Fixing a Grid on What Has No Pulse: A Rhythm Codex for Field Material

Arnaud Quercy

August 29, 2026

Ideamorphic Studies Papers — codex

A recording of an environment keeps no time, and yet a rule can be made to write it as bars, beats and sixteenths. This paper states that rule completely enough to be re-executed without access to any code, and freezes it as the Rhythm Codex for Field Material, version 1.0. Two detectors are specified and chosen by material rather than by taste — spectral flux for impacts, syllable nuclei for speech — after which each bar keeps its own strongest four onsets, a selection iterated to a fixed point because a bar's length depends on the tempo the selection produces. The first and last kept onsets anchor the span, so choosing a number of bars is choosing the tempo, and the number chosen is the one that makes the eighth note the median value. Six rules tried and withdrawn are recorded with the measurement that ended each. The paper's own result is the grid error: the mean distance, in sixteenths, between a kept accent and the nearest grid position, which is zero if the material sits on the grid and a quarter if it sits at random. Eight of nine executions return between 0.221 and 0.244. What this paper computes is the distance between the events the rule detects and the grid it lays over them — and that distance sits at chance: the grid adds no order the detected events did not have. This establishes something the research notes do not: that measurability does not entail informativeness. A gap can be a genuine number and the number can report that the prescription is doing no work.

A deterministic rule that writes any recording as bars, beats and sixteenths — and a measure of how far the material stands from the grid it is written on.

Arnaud Quercy — Ideamorphic Studies
Ideamorphic Studies Papers, PUB-PAP0011 — Specifies the Rhythm Codex for Field Material, version 1.0

1. What This Paper Does

A fragment of sound can be made to yield three different things: a set of pitches, a rhythm, or a sequence of forms. Three rules, one material. They are specified separately, and this paper covers the second — the rule that takes a fragment and returns a notated rhythm.

The separation rests on the decision declared in Note 01: a mode is a ruled space, individuated by its rules and not by the sensory channel [1]. That decision is not an established result, and nothing here depends on the part of it that is scheduled to reopen.

The paper states the rule in full, names its frozen parameters and where each came from, records the rules tried and withdrawn, and declares the scope of the evidential protocol. It carries one result the companion specification [2] could not: the gap between the rule posed and the material it is posed on is a **measurable quantity here**, it has been measured, and its value says the grid is not in the recording. That is the paper's own finding, and §10 states what follows from it. The rule has been executed on nine scopes; that is not a corpus.

2. Why the Instrument Exists

There is no pulse in this material to find, and §5 gives the measurement that says so in the paper's own terms: across nine executions the accents fall almost exactly as far from the grid as they would if they ignored it. Beat-tracking has nothing to track. A café terrace does not keep time.

The rule fixes a time anyway, and its worth lies entirely in being reproducible.

What it is for is audible rather than argumentative. The signature has to be playable **over the original recording** — a count-in, then clicks on the accents, so that a listener standing in front of the work hears the starting point rather than being told about it. That requirement is what forces notation instead of a list of timestamps: a click track is a rendering, and a rendering needs a tempo, a meter and note values.

A recording has none of the three. The rule assigns all three. That is not a discovery about the recording; it is an institution, and it binds only the emitter [3].

3. The Rule

Stated completely, so that it can be re-executed without access to the code.

Input. One recording; optionally a time range within it; and a choice of detector. Everything below operates on that range and on nothing else.

Step 1 — Choose the detector. Two are available, and the choice is made by material rather than by taste. They are different measurements, not two settings of one.

- **Flux** measures how fast the spectrum changes: the sum, across the spectrum, of the positive part of the frame-to-frame difference. It is right for crockery, a chime, a door — material whose spectrum tips over in an instant.
- **Nuclei** measures band-limited energy instead: the sum of energy between 300 and 3000 Hz, in dB, smoothed to 50 ms. It is right for speech, and it aims at the middle of a vowel rather than at an attack.

Flux is the default. Nothing about the material makes it the natural first choice — it is the one that was built first, and it is fixed as the default so that the choice of detector is a declared decision rather than a silent one.

Step 2 — Find the onsets. Under flux, keep every local maximum of the flux curve that exceeds the local median plus 1.5 × MAD (median absolute deviation, scaled by 1.4826), the local window being ±0.5 s, with no two onsets closer than 50 ms. Under nuclei, keep every local maximum of the smoothed band energy that stands above the 35th percentile of the scope, with no two closer than 120 ms.

Step 3 — Select, per bar, iterated to a fixed point. Each bar keeps its own strongest four onsets. The bar is the unit being quantised, so it is also the unit of selection.

Three consequences must be stated, because they decide the counts a re-execution will obtain. First, the span runs from the first kept onset to the last, and *n* bar lengths laid along it open ***n* + 1** selection buckets — the last one holding whatever falls after the final bar line, which is the last onset itself. A scope of *n* bars therefore yields up to 4 × (*n* + 1) accents, and fewer wherever a bucket held fewer than four onsets. Second, the floor of eight events applies to the **provisional** pass below, never to the final set, which may end under it on a very short scope. Third, the ceiling of four hundred applies to the per-bar selection only.

The selection is circular — a bar's length depends on the tempo, which depends on the selection — so it is iterated: a provisional tempo is taken over all onsets, giving a bar length, which re-selects, which gives a new tempo. The iteration stops when the bar length settles within 10 ms, or after six passes.

Under the nuclei detector the per-bar quota is **not applied**, and neither is the ceiling: every nucleus is kept. A syllable is already a discrete event, not a peak selected out of a continuum; it is counted, not chosen.

Step 4 — Fix the anchors, and with them the tempo. The first kept onset is beat 1 of bar 1; the last is beat 1 of the final bar. The span between them is therefore fixed, and choosing a number of bars *is* choosing the tempo.

The number chosen is the one whose tempo makes **the eighth note the median value**: take the median gap between kept onsets, require that it be written as an eighth, and the beat is twice that. Among the bar counts whose resulting tempo falls between 30 and 220 BPM, the one nearest that target wins. Tempo and note values are one knob — halving the tempo just writes the same rhythm in finer values — so the rule fixes the knob in the middle of the range rather than at either end.

Step 5 — Quantise. Round each onset to the nearest sixteenth of the resulting grid, and read off its bar, beat and sixteenth. Where several onsets land on the same sixteenth, they are marked as an **ornament** on that sixteenth and the grid is not deepened.

Step 6 — Measure loudness, and read dynamics from it. For each kept accent, take the RMS of the signal in a ±40 ms window and express it in dB relative to the median accent of the scope. This is measured identically under both detectors, because the two do not report comparable strengths: flux returns a rate of spectral change, nuclei returns band energy. The detector says *where*; this says *how loud*. Dynamics follow directly: the median accent is `mf` by construction, and each 2 dB moves one rung along `ppp pp p mp mf f ff fff`. A mark is written only where it changes.

4. The Frozen Parameters

Twenty-two values fix the rule. Changing any of them changes the codex and must produce a new version rather than a silent rewriting of past results.

Parameter	Value	Why this value
Analysis rate	22 050 Hz	Onsets need time resolution, not bandwidth.
Meter	4/4	Stipulated, exactly like the 440 Hz of the pitch codex. The same span reads as 57 bars in 4/4 and 76 in 3/4.
Grid	sixteenths	The target, and the floor of readability.
Target median value	the eighth	Fixes the tempo knob at the middle of its range rather than at an end.
Tempo range	30–220 BPM	Outside it the notation stops being playable.
Accents kept per bar	4	Read off four test recordings as the density that made a thirty-second excerpt legible.
Event floor	8	Floors the provisional pass only, so a short excerpt still reaches a first tempo.
Event ceiling	400	Caps the per-bar selection, so a long flux scope does not become a wall. Not applied under nuclei.

Parameter	Value	Why this value
Fixed-point passes	6	The iteration settles in two or three; six is the give-up point.
Fixed-point tolerance	10 ms of bar length	Below this the selection no longer changes.
Flux — frame length	1024 samples	A 46 ms window: short, to place an onset in time.
Flux — hop	256 samples	An 11.6 ms step.
Flux — threshold	local median + 1.5 × MAD	Relative, because no absolute value of "marking" exists across materials.
Flux — local window	±0.5 s	What "local" means in the threshold above.
Flux — minimum gap	50 ms	Two onsets closer than this are one event.
Nuclei — band	300–3000 Hz	Where vowel energy lives.
Nuclei — smoothing	50 ms	The syllabic timescale.
Nuclei — minimum gap	120 ms	No language runs past about eight syllables per second.
Nuclei — voicing floor	35th percentile	Below it the frame is not voiced enough to carry a nucleus.
Nuclei — hop	128 samples	Half the flux hop: a nucleus is broad, but its centre must still be placed.
Loudness window	±40 ms	Short enough to belong to the accent, long enough to be an RMS.
Dynamic step	2 dB per rung	Eight rungs, `mf` at the median by construction.

The meter, the grid, the target value and the accent density are stipulations. They were not derived and they are not optimal in any sense the instrument can establish.

5. What the Instrument Records

For each execution: the version of the codex; the detector used; the tempo, the number of bars, the meter and the grid; the median gap and the target tempo it produced; the grid error; every kept accent with its time **on the file's own timeline**, its bar, beat and sixteenth, its strength relative to the median onset, its loudness in dB relative to the median accent, and whether it is part of an ornament; the number of ornament groups; the total onsets found and their density per second; and the twenty-two parameters.

Since 29 August 2026, every execution also carries two blocks that belong neither to the frozen parameters nor to the output. **The recording**: its path, the SHA-256 hash of the file and its duration — what tells two takes apart once a result has left its node, and what catches a long take that loops. **The execution**: the emitter's decisions on that material — the range analysed, as start and end **on the file's own timeline**, the transposition in semitones, the onset detector and the number of shape segments. None of these is a parameter of the codex and none is a result; without them a transposed or trimmed result could not be replayed. Executions made before that date carry neither block.

Reference values, nine scopes:

Scope	Detector	Tempo	Bars	Accents	Onset density	Grid error
A	flux	81.7	3	16	7.66 /s	0.225
B	flux	192.0	3	16	5.67 /s	0.236
C	flux	30.0	87	352	5.83 /s	0.236
D	flux	30.4	10	44	7.72 /s	0.240
E	flux	133.4	4	18	4.46 /s	0.224
F	flux	42.4	10	44	4.70 /s	0.232
G	flux	206.7	1	7	5.36 /s	0.149
H	nuclei	143.2	4	31	4.10 /s	0.221
I	nuclei	157.2	89	564	4.13 /s	0.244

The grid error is the quantity worth attending to, and §10 turns on it. It is the mean distance, in sixteenths, between a kept accent and the nearest grid position — 0 if the material sits exactly on the grid, 0.25 if it sits at random with respect to it, since the distance from a uniformly placed point to the nearest integer averages a quarter.

Eight of the nine scopes fall between 0.221 and 0.244. Scope G, at 0.149, holds seven accents in a single bar and is too small to weigh. Scope I shows the nuclei exemption of Step 3 at work, and not by exceeding a limit: no per-bar quota and no ceiling apply there, so the accents simply *are* the onsets. 89 bars at 157.2 BPM span 136 s, which at the reported 4.13 onsets per second is the 564 accents written — one for every nucleus found.

6. What the Instrument Does Not Record

No pulse. The rule does not claim to find one, and §5 gives the measurement that says there is none to find. The tempo it returns is assigned, not detected.

No meter. 4/4 is stipulated. Nothing in a recording chose it, and a different meter would return a different notation of the same material — 57 bars against 76 for one span already tested.

No absolute accent. Strength is reported relative to the median onset of its own scope, never on a scale that crosses recordings. Nothing in speech exceeds twice the median onset while a chime reaches 8.8 times; a fixed threshold therefore means one thing in one recording and another in the next.

No finer subdivision than the sixteenth. Coincident onsets are named as ornaments rather than resolved, so no execution can report thirty-seconds or sixty-fourths.

No verdict on aptness. A stipulated rule is neither true nor false [4]. What is available is the extent of its execution.

7. Rules Tried and Withdrawn

Six, each recorded with the measurement that ended it.

Ranking the whole scope at once. The obvious selection takes the strongest *n* onsets over the whole range. The loudest passage then swallows the quota: 15 of 40 accents landed in the first three seconds, and raising the density only thickened that same burst. The rest of the piece stayed silent. Selection had to become per-bar, because the bar is the unit being quantised.

A fixed strength threshold. If per-bar quotas distort, an absolute threshold should not. It fails for the reason given in §6: the same multiple of the median means a loud impact in one material and nothing at all in another. Withdrawn as unmeaning rather than as inaccurate.

A fixed number of events. Ten accents for any scope makes the tempo a property of where the recording was cut. Two framings of one recording returned 69.9 and 38.2 BPM. Under a per-bar density the same two framings returned 69.9 and 63.9 — the tempo became a property of the material instead.

Deepening the grid to resolve coincident onsets. Subdividing until every onset gets its own position let a single tight pair drag an entire piece down to sixty-fourths. Finer values are only meaningful on isolated events. Clusters are named, not subdivided.

Flux on speech. A syllable is not a transient: it is a vowel, a smooth swell of energy, and flux has no edge to find. Lowering the threshold to 1.0 × MAD only harvested consonant noise and within-syllable ripple — 39 "syllables" that landed nowhere near the syllables. The nuclei detector exists because of this, and on a French announcement it gives 4.16 syllables per second, confirmed by ear against the recording.

Capping nuclei per bar. Applying the per-bar quota to the nuclei detector dropped 31 syllables to 8. A syllable is not selected out of a continuum; it is already discrete. The quota belongs to flux alone.

8. Three Strata, and the Scope of the Protocol

The division into strata is Note 04's — acquired technique, stipulated codex, enacted idea of the work [5]. The first term is replaced here as in the companion specification, and the replacement is a debt to Note 05: *the given* is what a sensor measures at the surface **without the codex**, and *the instituted* is the rule that elects relations within that surface, which only the divulged codex deciphers [6].

The given — the detection curve of the file (spectral flux, or smoothed band energy) at the declared frame and hop, and the signal's RMS. Facts of the file.

The stipulated codex — the detector's thresholds and minimum gaps, the per-bar selection, meter, grid, target value, tempo range. It produces the onsets kept, their written positions and their dynamics. It produces no interpretation of that score.

The idea of this particular work — which detector suits this material, what the score is played against, what the practitioner's notes say the fragment is. Not deduced.

Here the strata come apart in a way they did not for pitch, and it is the paper's own result. Note 05 attaches a condition to any gap between a rule and its realization: it is a measurable magnitude only where the **arrival** material carries a metric [6]. The arrival material of this codex is a notated score, whose metric is the grid itself — but the grid error of §5 is not the prescribed/realized gap of Note 05. It is the distance between the events the rule detects and the grid it lays over them — and that distance sits at chance: the grid adds no order the detected events did not have.

Note 05's condition therefore does not apply here. What exists is another quantity, measured on nine scopes: between 0.221 and 0.244 against a chance value of 0.25. The material stands almost exactly as far from the grid as it would if it ignored the grid completely.

The evidential protocol reads the first stratum only. That this distance is near chance is a fact about the rule's output on the given; that the grid is worth imposing anyway is not a claim the protocol can settle, and is not made here.

9. The Chain of Gaps

- from the recording to the range chosen within it, and to the detector chosen for it — the two links entirely the practitioner's, and the two the rule takes as given;
- from the event in the room to the microphone;
- from the microphone to the encoded file, with its lossy compression;
- from the file to the decoded range;
- from the range to the onset curve, with its window and its hop;
- from the curve to the kept onsets, through a threshold that is local and relative;
- from the kept onsets to the tempo, through a median and a bar count;

- from the tempo to the written position, through the rounding to the nearest sixteenth.

The last link is **reported rather than inferred**: the grid error is what that rounding costs, averaged. The others are not decomposable with what exists.

10. What the Device Contributes, and the Question It Makes Askable

Note 05 establishes that the gap between a prescribed rule and a realized surface is a measurable magnitude only where the arrival material carries a metric: it holds for colour, which has chromatic coordinates, and not for the regressed word of the Feigned Tongue, whose criterion is an audible curve rather than a number. Measurability, the note concludes, is indexed to the codex and is not a property of the instituted invariant in general [6].

What is here is one step further along, and it is the paper's own. This is the first codex in the practice on which a distance has actually been **computed** — the distance between the events the rule detects and the grid it lays over them — and that distance sits at chance: the grid adds no order the detected events did not have. Note 05 says measurability is indexed to the codex. This says something the note does not: that **measurability does not entail informativeness**. A gap can be a genuine number and the number can report that the prescription is doing no work at all. A codex can be fully measurable and the measurement can be empty.

That distinction has a consequence for the device *The Prescribed and the Realized* specifies. It is built to record a gap; it is not built to say whether a recorded gap carries information. A corpus of executions carrying their grid errors is what would let a threshold be posed — below which a measured gap is indistinguishable from no rule at all.

Nine scopes cannot pose it. The instrument records the quantity so that a corpus can.

11. What This Paper Does Not Claim

The instrument writes a recording as a score under a stipulated rule. It establishes nothing about the recording's rhythmic content, nothing about how a work made from it is received, and nothing about whether the rule is apt.

It does not claim that the tempo it returns is present in the material. One measurement says otherwise, and it is stated twice rather than being two: the accents fall near-randomly against the grid the tempo implies — §5 reports it, and §2 rests on it. It is written here rather than left to be discovered.

It does not claim that near-chance grid error makes the rule worthless. A stipulated rule is neither true nor false; what the number reports is how much of the writing is the writing's own — and on this material, most of it.

The rule changed repeatedly while it was being built, until it stabilised in July 2026 for the works of the exhibition *Mémoire Urbaine et Collective* (Orangerie du Sénat, Paris, 28 August – 8 September 2026). This paper, published later, fixes it as version 1.0. Any change to the twenty-two values, or to Steps 1 to 6, produces a new version; a result stays attached to the version that produced it, and nothing already executed is rewritten.

REFERENCES

[1] Quercy, A. (2026). *The Passenger Without a Vehicle — Relational Content and Multiple Execution*. Research Notes, Series N, Note 01 (P2). Ideamorphic Studies. https://doi.org/10.5281/zenodo.22772473

[2] Quercy, A. (2026). *Folding a Recording onto a Fundamental: A Pitch Codex for Field Material*. Ideamorphic Studies Papers, PUB-PAP0010. https://doi.org/10.5281/zenodo.22774915

[3] Quercy, A. (2026). *Instituted Unity: A Rule Suffices to Bind What No Space Grounds*. Research Notes, Series N, Note 02 (P9). Ideamorphic Studies. https://doi.org/10.5281/zenodo.22769643

[4] Quercy, A. (2026). *The Prescribed and the Realized: A Measurement Device for a Chromatic Codex*. Ideamorphic Studies Papers, PUB-PAP0004.

[5] Quercy, A. (2026). *What Precedes Execution: Direction, Not Figure*. Research Notes, Series N, Note 04 (P1). Ideamorphic Studies. https://doi.org/10.5281/zenodo.22771429

[6] Quercy, A. (2026). *Two Invariants: The Given and the Instituted*. Research Notes, Series N, Note 05 (P3 + P3a). Ideamorphic Studies. https://doi.org/10.5281/zenodo.22773025

CITE THIS WORK

Quercy, A. (2026). Fixing a Grid on What Has No Pulse: A Rhythm Codex for Field Material. Art Quam Anima Publishing New York. https://doi.org/10.5281/zenodo.22775346

DOI (this version)	10.5281/zenodo.22775347
DOI (all versions)	10.5281/zenodo.22775346
Deposited	Version 1.0 · 2026-09-15
License	CC BY 4.0

ISSN: applied for — U.S. ISSN Center, Library of Congress

Published by Art Quam Anima Publishing New York,
an assumed name of AQA Publishing LLC

c/o Northwest Registered Agent, 418 Broadway Ste N
Albany, NY 12207, USA
+1 917-764-5470

NY DOS ID 7624376 · Assumed Name ID 7658709

publishing.artquamanima.com

Last updated: 2026-09-15

Persistent URI: https://publishing.artquamanima.com/en/papers/2026/08/fixing-a-grid-on-what-has-no-pulse-a-rhythm-codex-for-field-material-5e8m.pdf