ein Takt ist keine Eigenschaft des Schalls, sondern eine Zuschreibung des Hörers.
- **„Kippfigur“ macht daraus eine Form.** Acht gleich laute Achtel, in allen 45 Takten wörtlich dieselben. Nur der Rahmen — Bass, Schlagwerk, Klavier — setzt seine Akzente um.
- **Dieselben acht Töne** heissen dadurch nacheinander 3+3+2, dann 2+2+2+2, dann wieder 3+3+2, aber ein Achtel später. Am Anfang und am Ende steht die Figur ohne Rahmen: unentschieden, wie der Würfel auf dem Papier.
- **45 Takte, 120 Schläge je Minute, 90,000 Sekunden, Seed 21.** Der Film ist nicht die Illustration des Stücks, sondern seine Bühne — er hält denselben Vertrag wie die Musik.

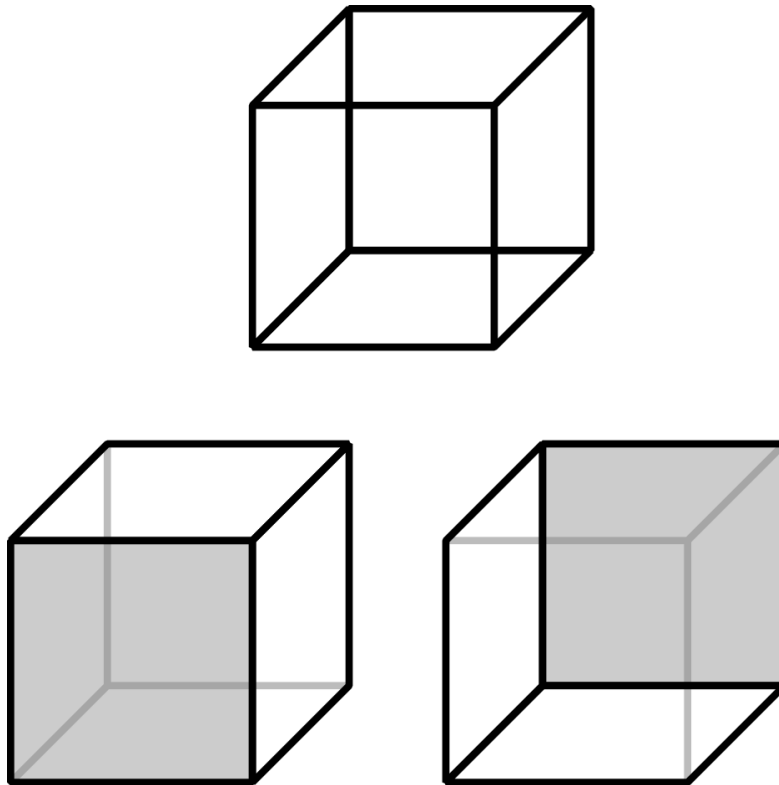
1 · Was eine Kippfigur ist

Es gibt Bilder, die zwei Antworten haben und sich für keine entscheiden. Das bekannteste ist ein Drahtwürfel, den der Schweizer Kristallograph **Louis Albert Necker** 1832 beschrieb: zwölf Striche auf einer Fläche, mehr nicht. Weil die Zeichnung keine Angabe darüber macht, welche der beiden Quadratflächen vorn liegt, muss das Auge sie ergänzen — und es ergänzt sie mal so, mal so.



Der Necker-Würfel. Schauen Sie ihn zwanzig Sekunden an, ohne etwas zu wollen: irgendwann springt er um. Sie haben nichts getan, und die Zeichnung erst recht nicht.

Was dabei passiert, lässt sich zeigen, indem man die eine und die andere Lesart einfärbt. Beide Male ist es **dieselbe Strichzeichnung**. Der Unterschied liegt nicht auf dem Papier, sondern in der Zuordnung, die der Betrachter mitbringt.

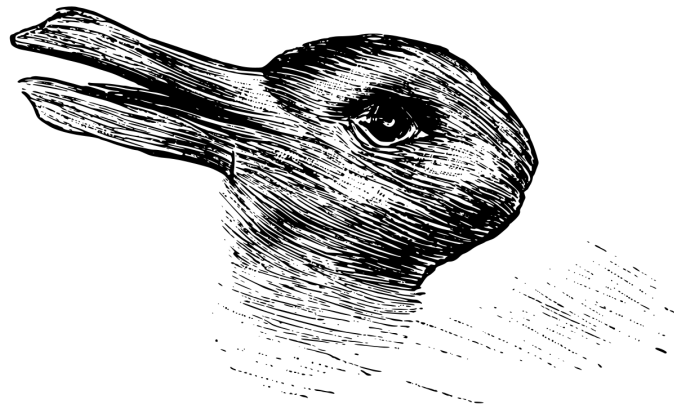


Oben die Zeichnung, wie sie dasteht. Unten die beiden Lesarten: einmal liegt die graue Fläche unten links vorn, einmal oben rechts. Kein Strich hat sich bewegt.

Drei Eigenschaften machen eine Kippfigur aus, und alle drei sind für das Folgende wichtig. **Erstens:** die Vorlage ändert sich nicht. **Zweitens:** es gibt zwei (oder mehr) vollständige, in sich stimmige Lesarten — keine ist der Fehler, keine ist die Wahrheit. **Drittens:** man kann sie nicht gleichzeitig haben. Das Umspringen kommt von selbst, und wer es erzwingen will, merkt schnell, dass er es allenfalls einladen kann.

Die zweite berühmte Kippfigur ist keine geometrische, sondern eine gegenständliche. Sie stand 1892 in den Münchner **Fliegenden Blättern**, mit der Bildunterschrift, die schon die Antwort gab. Der amerikanische Psychologe Joseph Jastrow machte sie um 1899 zum Versuchsobjekt, Ludwig Wittgenstein später zum philosophischen Beispiel für das „Sehen als“: Man sieht nicht erst Striche und deutet dann — man sieht sofort eine Ente oder sofort einen Hasen.

Welche Tiere gleichen ein- ander am meisten?



Kaninchen und Ente.

„Welche Tiere gleichen einander am meisten? Kaninchen und Ente.“ — Fliegende Blätter, München 1892. Was beim Hasen die Ohren sind, ist bei der Ente der Schnabel. Dieselbe Tinte, zwei Tiere.

Dazu kommt, weniger spektakulär, aber theoretisch am klarsten: die **Rubin-Vase** des dänischen Psychologen Edgar Rubin (1915), bei der dieselbe Kontur einmal als Gefäß und einmal als zwei Gesichter gelesen wird. Hier kippt nicht die Raumlage, sondern die Zuordnung von Figur und Grund: Was eben noch Umriss war, wird Hintergrund, und umgekehrt.

2 • Gibt es das auch für die Ohren?

Akustische Täuschungen sind bekannt, und die meisten funktionieren anders als eine Kippfigur. Der **Shepard-Ton** (Roger Shepard, 1964) scheint endlos zu steigen, obwohl er nach jeder Oktave wieder dort ist, wo er anfang — eine Schraube ohne Ende. Das **Tritonus-Paradox** von Diana Deutsch (1986) lässt ein Tonpaar für den einen Hörer steigen und für den anderen fallen. Beim **verbalen Transformationseffekt** (Richard Warren, Ende der 1950er Jahre) hört man ein in Schleife wiederholtes Wort nach kurzer Zeit als ein anderes, und dann als ein drittes.

Der Unterschied ist wichtig: Eine **Täuschung** sagt Ihnen etwas Falsches — der Ton steigt gar nicht. Eine **Kippfigur** sagt Ihnen zwei Dinge, die beide stimmen, und überlässt Ihnen die Wahl. Sie werden nicht getäuscht, Sie werden zuständig gemacht.

Von den genannten Fällen kommt der letzte der Sache am nächsten, und es gibt einen weiteren, der in der Hörforschung seit den 1970er Jahren als Standardfall gilt: das **auditorische Streaming**. Spielt man abwechselnd zwei Töne in der Folge hoch–tief–hoch, Pause, hoch–tief–hoch, so hört man je nach Tempo und Abstand entweder **eine** galoppierende Melodie oder **zwei** getrennte, parallel laufende Stimmen. Im Grenzbereich kippt die Wahrnehmung hin und her, ohne dass sich am Signal irgendetwas ändert. Albert Bregman hat diesen ganzen Bereich 1990 unter dem Namen **Auditory Scene Analysis** zusammengefasst: Das Ohr bekommt ein einziges Druckwellengemisch und muss daraus erst eine Szene bauen. Wo die Vorlage mehrdeutig ist, baut es mal so und mal so.

Für Musik interessanter als all das ist aber eine Mehrdeutigkeit, die niemand für eine Täuschung hält, weil sie so alltäglich ist: **die Frage, wo die Eins ist.**

3 · Die Eins steckt nicht im Klang

Ein Takt ist keine Eigenschaft des Schalls. In der Luft gibt es Anschläge und Abstände, sonst nichts. Dass acht gleich laute Anschläge sich zu „zwei Vierteln plus zwei Vierteln“ ordnen oder zu „drei plus drei plus zwei“, ist eine Leistung des Hörers, nicht des Lautsprechers. Solange kein Anschlag hervorsticht, ist die Reihe **metrisch unbestimmt** — genau wie die zwölf Striche des Würfels räumlich unbestimmt sind.

Musiker nutzen das seit Jahrhunderten. Die **Hemiole** setzt zwei Dreiergruppen gegen drei Zweiergruppen, ohne einen Ton zu ändern. Das lateinamerikanische **Tresillo**-Muster 3+3+2 teilt dieselben acht Achtel anders auf als der Marsch. In Leonard Bernsteins „America“ aus der *West Side Story* wechselt derselbe Takt zwischen der Sechser- und der Dreierlesart hin und her; das ist der ganze Witz der Nummer. Solche Stellen sind mehrdeutig — aber sie sind es meist nur nebenbei, für ein paar Takte, als Würze.

Die Frage, die dieses Stück stellt, ist eine andere: **Was passiert, wenn man daraus die ganze Form macht?** Nicht ein rhythmischer Trick an einer Stelle, sondern neunzig Sekunden, in denen die mehrdeutige Figur unangetastet bleibt und ausschliesslich ihre Umgebung wechselt — so, wie ein Museum denselben Necker-Würfel dreimal an die Wand hängt und nur die Beschriftung austauscht.

4 · Die Figur

Das Stück steht auf acht Achteln, die 0,25 Sekunden auseinander liegen: ein Takt von zwei Sekunden bei 120 Schlägen je Minute. Diese acht Anschläge sind in allen 45 Takten dieselben — dieselben Tonhöhen, dieselben Stellen, dieselbe Anschlagstärke, dieselbe Länge. Das ist keine Beschreibung, sondern eine Bedingung, gegen die die Datei geprüft wird: Wenn ein einziger Anschlag in einem einzigen Takt abweicht, ist das Stück kaputt.



Schaubild 1 — die Figur. Die Höhe eines Punktes ist seine Tonhöhe. Diese acht Anschläge stehen im ganzen Stück an derselben Stelle, in derselben Reihenfolge und mit derselben Anschlagstärke. Sie tragen keine Betonung — genau deshalb kann der Rahmen ihnen eine geben.

Entscheidend ist, was die Figur **nicht** hat: eine eigene Betonung. Alle acht Anschläge sind gleich laut. Wäre einer lauter, hätte die Figur selbst schon eine Eins, und der ganze Versuch wäre entschieden, bevor er anfängt. Der Necker-Würfel wird ja auch nur deshalb zur Kippfigur, weil der Zeichner die verdeckte Kante nicht gestrichelt hat.

Die drei Tonhöhen — D, Fis, A — hat der Würfel bestimmt, nicht ich: Das Haus arbeitet mit einem festen Startwert (hier: 21), aus dem alle Zufallsentscheidungen abgeleitet werden, damit dasselbe

Programm immer dasselbe Stück ergibt. Herausgekommen ist ein D-Dur-Dreiklang über einem gehaltenen A, das die Diva durchsingt. Auch das ist unentschieden: Man kann es als D-Dur mit dem A im Bass hören oder als ein schwebendes A mit einer grossen Sexte darüber. Der Würfel hat die Sache getroffen.

5 • Die drei Lesarten

Der Rahmen besteht aus Orgelbass, Schlagwerk und Klavier. Er fügt der Figur nichts hinzu — er benutzt nur ihre eigenen drei Töne und deren Oktaven — und er hat genau eine Aufgabe: zu bestimmen, **wo die Gruppen anfangen**. Drei Möglichkeiten kommen vor.



Schaubild 2 — die drei Lesarten. Oben die Figur, darunter dreimal dieselbe Figur mit den Klammern des Rahmens. Die goldenen Punkte sind die Gruppenanfänge: dort setzen Bass, Basstrommel und Klavier ihre Akzente. Es ist kein Ton dazugekommen, keiner verschwunden, keiner verschoben. Nur die Klammern sind umgesprungen.

In der **ersten Lesart** liegen die Akzente auf dem ersten, vierten und siebten Achtel: drei, drei, zwei. In der **zweiten** auf jedem zweiten Achtel: vier gleiche Paare. Die **dritte** ist die eigentliche Zumutung — sie ist derselbe Dreier-Dreier-Zweier wie die erste, aber um ein einziges Achtel nach hinten gerückt. Was eben noch der Auftakt war, ist jetzt die Eins; was die Eins war, hängt hinten über. Dieselbe Figur, um eine Sechzehntel Sekunde anders eingerahmt, und sie klingt wie eine andere Melodie.

Der Puls selbst bleibt die ganze Zeit gleich: Die Hi-Hat schlägt in jedem Formteil alle acht Achtel gleich laut durch. Sie ist das Metronom, das sich nicht einmischt. Nur Basstrommel, Bass und Klavierstücke ziehen um — und mit ihnen die Eins.

6 • Der Ablauf

Die neunzig Sekunden sind in sieben Teile geschnitten. Anfang und Schluss sind rahmenlos: die Figur allein, undeutbar, wie der Würfel auf dem weissen Blatt. Dazwischen bekommt jede Lesart

genau zehn Takte — acht am Stück und zwei im „Flackern“, wo jeder einzelne Takt anders gelesen wird und das Ohr nicht zur Ruhe kommt. Im vorletzten Teil laufen zwei Lesarten gleichzeitig: Der Bass liest Dreier-Dreier-Zweier, das Schlagwerk liest vier Paare, beide gleich laut. Dort muss der Hörer sich entscheiden, und er kann sich anders entscheiden als sein Nachbar.

GOLDENE STRICHE = UMSCHLAG (ein Blechruf, kein Übergang)

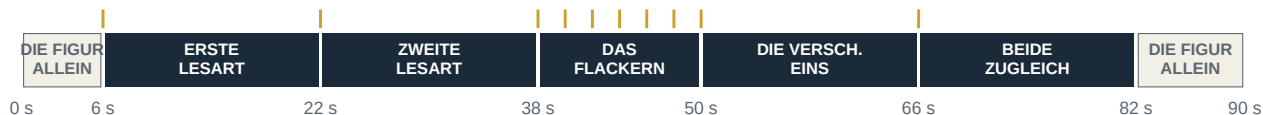


Schaubild 3 — der Ablauf. Die hellen Felder sind die Stellen ohne Rahmen: Anfang und Schluss. Dazwischen liegen die drei Lesarten, das Flackern, in dem jeder Takt anders gelesen wird, und der Teil, in dem zwei Lesarten gleichzeitig laufen.

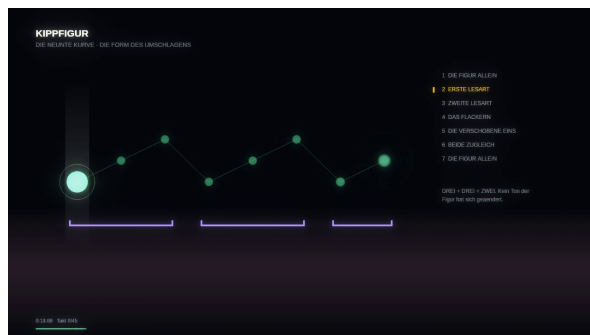
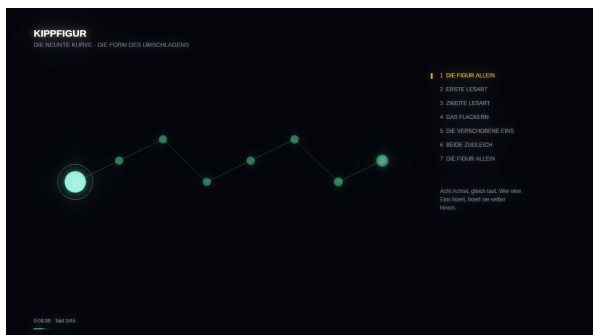
Jeder Umschlag ist ein **Schnitt, kein Übergang**. Kein Anschwellen, keine Überblendung: Der alte Rahmen hört mit dem Takt auf, der neue steht ab der Taktgrenze. Angesagt wird er durch genau einen Blechruf auf der Eins — sechs Hörner, ein einziger Akkord, und sonst nie im Stück. Das ist die akustische Entsprechung zu dem, was beim Würfel passiert: Er kippt nicht langsam, er ist plötzlich anders.

7 • Wie man es hört

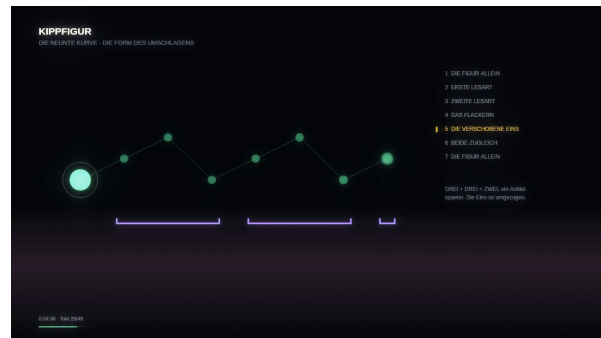
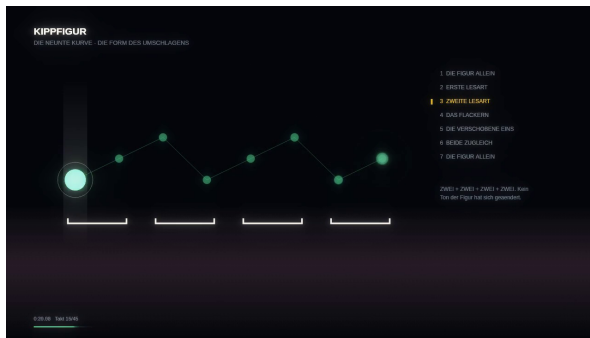
Für Laien ist das Stück leichter zugänglich, als es klingt, wenn man beim Hören eine einzige Sache tut: **mitklopfen**. Nicht jeden Anschlag — nur das, was sich als Eins anfühlt.

EIN KLEINER VERSUCH IN VIER SCHRITTEN

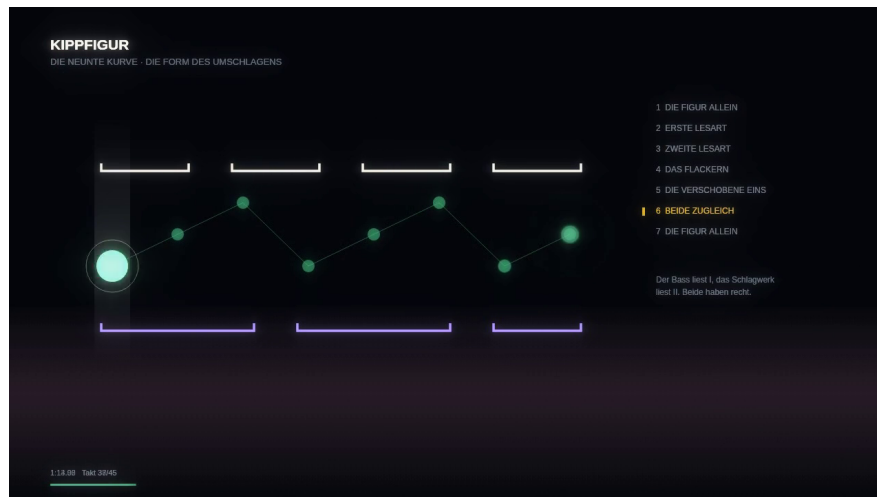
- **Sekunde 0 bis 6:** die Figur allein. Klopfen Sie mit. Wo Sie die Eins hinlegen, ist Ihre Entscheidung — das Stück gibt sie nicht vor.
- **Sekunde 6 bis 22:** der Rahmen kommt. Ihr Klopfen wird sich still nach ihm richten. Achten Sie darauf, dass die acht Töne der Melodie dabei genau dieselben geblieben sind.
- **Sekunde 22:** ein Hornstoss, und die Gruppierung ist eine andere. Dieselbe Melodie klingt jetzt gerader, marschierender. Sie hat sich nicht geändert.
- **Sekunde 66 bis 82:** beide Lesarten laufen gleichzeitig. Versuchen Sie, absichtlich von der einen auf die andere zu wechseln. Das ist genau die Anstrengung, die man beim Necker-Würfel kennt.



Links Sekunde 4: die Figur ohne Rahmen — acht Punkte, keine Klammern, nichts entschieden. Rechts Sekunde 14: die erste Lesart, drei violette Klammern unter denselben acht Punkten.



Links Sekunde 30: die zweite Lesart, vier Klammern. Rechts Sekunde 58: die verschobene Eins — die Klammern beginnen unter dem zweiten Punkt, die letzte ragt über den Takt hinaus.



Sekunde 74: beide zugleich. Oben die Klammern des Schlagwerks (vier Paare), unten die des Basses (drei, drei, zwei). Der Hörer entscheidet, welche er glaubt.

8 • Das Bild hält denselben Vertrag

Der Film ist keine Illustration und keine Animation, die zur Musik dazuerfunden wurde. Er ist die Bühne des Stücks: dieselbe Partitur, Bild für Bild abgespielt. Was man sieht, ist die Reihe der acht Punkte — ihre Höhe ist ihre Tonhöhe —, durch die das Licht in jedem Takt gleich wandert. Der Rahmen erscheint als **Klammern**, die im Umschlag ohne Übergang umspringen. In den beiden rahmenlosen Teilen gibt es keine Klammern; dort gruppiert das Auge selbst, genau wie das Ohr. Die Klammern sind damit nicht Dekoration, sondern das, was die Musik dem Hörer antut, in sichtbarer Form.

9 • Ein Befund aus der Werkstatt

Ein Punkt, der beim Bauen aufgefallen ist und der über dieses Stück hinaus gilt. Die Vertonung entstand in MuseScore mit den Klangbibliotheken „Muse Sounds“. Diese Bibliotheken nehmen ihre Lautstärke aus den **Notenzeichen** der Partitur, nicht aus den Anschlagstärken der MIDI-Datei. Ein importiertes MIDI hat keine Zeichen — also spielt alles gleich, nämlich mittellaut. Das Ergebnis war messbar: Die Figur, gespielt von einer Harfe, lag in ihrem eigenen Frequenzband (560 bis 920 Hertz) **25,6 dB unter dem Rest**. Sie war nicht leise, sie war weg. Ein Stück, dessen ganze Idee darin besteht, dass eine unveränderte Figur durch alles hindurch trägt, hatte die Figur verloren.

MASSNAHME	WIRKUNG
Regler 11 (Expression) des Rahmens von 84 auf 30	rund 4 dB
Mischpult-Fader: Harfe +12 dB	rund 8 dB
Mischpult-Fader: Schlagwerk –12, Klavier –8, Orgel –6 dB	rund 4 dB
Abstand Figur zum Rest, vorher / nachher	25,6 dB → 6,4 dB

Die Lehre daraus steht jetzt in den Hausnotizen: **Bei Muse Sounds ist die Balance Mischpult-Arbeit, nicht Anschlagstärke** — und gemessen wird im Frequenzband des betroffenen Instruments, nicht am Gesamtpegel. Am Gesamtpegel sah alles ordentlich aus, während die Hauptsache unhörbar war. Ein zweiter, kleinerer Fund: MuseScore öffnet eine geänderte MIDI-Datei unter demselben Namen aus seinem Zwischenspeicher — zweimal kam bytgleich die alte Vertonung heraus. Wer eine neue Fassung hören will, muss sie unter einem neuen Namen ablegen.

10 • Was daran neu ist und was nicht

Der Ehrlichkeit halber, weil die Frage berechtigt ist: **Das Phänomen ist nicht neu.** Metrische Mehrdeutigkeit ist ein alter Bestand der Musiktheorie, und dass Wahrnehmung zwischen zwei Deutungen kippen kann, ist in der Hörforschung seit Jahrzehnten beschrieben. Wer sagt, er habe die akustische Kippfigur erfunden, sagt die Unwahrheit.

Neu — jedenfalls in dieser Strenge selten — ist die **Bauweise**. Üblicherweise ist metrische Mehrdeutigkeit eine Stelle: ein paar Takte, in denen ein Komponist mit der Eins spielt, und die Melodie ändert sich dabei mit. Hier ist die Mehrdeutigkeit die Form selbst, und die mehrdeutige Figur ist buchstäblich unantastbar: Man kann in der Datei nachzählen, dass alle 45 Takte identisch sind. Damit wird aus einem musikalischen Effekt ein **kontrollierter Versuch** — dasselbe, was der Necker-Würfel für das Auge ist, mit dem Unterschied, dass ein Bild stillsteht und Musik durch die Zeit läuft. Der Hörer kann die beiden Lesarten nicht nebeneinander legen; er bekommt sie nacheinander und muss sich auf sein Gedächtnis verlassen. Das ist die eigentliche Schwierigkeit — und der Grund, warum das Stück am Anfang und am Ende dieselbe rahmenlose Figur zeigt.

Der Formpunkt in einem Satz: **Es ändert sich kein Ton, und trotzdem ändert sich alles.** Wer am Schluss dieselben acht Achtel hört wie am Anfang und dabei etwas anderes hört, hat den Beweis geführt, dass ein Teil der Musik nicht in der Musik steckt, sondern in ihm.

11 • Zahlen und Dateien

GRÖSSE	WERT
Länge	45 Takte à 2,0 s = 90,000 s; der letzte Ton endet auf 90,000
Tempo, Raster	120 Schläge/Minute, 4/4, acht Achtel je Takt (0,25 s)
Startwert (Seed)	21 — derselbe Seed ergibt immer dasselbe Stück
Die Figur	D5 Fis5 A5 · D5 Fis5 A5 · D5 Fis5, alle mit Stärke 0,78
Anschläge gesamt	1096, davon 360 auf die Figur (8 in jedem Takt)
Umschläge	10, davon 3 mit Becken; 30 Blechruf-Anschläge, sonst keine
Tonspur	93,1 s (90 s Form + Hall), Spitze –0,8 dBFS, Effektivwert –16,4 dBFS

Im Ordner liegen die Partitur als Programm, die Prüfung mit den acht Vertragspunkten und zehn Gegenproben, der MIDI-Bau, die Bühne als HTML, der Bildrechner und der Pegelabgleich; dazu die MIDI-Dateien in Haus- und Orchesterbesetzung, die Tonspur, die Bühne und der Film in zwei Auflösungen. Der Werkstand steht in **STAND_KIPFFIGUR.md**.

Belege und Vorlagen

Louis Albert Necker: Beobachtung zum Umspringen perspektivischer Zeichnungen, 1832 — daher der Name Necker-Würfel.

„Welche Thiere gleichen einander am meisten? Kaninchen und Ente.“ — Fliegende Blätter, München 1892; von Joseph Jastrow (um 1899) in die Psychologie und von Ludwig Wittgenstein (Philosophische Untersuchungen, 1953) in die Philosophie übernommen. Gemeinfrei.

Edgar Rubin: Figur und Grund („Rubin-Vase“), 1915.

Roger N. Shepard: der endlos steigende Ton, 1964. — Diana Deutsch: das Tritonus-Paradox, 1986. — Richard M. Warren: der verbale Transformationseffekt, ab Ende der 1950er Jahre.

Albert S. Bregman: Auditory Scene Analysis, 1990 — Standardwerk zum auditorischen Streaming und zu bistabilem Hören.

Die Abbildungen des Necker-Würfels und der Hase-Ente-Zeichnung sind gemeinfreie Vorlagen; die Standbilder stammen aus dem Film zum Stück.

Werkreihe „Die Band“ · Die neunte Kurve · 28. August 2026

Autorschaft: Konzept, Regie und redaktionelle Prüfung — Olaf Engel. Komposition, Programm, Bild und Text — Claude (Anthropic).