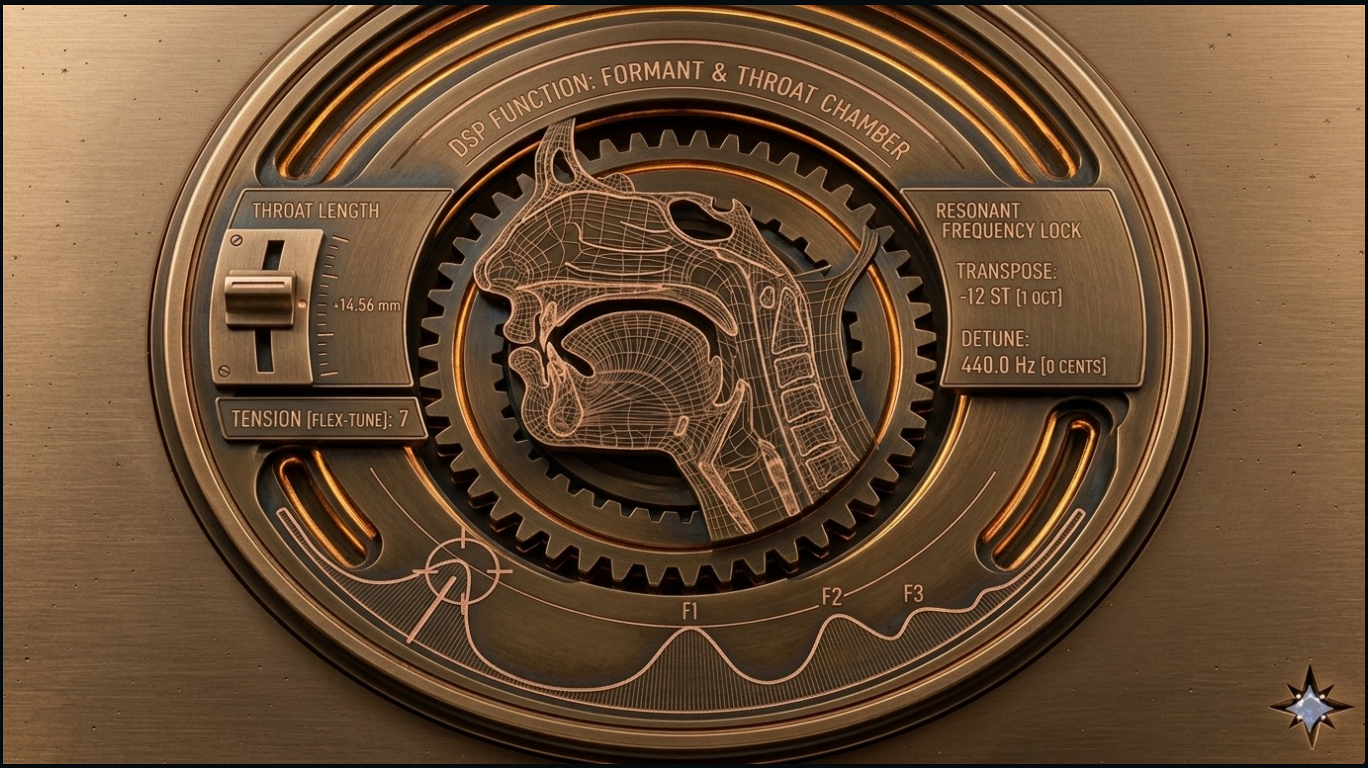
MARTINVILLE // MAGIC GEAR DIVISION

APPLETON, WI



WORM GEAR SPEED REDUCER

OPERATOR'S MANUAL

MODEL GRX-A1 · INDUSTRIAL AVIONICS SERIES

PITCH CORRECTION & FORMANT INSTRUMENT

MORTVIS NON SVNT MORTVIS

CONTENTS

Model GRX-A1 · Industrial Avionics series Martinville // Magic Gear Division

A pitch correction and formant instrument. The face is a worm-gear speed reducer because that is what pitch correction *is*: a high-speed input driven down through a fixed ratio into a slower, steadier output. Every control on the plate is a real parameter, not decoration.

Four of the things this instrument does are not available anywhere else in this form, and they are the reason it exists rather than being one more corrector. They are documented first, at length, in **Part I**. Everything needed to operate the plugin day to day is in **Part II**. Reference tables are in **Part III**.

The technology in Part I is proprietary and built in-house; see §0.3.

0. Contents

Where do I go for…

I WANT TO…	GO TO
Load a file and play it	**§6 The transport**
Pause — the PLAY key is also the PAUSE key	**§6.2**
Export a rendered WAV	§6.6
Read the top bar	§7 The status ribbon
Turn correction on and off	§8 HELIX
Make it snap harder or softer	§9 TENSION
Set the key it corrects to	§10 KEY and SCALE
Change speed without changing pitch	§11 DRIVE RATE, and §1
Change pitch without changing speed	§11.4 TRANSPOSE and MAGIC GEAR
Change the singer's throat length	§2 The throat chamber
Open the spectrum, the tuning grid, the EQ	**§12 The ADVANCED face**
Recall a voice	§13 VOCAL PRESET bank
Change the skin, read the manual, reset	**§16.1 The SYS menu**
Find out what a control does	Hover it — or §15.1
Something is wrong	§17 Troubleshooting

Front matter

Part I — The four engines

Sections in bold are new in this revision. §6, §12 and §16.1 document surfaces that shipped without documentation: the transport, the ADVANCED face and the system menu. Part I keeps its numbering; Part II is renumbered from §6 onward to put the transport where it belongs, which is first — you load a file before you touch anything else.

0.1 About Martinville

The name is not a founder

There is no Mr. Martinville at a desk in Appleton. **Martinville is a roll of names** — the engineers and instrument-makers who were right, were early, and were forgotten, absorbed without credit, or dismissed as cranks until long after they were dead. The division is named for them collectively. Edouard-Leon Scott de Martinville stands at the head of the list because his case is the clearest, not because he is the only one.

Edouard-Leon Scott de Martinville (1817-1879), a Paris printer, could not accept that photography had learned to fix light while nothing could fix sound. In **1857 he patented the phonautograph**: a horn onto a membrane, a membrane carrying a stylus, the stylus dragging through lamp soot on a hand-cranked cylinder. He built it out of anatomy — the membrane modelled on the eardrum, and when a brass horn proved too resonant he threw it out and sculpted a plaster conduit imitating the curves of the human ear canal. He was not trying to play sound back; that thought is nowhere in his patent. He was making sound **visible**, and expected people would learn to read a trace the way they read a printed line. He recorded *Au Clair de la Lune* in **1860**, seventeen years before Edison. Credit went elsewhere. He died insisting he had got there first.

Al-Jazari (1136-1206), engineer to the Artuqid court, wrote the *Book of Knowledge of Ingenious Mechanical Devices* in 1206 — fifty machines described well enough to build from. In it are the camshaft, the crankshaft coupled to a connecting rod, segmental gears, escapement mechanisms, and programmable automata whose sequence could be changed by moving pegs. He is a founder of mechanical engineering and of robotics by any honest reading, and sat outside the Western canon for the better part of eight centuries.

The Banu Musa brothers — Muhammad, Ahmad and al-Hasan — working at the House of Wisdom in ninth-century Baghdad, wrote the *Book of Ingenious Devices*, around a hundred mechanisms including automatic valves, self-trimming lamps, and devices that regulate themselves without attention. That is feedback control, described in the 800s.

Athanasius Kircher (1602-1680), the Jesuit polymath, wrote *Musurgia Universalis* in 1650 — the first substantial treatise on acoustics, covering echo, resonance, harmony, and the architecture of sound. He was among the most famous scholars in Europe in his lifetime and spent the following three centuries filed under crank. The reassessment is recent and ongoing.

Alexander Bain (1810-1877), a Scottish clockmaker, patented the electric clock and then, in **1843**, the **facsimile machine** — a working fax, three decades before the telephone. He lost his patent fights, was outmanoeuvred commercially, and died on a civil list pension in reduced circumstances.

Five people who built the thing first and did not get to keep it.

MORTVIS NON SVNT MORTVIS

That is the motto, and it is not decoration. In 2008, researchers at the First Sounds collective, working with Lawrence Berkeley National Laboratory, optically scanned Scott's soot traces and turned the lines back into sound. A hundred and forty-eight years after it was sung, into an instrument built by a man who did not believe playback was possible, a voice came back. The work outlives the worker. Set a thing down carefully enough and it is recoverable by people not yet born, whether or not its maker ever gets to hear it.

That is the standard this division holds itself to, and it is the reason this manual states its constants, formulas and thresholds exactly rather than gesturing at them. Documentation is the soot.

The mechanism

The gear is among the oldest machine elements we have, and the **worm gear** — a screw driving a toothed wheel — is attributed to **Archimedes of Syracuse** in the third century BCE. It survives unchanged in principle because it solves a hard problem in one stage: enormous speed reduction, and a drive that cannot be pushed backwards through. Set it and it holds.

That gearing was never primitive. The **Antikythera mechanism**, built around 150-100 BCE and pulled off the sea floor of a Mediterranean wreck, carries more than thirty bronze gears and predicts eclipses and the positions of the heavens. Nothing of comparable complexity is known again for something like fourteen hundred years — and the gap is not because the knowledge was disproved. It was mislaid. Al-Jazari and the Banu Musa are on the list above precisely because they are what filled that gap while Europe was not looking.

This is why the division is built on the gear, and why this instrument wears a speed reducer on its face rather than a picture of a waveform. Pitch correction *is* a reduction drive: a fast, unstable, high-energy input taken down through a known ratio into a slower, steadier, held output. That is not a metaphor applied after the fact. It is what the process does, and every name on the plate follows from it — lag, spool, backlash, wear, surface finish, alignment.

Why the two belong together

Scott worked in mechanical acoustics because it was the only medium available to him: a membrane, a stylus, a crank, and the shape of an ear cut into plaster. He made sound legible, and he modelled the organ of hearing to do it.

This division builds signal processing the same way — as machines whose working can be read off the face. The throat chamber of §2 models the organ of speech and reports it in millimetres. The alignment band of §1.11 reports the true state of the drive instead of decorating it. The GATE card of §8 shows the detector's confidence *and* its uncertainty rather than asking you to trust a lamp.

Scott made the invisible visible and was not believed. Everything here is built so that what it is doing can be seen.

0.2 The system, and its named parts

Nothing in this instrument is named after a marketing exercise, and none of these names will be familiar, because they are ours. This is the whole vocabulary in one table. Every one of them has its own section later.

NAME	WHAT IT IS	WHERE
HELIX	The master engage, and the three-state status lamp that reports it. Red bypass, green tracking, blue locked. Nothing downstream runs until HELIX is green.	§8
AURA	The pitch detector. Ours, and a two-state Kalman tracker — it carries pitch and the rate pitch is moving, which is why it follows a scoop instead of fighting it. It is what decides whether the lamp goes blue.	§4
GATE [HELIX LOCK]	The card that shows AURA's working — confidence, variance, frame counts, and which detector is driving. The lamp says what; this says why.	§8

NAME	WHAT IT IS	WHERE
TENSION	The one macro that decides how hard the plugin corrects. RELAX to TIGHTEN, 0 to 10. If you learn one control, learn this.	§9
MAGIC GEAR	The mode that decouples time from pitch. Ships OFF.	§1
CYCLOIDAL DRIVE ALIGNMENT	The live arc and rosette around the face that report the drive's state. Instrumentation, not ornament.	§1.11
VT [THROAT CHAMBER]	The card reporting vocal tract length in millimetres, and the measured formants.	§2
MECHANICAL_LAG	Retune speed in ms. Named for what it is mechanically: the lag before the drive catches up.	§15
INERTIAL_SPOOL	Pitch dead-zone in cents. The slack the drive tolerates before it acts.	§15
SURFACE_FINISH	The 6.5 kHz shelf. How polished the metal is.	§3
MIDRANGE_CUT	The 3 kHz bell. Fry and sibilance tamer.	§3
Alloys	BRASS, GOLD, PLATINUM — the pinion-hub selector. Each pins a wear profile and loads a voice.	§14

NAME	WHAT IT IS	WHERE
Wear profiles	HARD_LASH, SINE, SAWTOOTH — the *character* of the backlash, not its quantity.	§14
PATIENT ZERO	The preset that is essentially the untouched material. Your A/B reference.	§13

The through-line, if you want one sentence: the machine is a speed reducer, so everything is named for the mechanics of a speed reducer — lag, spool, backlash, wear, surface finish, alignment. The names are not decoration laid over DSP terms. They are the metaphor the instrument was designed inside, and each maps onto exactly one real parameter.

0.3 Proprietary technology

Everything in Part I was designed and written in-house for this instrument. None of it is a wrapper around a third-party engine, a licensed library, or a stock algorithm dropped in behind a new faceplate. What that buys you is an instrument whose readouts are measurements rather than decorations: the throat lever reads in millimetres because there is a real geometry behind it, and the formants beside it are measured off your voice, not inferred from the knob.

SYSTEM	WHAT IT IS	OURS
AURA	The pitch and voicing detector. It reports how strongly periodic the signal is and tracks where the pitch is heading, on its own dedicated thread. Ported from AURA-TONE™.	§4
Time-stretch	Changes duration without changing pitch, aligning each grain to the one before it so the waveform stays continuous. Written from scratch, no third-party code.	§1.4
MAGIC GEAR	The decoupling of the time and pitch axes, and the arithmetic that composes them without cross-talk.	§1.7

SYSTEM	WHAT IT IS	OURS
Throat chamber model	Formant shift expressed as vocal tract geometry, calibrated so that the zero position lands on real adult anatomy.	§2
Spectral bench	The real-time analyser and the conditioning stage, positioned so the corrector's own artefacts are visible and addressable in the box.	§3
Cycloidal drive alignment	The live instrumentation band — hypotrochoid rosette, alignment arc, travelling pip — driven from actual transport state.	§1.11

On AURA specifically. AURA predates this instrument. It is the detection stage of **AURA-TONE™**, developed in-house across three versions before Worm Gear was designed, and ported here onto its own thread. It is the reason the lock lamp means something: it reports a true measure of how periodic your voice is, rather than a number that can never reach its own threshold. See §4.

0.4 How it compares to what you already own

Most people reading this already have a pitch corrector. This section says plainly what Worm Gear does that yours does not, and — just as plainly — what it does not do that yours does.

Compared against Auto-Tune Pro 11 and Melodyne 5 Essential, verified 13 August 2026 against each vendor's own documentation and current store pricing. Both are excellent tools built by people who know exactly what they are doing. Feature sets and prices change; this chart is dated for that reason.

What it costs

	WORM GEAR	AUTO-TUNE PRO 11	MELODYNE 5 ESSENTIAL
Perpetual licence*	**$149 / $199 / $299**	$450*	$99*
Subscription option*	none offered	$210 / year, or $17.50 / month billed annually	none
Runs without a DAW	**yes, full standalone**	no	yes
Corrects in real time	**yes**	yes	no — offline editing

* **As of August 2026.** Competitor pricing is theirs to change and is not warranted here. Each figure is that vendor's own list price on the date this chart was verified.

The throat chamber

This is the reason the instrument exists, and it is the one thing here you cannot do anywhere else.

	WORM GEAR	AUTO-TUNE PRO 11	MELODYNE 5 ESSENTIAL
Formant shifting	yes	yes	no
Shown as a physical length in millimetres	**yes**	no	no
Live F1 / F2 / F3, measured, in hertz	**yes**	no	no
Measured against the physics that predicts them	**yes**	no	no
Spectral envelope of your own voice, live	**yes**	no	no

The difference is not that we have a formant control. Antares has shipped throat modelling for years and it is good. The difference is that this one tells you *what you have built*: the lever reads in millimetres of vocal tract, and the three resonances it produces are measured out of your actual voice and printed in hertz beside the quarter-wave figure that predicts them — see §2.

Every other corrector gives you an abstract slider from -10 to +10. This gives you a scale, a measurement, and the arithmetic connecting them.

Tuning

	WORM GEAR	AUTO-TUNE PRO 11	MELODYNE 5 ESSENTIAL
Chromatic, major, minor	yes	yes	yes
Modal and exotic scales	yes	yes	basic only
Alternate tunings — 5-EDO, 19-EDO, 31-EDO, Bohlen-Pierce	**yes**	no	no
Historical temperaments — Vallotti, Young, Werckmeister, the meantones	**yes, ten**	no	no
Per-note microtuning grid, ±100 ¢	**yes, all twelve**	no — its Detune is global	no
Per-note bypass and remove	**yes**	yes	no
Scala (.scl) import	**yes**	no	no

All four rows retune the corrector — they are not displays. Choose Werckmeister III and the note your voice is pulled to genuinely moves, by the cents that temperament specifies, per note name. Drag a single column and only that pitch class moves, in every octave.

A corrector that only knows equal temperament can only ever make a voice sound *modern*. See §12.2.

Time and pitch

	WORM GEAR	AUTO-TUNE PRO 11	MELODYNE 5 ESSENTIAL
Transpose	yes, ±12 st	yes	yes
Tape-style varispeed, 0.5×…2×	**yes**	no	no
Break pitch off tempo on a switch	**yes — MAGIC GEAR**	no	yes, offline
Per-note time editing	no	yes	yes

Worm Gear and Melodyne solve opposite halves of this problem. Melodyne moves a note after the fact, note by note, with all the precision that allows. Worm Gear moves the entire transport and then decides — live, while you listen — whether pitch rides along with it. One is an editor. This is an instrument. If you need to nudge the third syllable of bar 14, use Melodyne. See §11.

The working surface

	WORM GEAR	AUTO-TUNE PRO 11	MELODYNE 5 ESSENTIAL
Real-time spectrum analyser built in	**yes**	no	no
Conditioning EQ inside the corrector	**yes**	no	no
Import a file and work with no host at all	**yes**	no	yes
Render to disk from the plugin	**yes**, 44.1/48/96 k · 16/24/32-bit	no	yes
Factory voice bank	22, whole-instrument recall	presets	presets

What it does not do

Stated here rather than left for you to discover.

	WORM GEAR	AUTO-TUNE PRO 11	MELODYNE 5 ESSENTIAL
VST3	yes	yes	yes
CLAP	**yes**	no	no
AU	**not yet**	yes	yes
AAX	**not yet**	yes	yes
ARA2	**not yet — planned for v2**	yes	yes
macOS	**not yet**	yes	yes
Per-note graphical pitch editing	**no**	yes	yes
Polyphonic editing	**no**	no	no

Worm Gear is Windows today: VST3, CLAP, and a full standalone. If you are on a Mac, or you need ARA2 so the corrector sits inside the timeline, or you need to draw a pitch curve by hand, this is not the tool for that job yet, and no chart should pretend otherwise. **ARA2 is the headline of version 2.**

One line

Auto-Tune tunes a voice. Melodyne edits one. **Worm Gear lets you build the throat it came out of — and shows you the measurements.**

PART I

WHAT NOTHING ELSE DOES

1. The speed reducer: time and pitch, driven separately

1.1 What a speed reducer actually is

A worm gear speed reducer is a worm — a screw — driving a wheel. The worm turns fast and weakly. The wheel turns slow and strongly. The ratio between them is fixed by tooth count and cannot slip. It is the most honest mechanism in the machine shop: one shaft in, one shaft out, a known number between them, and no way to cheat it.

The face of this plugin is drawn as that mechanism and drawn honestly. The worm wheel carries **42 teeth**, the pinion carries **12**, the tooth module matches across both, and the pinion counter-rotates exactly **3.5 times** for every turn of the wheel. That is a genuine 3.5:1 reduction, not a decorative number painted on a graphic.

It matters because the whole premise of this section is that the gear train is the control surface. When you read **7:1** on the plate, you are geared down to half speed. When you read **1.75:1**, you are geared up to double. The mechanism and the parameter are the same statement.

PLATE READING	SPEED RATIO	WHAT THE AUDIO DOES
1.75:1	2.0×	Double speed
3.5:1	1.0×	Unity — the model's nominal reduction
7:1	0.5×	Half speed

The alignment arc around the face lights in proportion to lever position — half lit at unity, filling as you gear down — and the travelling pip runs at a rate proportional to the speed ratio, so the face is telling you the truth about the transport even when you are not looking at the numbers.

1.2 The problem this solves

Every mechanical playback device ever built couples speed to pitch. Tape run slow drops in pitch. A turntable at 33 instead of 45 drops in pitch. This is not a design flaw, it is arithmetic: if you read the same waveform out at a different rate, every period in it is stretched by the same factor, so every frequency in it scales by the inverse.

Software inherited the problem verbatim. A plain playback-rate control is exactly a tape machine — one lever, both outputs welded together. Set it to half speed and you get an octave down with it, whether you wanted the octave or not.

That coupling is the single most common reason a good idea dies in a session. You want the phrase slower so you can hear it. You get it slower and an octave down, and now you cannot hear it either. You want the beat a semitone flatter to sit with a sample. You get it a semitone flatter and four percent slower, and now it fights the grid.

Auto-Tune cannot do this. Melodyne can, élastique can, Waves SoundShifter can — but they are offline processes or tab-based editors. Not one of them is a lever you drive in real time on the face of the instrument. Doing it as a gearbox you actually operate is unique to this plugin.

1.3 The two levers

Turn **MAGIC GEAR** on. Then:

LEVER	LOCATION	RANGE	AXIS IT DRIVES	WHAT STAYS PUT
SPEED RATIO	Steel lever, centre face	0.5×…2.0×	Time. Tempo, duration, groove.	Pitch. Every note stays where it was.
TRANSPOSE	Knob	±12 semitones	Pitch.	Time. Duration and tempo do not move.

They are independent. You may drive either alone or both together, and they compose cleanly — half speed and up a fourth is a legitimate, single, coherent setting, not a compromise between two fighting processes.

MAGIC GEAR is a toggle, and it ships OFF. The button sits by the transport and reads **MAGIC GEAR · OFF** until you click it. **With it off, the steel lever is ordinary varispeed** — speed and pitch coupled, exactly like tape. That is deliberate; varispeed is a sound people want on purpose. But it means that if you drive the lever expecting decoupling and never armed the toggle, you will hear pitch dragging along with speed and reasonably conclude the feature does not work. Arm it first. The button lights when it is live.

1.4 How it works: SOLA

The decoupling is done by **SOLA** — Synchronous Overlap-Add — written in-house for this instrument. No third-party library is shipped inside the plugin doing work nobody can inspect.

The idea is old and sound. To make audio longer without making it lower, you do not stretch the waveform — stretching the waveform is exactly the thing that lowers it. Instead you cut the signal into short frames, and you lay those frames back down with *more space between them* than you took them with. The frames themselves are untouched, so the periods inside them are untouched, so the pitch is untouched. Only the rate at which new material arrives has changed.

To make audio shorter, you lay the frames down with less space between them than you took them with, and some material is skipped. Same principle, opposite sign.

TERM	MEANING HERE
Analysis hop	How far you advance through the *input* between frames.
Synthesis hop	How far you advance through the *output* between frames.
Stretch factor S	Synthesis hop divided by analysis hop. S above 1 is longer and slower; S below 1 is shorter and faster.

1.5 Why the "S" in SOLA is the whole trick

If you simply overlap-add frames at the new spacing, you get a specific and very recognisable failure: a hollow, flanged, chorused quality, worst on sustained vowels and held notes. It is the sound of naive time-stretch and everyone has heard it.

The cause is phase. Two overlapping frames of a periodic signal only reinforce each other if their waveforms line up. Move one of them by half a period and the crests of one sit in the troughs of the other, and the overlap region partially cancels. Because the misalignment drifts continuously as you advance through the file, the cancellation sweeps through the spectrum. That sweep is the flanging you hear.

Synchronous overlap-add fixes it by refusing to place a frame blindly. Before laying each frame down, the engine runs a bounded cross-correlation search: it slides the incoming frame back and forth by up to **±48 samples** and measures, at each offset, how well the frame's leading half correlates with the tail of the frame already written. It then commits to the offset with the highest correlation — the one where the waveforms genuinely align.

That is the entire difference between a time-stretch that sounds like the source and one that sounds like a broken chorus pedal.

1.6 What it costs, and where you will hear the limit

The stretch is not free, and the manual would be lying if it said otherwise. Two places you can hear it:

Low voices at extreme stretch. The engine aligns each frame to the one before it by searching a window about a millisecond wide. At high female and child pitches a millisecond is a whole pitch period or more, so the alignment is always findable. At a low male fundamental a period is nearer ten milliseconds, so the search has proportionally less room. Push a deep voice to the far end of the lever and that is the first thing to give.

Percussive material. The analysis frame is about twenty milliseconds. That is chosen to hold several periods of a low voice — but a transient shorter than the frame is being asked to sit inside one, and hard consonants soften slightly at large stretch factors. This is a vocal instrument, and the frame is sized for voices.

What it does not cost you: level. There is no compensation gain anywhere in the stretch path — nothing normalising, nothing riding level. So nothing pumps, nothing breathes, and the start and end of the buffer need no fade. The stretch is transparent in level by construction rather than by correction.

1.7 Two levers, two outputs, no cross-talk

This is the part that makes both levers work at once.

The **speed lever** sets duration. The **TRANSPOSE** control sets pitch. With **MAGIC GEAR** engaged they do not interact:

YOU MOVE	WHAT CHANGES	WHAT DOES NOT
Speed lever	Duration	Pitch stays where it was
TRANSPOSE	Pitch	Duration stays where it was

That independence is the whole point of the section, and it is not how playback normally behaves — a tape machine, a turntable and a plain playback-rate control all weld the two together. Here they are separate axes, and §11.4 covers driving them.

With MAGIC GEAR disengaged the two are deliberately re-coupled, because that tape behaviour is musically useful and sometimes exactly what you want.

1.8 Worked settings

GOAL	MAGIC GEAR	LEVER	TRANSPOSE	RESULT
Learn a fast line by ear	ON	0.5× (7:1)	0	Half speed, pitch dead still.
Tighten a dragging take	ON	1.08×	0	Eight percent faster, pitch untouched.
Drop a sample to fit a track	ON	1.0×	-1	A semitone down, duration and groove identical.
Male read toward a lighter voice	ON	1.0×	+3	Up a minor third, tempo intact. Pair with §2.
Half speed AND up a fourth	ON	0.5×	+5	Both axes at once. They do not fight.
Deliberate tape varispeed	**OFF**	0.5×	—	Coupled. Half speed, octave down. The old sound, on purpose.

1.9 Why it renders offline, and what that feels like

The correlation search is the expensive part. Cost scales as frame size times frame count times search range, and on a three-minute mono 48 kHz buffer that lands around **one to three seconds** on current hardware.

That is far too slow to run per-audio-callback, and pretending otherwise would mean dropouts. So the engine does not pretend. **The stretch renders offline, once, on knob release** — not continuously while you drag. Expect a brief pause when you let go of the lever, then clean playback at the new setting. It is a commit, not a scrub.

The upside of accepting that constraint honestly is that the render can afford to be good. It is not racing a deadline, so it uses the full correlation search on every frame rather than a cheap approximation.

1.10 Known limitation, current build

Pitching **up** uses a resample that is not yet band-limited. Content above the new Nyquist folds back into the audible band as aliasing, heard as a metallic grit on bright material. Aggressive up-shifts show it most. Sustained tones may also smear slightly under large stretch factors, which is the SOLA search range of §1.6 reaching its limit. > Down-shifts are clean. Moderate up-shifts are clean. Both issues are known and both are addressed by the phase-locked spectral engine that replaces this path. Until then: if an up-shift grits, shift down from a higher source instead.

1.11 Reading the face: the cycloidal drive alignment display

The band of instrumentation around the face is labelled **CYCLOIDAL DRIVE ALIGNMENT**, and every part of it is live. None of it is a texture. If you learn to read it you can run the transport without looking at a number.

Why a cycloidal drive at all, next to a worm gear. They are the two mechanisms in the reduction-gear world that solve the same problem from opposite ends. A worm drive gets its ratio from a screw meshing a wheel — enormous reduction in one stage, and it cannot be back-driven. A cycloidal drive gets its ratio from a lobed disc rolling inside a pin ring, driven by an eccentric — also enormous reduction in one stage, but with very low backlash and high shock tolerance, because load is shared across many contact points at once rather than concentrated on one tooth pair.

Naming the two axes after the two mechanisms is not arbitrary. The worm gear is the *correction* path: fixed ratio, non-backdrivable, it holds what you set it to. The cycloidal drive is the *transport* path: continuously variable, low lash, tolerant of being shoved around while it runs. That is exactly how the two halves of this plugin behave.

The rosette. The dense curve filling the disc is a genuine hypotrochoid — the curve a point traces when one circle rolls inside another, which is the generating curve of a cycloidal drive's disc profile. It is drawn from the parametric form

- x = (1 - Rr) cos(th) + d cos(k · th)
- y = (1 - Rr) sin(th) - d sin(k · th)

where **Rr** is the rolling ratio of the inner circle to the outer, **d** the tracing offset, and **k = (1 - Rr)/Rr** — the relation that closes the curve on itself rather than leaving it open. Choose Rr and d badly and you get a spirograph doodle; the proportions here are tuned by eye against real cycloidal disc profiles until the lobe count and the density read as a machined part.

It is traced in two passes — a wide soft pass for the glow, a fine crisp pass over it — and the hue drifts along the curve rather than sitting flat, so the disc reads as lit rather than printed.

And it turns with the drive. The rosette's phase advances at a rate proportional to the speed ratio. Gear down and it slows; drive fast and it races. It is a tachometer you read peripherally.

The alignment arc. The single most useful thing on the face.

ELEMENT	WHAT IT REPORTS
Dim rail	The full travel of the lever. The part you have not used.
Lit length	Lever position. **Half-lit is unity, 3.5:1.** Filling as you gear down.
Travelling pip	Runs the rail at a rate proportional to the speed ratio, wrapping continuously. It visibly races when you drive fast and crawls when you gear down.
Tick ring	46 ticks, every fourth one major, giving you a coarse scale to read position against.

The lit length and the pip speed are two different readings of the same drive. Length is *where the lever is*; pip rate is *how fast the transport is actually moving*. On a healthy transport they agree. The pip is also the fastest way to confirm playback is genuinely running when the material is quiet.

The numeric readouts sit with the lever:

READOUT	MEANING
RATIO n:1	The live gear ratio, computed as 3.5 divided by the speed ratio. Green. This is the mechanical statement.
VARISPEED ×n	The same setting as a plain multiplier. Dim. This is the audio statement.
Tick labels	1.75:1 at the top (2×), **3.5:1** at centre in green (unity), 7:1 at the bottom (0.5×).

Both numbers describe one lever. The green one is the machine's language, the dim one is the DAW's. They can never disagree.

> **The label says VARISPEED whether or not MAGIC GEAR is armed.** It is reporting the lever, not the mode. Check the MAGIC GEAR button for which engine the lever is actually driving — see §1.3.

2. The throat chamber: formant as physical geometry

2.1 The thing nobody else gives you

Every corrector on the market has a formant control. It is invariably a knob labelled *formant*, running from some negative number to some positive number, with no unit, no reference, and no statement of what it is physically doing. You turn it until it sounds right. There is nothing else you can do with it, because you have not been told what it is.

This instrument gives you the same shift expressed as **the length of the resonant chamber that would produce it, in millimetres, against a stated nominal.** The **VT [THROAT CHAMBER]** card reads out chamber length, the nominal it is measured against, the live measured formant peaks **F1 / F2 / F3** in hertz, and the **F2/F1** ratio.

That converts an aesthetic knob into an instrument reading. You are no longer turning something until it sounds right; you are setting a vocal tract to a length and watching the resonances land where the geometry says they must.

2.2 The physics it is built on

A vocal tract, to a first approximation that is good enough to be useful and standard in the literature, is a uniform tube closed at the glottis and open at the lips. A tube with those boundary conditions is a quarter-wave resonator. Its resonances fall at

$$F(n) = (2n - 1) \times c / (4L)$$

where **c** is the speed of sound, about 343 m/s in warm moist air, **L** is the tube length, and **n** counts the resonances 1, 2, 3. Those resonances are the formants.

Two consequences follow immediately, and they are the two facts this whole chapter rests on.

First: formant frequency is inversely proportional to tract length. Halve the tube, double every formant. This is why a child's voice sits high in a way that has nothing to do with the pitch they are singing — the chamber is short.

Second: the nominal is not arbitrary. Put **L = 175 mm** into the formula:

FORMANT	CALCULATION	RESULT
F1	343 ÷ (4 × 0.175)	490 Hz
F2	3 × 343 ÷ (4 × 0.175)	1470 Hz
F3	5 × 343 ÷ (4 × 0.175)	2450 Hz

Those are the textbook neutral-schwa formants of an adult male vocal tract — the values that appear in every phonetics text as roughly 500, 1500 and 2500 Hz. The 175.00 mm printed on the card is the length that puts the instrument's zero exactly on the neutral adult male vowel. Zero on this control is a real physical position, not an arbitrary centre detent.

2.3 The conversion, exactly

The chamber length for a shift of **t** semitones is

L(t) = 175 × 2^(-t/12) millimetres

and the **THROAT delta** readout shows the difference from nominal, which is 175 × (2^(-t/12) - 1).

The elegance is that the two readouts are the same statement in different units. Because formant frequency goes as 1/L, scaling the length by 2^(-t/12) scales every formant by 2^(+t/12), which is by definition a shift of t semitones. The millimetre readout and the semitone readout can never disagree, because they are algebraically the same number wearing different clothes.

FORMANT	CHAMBER LENGTH	THROAT DELTA	F1 LANDS AT	CHARACTER
-12 st	350.00 mm	+175.00 mm	245 Hz	Impossible anatomy. Cavernous, sub-human, an effect.
-6 st	247.487 mm	+72.49 mm	346 Hz	Very large chest. Heavy, chesty, oversized.
-3 st	208.17 mm	+33.17 mm	412 Hz	Large adult male. Broadens and darkens a thin read.
-1.4 st	190.00 mm	+15.00 mm	451 Hz	Big-framed adult male.
0 st	175.00 mm	0.00 mm	490 Hz	Nominal adult male. Neutral.
+3 st	147.10 mm	-27.90 mm	583 Hz	Adult female range.
+5 st	131.02 mm	-43.98 mm	654 Hz	Small adult, older child.
+8 st	110.19 mm	-64.81 mm	778 Hz	Child.
+12 st	87.50 mm	-87.50 mm	980 Hz	Half the nominal chamber. Every formant exactly an octave up.

2.4 Using it as an instrument rather than a knob

Because the axis is calibrated to real anatomy, you can work from a target instead of from trial and error. Typical adult male tracts run about 170 to 180 mm, adult female about 140 to 150 mm, and children roughly 100 to 120 mm. Read the target off the card and set the shift that lands on it.

JOB	SET FORMANT TO	REASONING
Male read toward a female tract	**+3 st**	175 mm to about 147 mm, which is inside the adult female range.
Female read toward a male tract	**-3 st**	The same move in reverse, 147 to 175 mm.
Undo chipmunking after a big up-transpose	**negative, about -0.7× the transpose**	A transpose up drags formants up with it. Push the chamber back out to hold the body where it was.
Undo the giant effect after a big down-transpose	**positive, same logic**	Same principle, opposite sign.

JOB	SET FORMANT TO	REASONING
Thicken a thin lead without touching pitch	**-1 to -2 st**	15 to 30 mm of extra chamber. Adds body, keeps the note.
Lift a dull background vocal	**+1 to +2 st**	Shortens the chamber, brightens without EQ and without moving pitch.

That fourth row is the pairing that makes Part I cohere. §1 moves pitch without moving time. §2 moves the body of the voice without moving pitch. Together they are three genuinely independent axes on one voice — **tempo, pitch, and physical size** — and you can set each without disturbing the other two.

2.5 Reading F1, F2 and F3

The three formant readouts are **measured, not predicted.** The engine takes the spectral envelope, picks its peaks, sorts them by frequency, and reports the lowest three in hertz. They are live. When they read **—**, no confident peak was found — usually silence, breath, or an unvoiced consonant, all of which is normal and not a fault.

They tell you two different things.

F1 tells you the chamber is doing what you asked. Set +12 and F1 should move toward roughly double where it sat. If it does not move, look at HELIX before looking at anything else — see §2.7.

F2/F1 tells you the vowel. The ratio of the second formant to the first is the classic coordinate of vowel space, and it is on the card because it is the number that stays meaningful while you move the chamber.

F2/F1	VOWEL REGION	SOUND
around 1.5 to 2	Open back	as in *father*, *thought*
around 3	Neutral	the schwa — the tube's own resting ratio
around 4 to 6	Mid front	as in *bed*, *bait*
around 7 to 10	Close front	as in *see*, *beet*

The useful property: shifting the chamber scales F1 and F2 by the same factor, so **the ratio is invariant under formant shift.** If you move the chamber and F2/F1 stays put, you have changed the size of the speaker and left the vowel alone, which is what you wanted. If F2/F1 moves, the vowel itself changed, and on a sustained note that means something else in the chain is interfering.

F3 - F1 is also displayed as a spread. It broadly tracks how much of the voice's brightness and presence is spread across the upper resonances.

2.6 The anatomical figure

The chamber artwork on the formant face is a single universal head rendered across every skin. This is intentional. The figure is the physical claim the control is making, so it is present on every finish rather than appearing on some and not others. It is composited in *lighten* so its dark background falls away over the gear beneath it, and where the artwork is present it *replaces* the code-drawn gear rather than stacking on top of it.

2.7 The one gate you must know about

FORMANT requires HELIX engaged (green). It does not require lock (blue).

The master engage gates the formant stage internally in the DSP. With the plugin bypassed, the formant control does nothing whatsoever — not a small amount, nothing.

This was confirmed rather than assumed: rendering the engine offline with the plugin bypassed, formant at 0 and formant at +12 produce **bit-identical** output files. There is no subtle residual effect to listen for. If HELIX is red, that knob is disconnected.

Green is sufficient. Waiting for blue is a common and pointless delay — see §8 for the difference between engaged and locked.

3. The spectral bench: an analyser and an EQ inside the corrector

3.1 Why this is here at all

A pitch corrector is supposed to hand you back a corrected signal and stop there. If the result is harsh, that is your problem, in your channel strip, with your EQ, after the fact.

That division of labour is tidy and it is wrong, for one specific reason: **correction changes the spectrum.** Snapping pitch resamples and re-phases material, and the audible cost lands in two narrow, entirely predictable places — a hardness around 3 kHz where fry and sibilance live, and a brittle glassy edge above about 6.5 kHz where the resampler's own character shows. Those artefacts are made *here*, by this process, and the information needed to hear them is *here* too.

So the conditioning is in the box, and so is the instrument for seeing it. The Advanced view carries a real-time FFT spectrum analyser and a two-band conditioning EQ, wired so that the analyser sits **after** the EQ. You are not adjusting blind and checking somewhere else. You move the control and watch the spectrum change.

3.2 The signal chain, in order

The output conditioning chain runs:

Correction engine -> 3 kHz peaking cut -> 6.5 kHz high shelf -> limiter -> analyser -> output

Two things follow from that ordering and both matter.

The analyser is last but one. It is fed from after the EQ and after the limiter, so the trace on screen is the signal that is actually leaving the plugin — not the raw corrector output, and not a tap from somewhere convenient upstream. What you see is what you get.

The limiter is inside the loop, always. It cannot be switched out. It is there as the safety net against clip-cracks, and because it sits before the analyser you can see it working rather than discovering it later.

3.3 The two bands, and why exactly these two

CONTROL	FILTER	CENTRE	Q	RANGE	THE PROBLEM IT SOLVES
MIDRANGE_CUT	Peaking	3 kHz	1.4	0 to -12 dB	Vocal fry and sibilant hardness in the 2–5 kHz band. A Q of 1.4 is roughly a one-octave bell, wide enough to take the whole hardness region down together rather than notching one squeak and leaving its neighbours.

CONTROL	FILTER	CENTRE	Q	RANGE	THE PROBLEM IT SOLVES
SURFACE_FINISH	High shelf	6.5 kHz	—	0 to -10 dB	The metallic ring and glassy polish above the shelf point. A shelf rather than a bell because the problem is a whole region tilting bright, not a resonance.

Neither is a general-purpose EQ and neither pretends to be. There is no frequency control on either, because the frequencies are not arbitrary — they are where this specific process puts its artefacts. Naming them **MIDRANGE_CUT** and **SURFACE_FINISH** rather than "3 kHz" and "6.5 kHz" is the honest labelling: they are jobs, not bands.

These are cuts only. Both run from 0 dB down. Neither will add anything, and that is deliberate — this is a conditioning stage for taking off what correction added, not a tone stack for building a sound. Boost belongs in your channel strip, after the plugin, where you have a real EQ.

3.4 The analyser

OPTION	VALUES	NOTES
FFT size	Selectable, 2048 by default	The resolution readout beside it reports the resulting **Hz per bin** live, so the trade is stated rather than implied.
Smoothing	Selectable	The analyser's time constant. Higher is calmer and slower; lower is twitchier and more truthful about transients.
Frequency axis	Log or linear	Log runs from 20 Hz to Nyquist and is the musical view. Linear spreads the top octaves out and is the view for hunting aliasing.
Floor	-90 dB	Bottom of the displayed window.
Peak hold	On / off	Holds the falling peak envelope per column. Turning it on clears whatever was held.
Freeze	On / off	Pauses the trace in place for comparison.
Waveform overlay	On / off	Superimposes the time-domain trace over the spectrum.

Hz per bin is the number to watch. At 2048 points and 48 kHz you get about 23 Hz per bin, which is coarser than the spacing between the low harmonics of a bass voice — fine for judging broad tilt, too coarse for resolving individual partials down low. Go up in FFT size when you need to see low-frequency detail and accept the slower response; go down when you are chasing something fast.

3.5 What to actually do with it

Freeze is the feature. The workflow that justifies the whole section is: set PATIENT ZERO from the preset bank (§13) to hear the material essentially untouched, freeze the trace, then switch to your working preset and compare the live trace against the frozen one. The difference between those two curves is precisely what correction did to the spectrum, isolated, on screen. That is the measurement no external analyser can easily give you, because it would need the before and after in the same window at the same gain.

Then condition against what you saw.

WHAT THE TRACE SHOWS	MOVE
A lump growing around 2–5 kHz after correction	MIDRANGE_CUT up until the lump returns to where PATIENT ZERO had it.
The whole top end tilting up and glassy	SURFACE_FINISH up. Shelf, so it takes the region down as a region.
Hash appearing above the source's own top end during an up-transpose	That is the aliasing of §1.10, not a tone problem. EQ will not fix it — reduce the up-shift. Switch the axis to linear to see it clearly.
Formant peaks not moving when you move the chamber	HELIX is bypassed. See §2.7.

And use it to read the throat. The spectral envelope on the formant face is the same analysis that drives the F1/F2/F3 readouts of §2.5. When you set a chamber length and want to confirm the resonances landed where the geometry says, the scope is the confirmation.

3.6 What it is not

To be precise, because precision is the point of this manual: the EQ is a pair of **biquad filters**, not an FFT-domain process. The FFT is the analyser — the instrument you measure with. Nothing here does spectral surgery, resynthesis, or per-bin gain.

The unusual thing is not the filter topology, which is ordinary and good. It is that a pitch corrector ships the measurement instrument and the conditioning stage together, positioned so that the artefacts the corrector itself creates are visible and addressable without leaving the plugin. That is the part nobody else does.

4. AURA — the detector

4.1 Where AURA comes from

Everything in the previous three sections assumes the plugin knows what pitch is being sung. That measurement is the hardest part of the whole instrument, and it is the part we already owned.

AURA is the detection stage of **AURA-TONE™**, our own pitch-correction technology, developed and revised across three versions. It predates this instrument. **Worm Gear is its first commercial iteration** — the engine came first, and the gearbox was built around it.

It earns its place here because a lock lamp is only as honest as the number behind it. A detector that cannot report certainty on a clean vocal is not an indicator, it is a decoration — and lowering the threshold to make it light up would be worse than leaving it dark.

4.2 What AURA is

AURA is our own pitch detector, running on its own dedicated thread so that analysis never competes with playback. It answers two questions:

MEASURE	QUESTION IT ANSWERS	WHAT IT MEANS
Voicing	*Is this a voice, or is it noise?*	How strongly periodic the waveform is, on a true 0-to-1 scale. On a clean vocal it goes where it should — near 1.
Certainty	*Do I actually know the pitch?*	How far the tracker's estimate has converged. High certainty means it has settled and stopped arguing with itself.

Those two measures drive the HELIX lamp, and between them they are the difference between an indicator and an ornament. **The lamp reports a measurement that can reach its own thresholds.**

Detection runs on its own thread rather than sharing one with playback and drawing. Periodicity analysis is expensive, and a detector that competes with the render loop for time is how you get crackle on the output — the failure it would be introduced to prevent.

4.3 What it is good at, and where it stops

AURA tracks **60 Hz to 1000 Hz** and reports nothing outside it. That is the singing range, and confining it there is what lets the detector run continuously without competing with playback.

It follows a voice rather than resisting it. A scoop into a note, a slide between two, a wide vibrato — these are movement, not error, and the tracker is built to recognise the difference between a voice going somewhere and a single bad frame. That is why correction here does not fight portamento, and why the lock survives ornament that makes a simpler detector let go.

It updates about **45 times a second**, which is fast enough that the readout moves with the voice rather than trailing it.

4.4 What it will not do

- **A pure test tone is not a voice.** Sine waves are perfectly periodic, so they read as pitched, but the voicing logic wants vocal material. This is the source of the "it ignores my test tone" report in §17. Test with a voice.
- **Breath and consonants drop the lock, by design.** Sibilants are not periodic. Watching HOLD reset during an s is the system working.
- **Silence reports silence.** Below the noise floor it does not guess.

4.5 What this means in practice

- **Blue means something.** It is a real statement about your voice, not a threshold quietly lowered until the lamp cooperated.
- **Sine tones do not trip it.** A pure tone is perfectly periodic, but the voicing logic wants vocal material. This is the source of the "it ignores my test tone" report in §17. Test with a voice.
- **Breath and consonants drop the lock by design.** Seeing HOLD reset during sibilants is the system working, not failing.
- **MIDI overrides the ear.** When SOURCE reads MIDI, the instrument is being told the pitch rather than measuring it, and the lock rows describe what it was told.

PART II

OPERATING

5. Quick start

1. **IMPORT** a WAV, AIFF or MP3, or arm **MIC**.
2. Click **HELIX** to engage. The lamp goes green (tracking), then blue (locked).
3. Set **KEY** and **SCALE**, or let the engine detect them on import.
4. Set **TENSION** — the one knob that decides how hard the plugin corrects.
5. Play. Adjust. **EXPORT WAV** when the take is right.

If you touch nothing else, TENSION and the VOCAL PRESET bank will cover most work.

For the speed reducer and the throat chamber, arm them explicitly: **MAGIC GEAR** must be toggled ON (§1.3), and **HELIX** must be engaged for FORMANT to do anything at all (§2.7).

5.1 What to feed it

Worm Gear corrects one voice at a time. That is not a limitation of this instrument in particular — it is what pitch correction *is*, and every corrector on the market says the same thing in its own manual.

Works:

- A single vocal track, reasonably well isolated.
- A single instrument playing one note at a time — a bass, a lead line, a monophonic synth.
- A stacked harmony **on separate tracks**, one instance per track.

Does not work:

- Two or more voices on the same track, including a double.
- A chord, or any instrument playing more than one pitch at once.
- A full mix.
- Heavy bleed — a vocal with the backing track audible in the same file.

The reason is not arbitrary. AURA finds a pitch by measuring the **period** of the waveform — how long it takes to repeat itself (§4). Two singers produce two periods at once and there is no single number that describes both, so the detector reports whichever is momentarily stronger and flips between them. The shifter then moves the whole track by whatever that reading said, which is wrong for both voices at every instant.

A wet vocal counts as bleed. Reverb and long delay are copies of the voice arriving late and at a different pitch relationship; they blur the period the same way a second singer does. Correct first, then add the space. If you must correct a wet take, expect the tracking to be less certain during tails.

Breath, noise and consonants

Breathy singing and noisy recordings can produce tracking errors. This is normal, it is not a fault, and **the instrument tells you when it is happening** rather than leaving you to guess — which is what the GATE card on the left is for:

READING	MEANS
PITCHED — YES	AURA has found a period. Correction is being applied.
PITCHED — NO	No usable pitch here. The audio is passed through untouched.
VOICE	AURA's confidence, 0 to 1. Sustained values below about 0.3 on notes you *are* singing means the detector is struggling.
STATE — TRACKING / LOCKED / BYPASS	Whether the gate is open, held, or off.

Consonants and breath are supposed to read as unpitched. They have no pitch to be wrong about, so the instrument leaves them alone deliberately — correcting a hard "t" would only buzz it. If PITCHED flickers off between words, that is the gate working, not failing.

If VOICE stays low across a *sung* phrase, the input is the problem, and in order of how often it is the culprit:

1. **Bleed or reverb** in the source. Fix the source.
2. **The take is very quiet.** Bring the level up before the plugin.
3. **The voice is genuinely breathy.** Some of this is unavoidable, and BRASS or a low TENSION will behave better on it than a hard tune — see §9.

6. The transport

The strip along the bottom of the instrument. It loads a take, plays it, and writes the result back out to disk. Everything on this row concerns the FILE — nothing here touches the correction.

You do not need it. Worm Gear is a plugin: put it on a track and it corrects whatever the host sends it. The transport exists so you can work on a take *without* a host — audition a preset, set the throat chamber by ear, check a drive ratio — and so that a standalone copy of the instrument is a complete tool rather than a demo.

6.1 IMPORT and MIC

IMPORT opens an audio file. WAV, AIFF and MP3 are accepted, and the format is read from the file's own header rather than from its extension — a .wav that is really an AIFF loads correctly, and a file with no extension at all loads correctly.

Importing does three things at once, before a note plays:

1. Decodes the file and resamples it to the session rate.
2. Measures the whole file's drive rate — see §11.
3. Measures its key, for HELIX — see §10.

The status line reports LOADED · <name> · m:ss immediately, and the two measurements overwrite it a moment later as they land. That order is deliberate: you learn the file was accepted before you learn what it contains.

MIC is live input. **It is not wired in this build.** There is no capture path yet, so the key is drawn and does nothing. It is left in place rather than hidden because it is coming, and a control that disappears between versions is worse than one that waits.

6.2 The six transport keys

They stay dim until a file is loaded — that dimming is the instrument telling you there is nothing to play, not that the key is broken. You can still hover any of them to read what they do.

KEY	WHAT IT DOES
■	Playhead to 0:00.
■	Back **five seconds**.
■ / ■	**Play, and pause.** One key, both jobs.
■	Stop, and return the playhead to 0:00.
■	On **five seconds**.
■	Playhead to the end.

THERE IS NO SEPARATE PAUSE KEY, AND THIS IS THE ONE THING IN THE TRANSPORT WORTH READING TWICE. The play key is the pause key. Press it while the file is running and playback pauses *where it is*; press it again and it resumes from that spot. The glyph changes from ■ to ■ so the key always shows what it will do next, and it lights while the transport is running.

Pause and stop are different, and the difference is your place in the take.

- **Pause** holds position. Play resumes from there.
- **Stop** returns to 0:00. Play starts from the top.

If you want to hear one phrase over and over while you move the throat lever, pause is the control you want — stop will send you back to the beginning every time.

The five-second steps are five seconds at any file length. They are not a percentage; a step means the same thing on a thirty-second loop as on a four-minute take, which is what makes them usable for finding a spot.

6.3 LOOP

Repeats the file instead of stopping at the end. **It lights while it is armed**, so its state is readable without playing to the end to find out.

Loop is how the instrument is meant to be set up. Arm it, start the take, and work the controls while it runs — the correction, the throat chamber and the drive ratio all respond as you move them, and nothing needs to be re-triggered.

6.4 The scrub strip

The waveform under the keys. Click anywhere on it to move the playhead there.

It is not a plain waveform: each column carries the minimum, the maximum **and** the RMS of that slice, and the colour rides the energy — cyan where the take is quiet through to red where it is loudest. You are looking at where the song gets hot, so the loud passage you want to test a setting against can be found by eye rather than by scrubbing for it.

6.5 The time readout

`0:00.0 / 0:00.0` — position, then total length.

The total changes with the drive ratio, because the drive ratio changes how long the file takes to play. At 1.75:1 a four-minute take runs two minutes; at 7:1 it runs eight. The number on the right is the length of what you are about to hear, not the length of the file on disk. See §11.

To the right of the clock is the take's own label — the file name, its sample rate and its measured length.

6.6 EXPORT WAV, and the two format selectors

EXPORT WAV renders the loaded take through the entire instrument at the current settings and writes it to disk. Not a capture of what is playing — a fresh render, from the top of the file, at whatever the controls say right now.

It stays dim until there is a file to render.

Pressing it asks where to save, then renders. A long take takes a moment; the status line reads `RENDERING …` while it works and `EXPORTED · 48 kHz 24-bit` when it lands.

What comes out is what you have been listening to. The render runs the same chain in the same order as live playback — drive ratio, AURA, the corrector, the throat chamber, MIDRANGE_CUT, SURFACE_FINISH, MIX and GAIN. If the preview sounds right, the file sounds right.

Two details that matter when you line the file up in a DAW:

- **It starts where the take starts.** The pitch shifter's own delay is removed from the head of the render, so the exported file is in time with the original rather than 53 milliseconds late.

- **It keeps the last note.** The shifter is still holding audio when the file runs out, and that audio is the release of the final phrase. The render keeps going until it has flushed.

SR — the exported file's sample rate

44.1 kHz · 48 kHz · 96 kHz.

How many times per second the file measures the sound. 44.1 kHz is what a CD uses and is right for almost everything. 48 kHz is the video standard, and is what most DAWs default to. 96 kHz is double 48, and is worth it only if something downstream is going to do heavy processing on the result.

This is the EXPORT's rate, not the session's. The rate the instrument is currently running at is shown in the ribbon at the top, and the two are independent — you can run a 48 kHz session and export at 44.1.

BITS — the exported file's bit depth

16-bit · 24-bit · 32-bit float.

How finely each of those measurements is recorded — in practice, the distance between the loudest and quietest sound the file can hold. 16-bit is CD. 24-bit is the studio standard and is the right default for anything you intend to mix further.

32-bit float is the one worth understanding, because it behaves differently in kind rather than in degree: it is the only setting that **cannot clip**. If the render peaks above full scale, 16- and 24-bit will flatten the tops of those peaks permanently, and 32-bit float will store them exactly as they were so they can be pulled back down later. If you have pushed GAIN, export in float.

7. The status ribbon (top bar)

Everything across the top. Left to right: what the instrument is set to, what it is, and how it is running.

7.1 The left group — what it is set to

The diamond, far left. Click it to step to the next skin; **right-click** it for the whole list. It is also the only always-visible sample of the current accent colour, which is why it is a diamond rather than a swatch — the same mark leads every menu row and both face tabs, and it means *a thing you can select*.

Eight finishes, darkest to lightest: DARK · GANGSTA · VENOM · ROSE PANTERA · TANGERINE DREAM · AURUM · ICE · PINSTRIPES. Cosmetic only — no audio effect — and your choice is remembered between sessions. §16.3.

CELL	WHAT IT IS
VOCAL PRESET	The factory bank. 22 voices, each a **whole-instrument** recall — not five knobs. §13.
KEY	The tonic the corrector snaps toward. Twelve chromatic roots. §10.
SCALE	Which degrees a note is allowed to land on. Fourteen scales. §10.
TRACKING	**A readout, not a control.** Retune speed in milliseconds. It reflects MECHANICAL_LAG, which TENSION drives — so this is the number that tells you how hard the instrument is currently tuning. It looks like the three selectors beside it and is the one that does not set anything.

7.2 The centre group — the two engage lamps

HELIX — the master engage, and the auto-key detector. Click to bypass, and double-click to detect the key of the loaded take. The lamp is the power lamp: dark is bypassed, green is tracking, blue is locked. Full detail in §8.

AURA-TONE™ — the badge for the pitch engine. Not a control. It lights when AURA has a confident lock, so at a glance it answers "is the detector actually following me" without opening the GATE card. §4.

7.3 The right group — how it is running

CELL	WHAT IT IS
FFT	The analyser's transform size. Set on the ADVANCED face. §12.1.
SR	**The session's** sample rate — whatever the host or the audio device is running at. Not the export rate; that is set separately at the bottom of the transport. §6.6.
LAT	The plugin's latency, in milliseconds — 2112 samples, so 44.0 ms at 48 kHz and 47.9 ms at 44.1. Reported to the host, which compensates for it automatically. It is not a delay you have to correct. §18.
MAT	The current alloy — BRASS, GOLD or PLATINUM. §14.3.
SIMPLE / ADVANCED	The face toggle. SIMPLE is the instrument you play; ADVANCED is the one you inspect. §12.
SYS	The system menu — skins, this manual, tooltips, restore defaults, and the build version. §16.1.

LAT is worth understanding once. Every pitch shifter has latency, because it cannot know what a waveform is doing until it has seen a period of it. Worm Gear reports its figure to the host so the host lines the track back up; the number is there to be *informative*, not actionable. If your DAW has delay compensation switched on — and it does by default — you do not need to do anything with it.

8. HELIX — the lamp, and what it is actually telling you

HELIX is the fleet-standard status lamp. It has three states:

STATE	COLOUR	MEANING
BYPASS	Red	Engine off. Dry signal passes through untouched.
TRACKING	Green	Engaged and hunting. Correction is active; pitch not yet confidently held.
LOCKED	Blue	Pitch is genuinely held. This is the working state.

Engaged is not the same as locked. Correction is live the moment the lamp is green. Blue only adds the statement "the detector is certain." Some controls — notably FORMANT, §2.7 — require *engaged*, not *locked*.

The GATE card — why the lamp is doing what it's doing

The lamp only says red/green/blue. The **GATE [HELIX LOCK]** card on the left says *why*, live:

ROW	MEANING
STATE	BYPASS / TRACKING / LOCKED.
VOICE	Smoothed voicing confidence (EMA). Compared against the thresholds.
PITCHED	Whether the detector currently has a real pitch, not just energy.
HOLD	Frames held toward lock, out of the 4 required.
SOURCE	Where the pitch is coming from: **AURA** when it is being heard, **MIDI** when it is being told. The frequency sits beside it, or `--` when there is nothing.
KALMAN P	Detector variance. Lower is more certain.

The lock rule: VOICE must hold above the LOCK mark for four consecutive frames to go blue, and only falls back to green after fifteen frames below the lower RELEASE mark. That gap is deliberate — it stops the lamp strobing on vibrato and breath. Tracker variance must also have settled.

You do not have to remember any of those numbers. The meter on the card draws the LOCK and RELEASE marks in place, so you can see how close you are rather than guessing.

When SOURCE reads MIDI, the instrument is being told the pitch rather than measuring it, and these rows describe what it was told.

Pure sine tones do not trip the detector. This is expected, not a fault — the voicing detector wants vocal material. Test with a real voice.

9. TENSION — the main control

RELAX <-> TIGHTEN. One knob, 0–10. If the plugin is fighting you, this is the fix. If you learn one control on this instrument, learn this one.

9.1 What it actually controls

TENSION is a **macro**. It does not correct anything itself — it drives the two parameters that decide how the correction behaves, together, along the curve that matters musically.

Those two are:

MECHANICAL_LAG — *how long the correction takes to arrive.* When you sing a note that is 40 cents flat, the corrector does not jump to the right pitch instantly; it glides there. This is that glide, in milliseconds. At 1 ms the note is simply *at* the target and the transition between notes is a step. At 250 ms the pitch is still travelling when the next syllable starts.

INERTIAL_SPOOL — *how wrong a note has to be before anything happens at all.* A dead zone, in cents. At 50 ¢ the corrector ignores everything inside a quarter-tone of a scale degree — which is most expressive singing. At 0 ¢ nothing is ignored: every deviation, including deliberate ones, is pulled.

Together they answer two different questions, and that is why one knob drives both. Tolerance decides *whether* to correct. Lag decides *how fast*. Tight means "correct almost everything, almost instantly." Relaxed means "correct only what is genuinely wrong, and take your time about it."

9.2 The whole curve

Every position, with the values it produces. You never need this table to use the knob — it is here so that when you write a number down, you can get back to it.

TENSION	MECHANICAL_LAG	INERTIAL_SPOOL	WHAT IT DOES TO A VOICE
0	250 ms	50 ¢	Barely intervenes. Scoops, slides and vibrato all survive intact. Catches only a genuinely missed note, and even then slowly.
1	144 ms	45 ¢	Repair, not production. You will not hear it working.
2	83 ms	40 ¢	Gentle. Holds a wandering line without flattening its shape.

TENSION	MECHANICAL_LAG	INERTIAL_SPOOL	WHAT IT DOES TO A VOICE
3	48 ms	35 ¢	The top of "natural". Vibrato still reads as vibrato.
4	27 ms	30 ¢	Audibly corrected on sustained notes, still organic on runs.
5	16 ms	25 ¢	Modern corrected-but-human. The pop centre of gravity.
6	9 ms	20 ¢	Clearly produced. The correction is part of the sound now.
7	5 ms	15 ¢	**The boot value.** Glossy and nailed, still recognisably a voice.
8	3 ms	10 ¢	Hard tune. Consonants start to click.
9	2 ms	5 ¢	Almost the full effect.
10	1 ms	0 ¢	The effect, not the repair. Every note stepped, nothing forgiven.

Retune follows an exponential curve, not a linear one, and the table shows why: the top half of the knob spends its whole travel between 16 ms and 1 ms, because that is where the audible change lives. The difference between 1 ms and 20 ms is the difference between a robot and a singer. The difference between 200 ms and 250 ms is nothing at all. A linear knob would waste half its rotation on distinctions no one can hear.

9.3 Which end of the knob you actually want

IF THE TAKE IS…	AND YOU WANT…	START AT
A good singer, slightly under on long notes	nobody to know you touched it	**1–2**
A good singer, a couple of genuinely wrong notes	the wrong notes fixed, everything else untouched	**0–1**
A strong performance you want to sound produced	modern radio vocal	**4–5**
A rap or sung-rap line	present and locked without sounding processed	**5–6**
Anything where the tuning IS the sound	the artefact, on purpose	**9–10**
A doubled or stacked harmony	the stack to agree with itself	**6–7**

9.4 When it is fighting you

Three complaints, and TENSION is the answer to all three.

"It sounds like it is chewing the vibrato." Tolerance is too tight — the corrector is treating the vibrato's own excursion as an error and pulling against it several times a second. **Come down.** Below about 4 the dead zone is wide enough that a normal vibrato lives inside it and is left alone.

"Every note arrives late" / "it smears between words." Lag is too long for the tempo. On a fast line the correction is still travelling when the next syllable begins. **Go up**, or set MECHANICAL_LAG directly to something under 10 ms and leave the rest alone.

"It sounds lazy." Both, at once, and this is the one people misdiagnose as a broken plugin. A wide dead zone plus a slow glide means most notes are never corrected and the ones that are arrive late. If a preset ever sounds like it is not doing anything, read MECHANICAL_LAG and INERTIAL_SPOOL before you touch anything else — see §9.6.

9.5 MESH FIT — retune as six named fits

Retune speed is a choice between **characters**, not a continuous sweep. Nobody has ever wanted 137 ms. Every voice in the preset bank clusters around a handful of values, so those clusters get names and a button each.

FIT	RETUNE	CHARACTER
PRESS FIT	0 ms	The tightest thing there is. No play at all.
TIGHT MESH	5 ms	Glossy and nailed, still a voice. The boot state.
RUNNING FIT	10 ms	Corrected and moving.
WORN MESH	20 ms	Audible give. Modern sing.
SLOP	60 ms	Loose. Placement survives.
FREEWHEEL	400 ms	Barely drives at all. Effectively through.

The names are machinist vocabulary because this is a speed reducer and the panel already speaks it — HARD_LASH, SLIP_CATCH, NO_BACKLASH. They are also **literally accurate rather than decorative**: retune speed is precisely how much play exists in the mesh before the gear takes up the load. A press fit is the tightest fit in engineering; a freewheel does not drive. You do not need to have touched a lathe for them to read correctly.

The chips set LAG ONLY. They do not touch the dead zone, which is what makes them useful next to TENSION rather than a second copy of it: set the character with TENSION, then pick the fit that suits the tempo, and the tolerance stays where the macro put it.

The chips are a convenience, not a cage. **MECHANICAL_LAG stays live** for anything in between, and landing off-step simply leaves no chip lit — the panel will not claim PRESS FIT while the engine is doing 3 ms.

9.6 The order that matters

TENSION overwrites MECHANICAL_LAG and INERTIAL_SPOOL the moment you move it. That is what a macro is. It means:

- Set **TENSION first**, then adjust the two knobs if you want to leave the curve. That order gives you the macro's judgement plus your correction.
- Do it the other way round and your two values are gone — the macro stamps its own over them, and nothing warns you, because every value it writes is legal.

The same order applies to the **VOCAL PRESET** bank: a preset that carries a tension applies it *before* its own hand-set lag and tolerance, so a voice dialled by hand recalls exactly as it was dialled. See §13.

Moving MECHANICAL_LAG or INERTIAL_SPOOL directly is always allowed. It simply means you have left the macro's curve on purpose, and TENSION's readout no longer describes what the engine is doing until you move it again.

10. KEY and SCALE

The corrector pulls each detected note to the nearest member of the chosen scale. Wrong scale, wrong targets — this is the single most common cause of a bad-sounding result.

Available scales (14): Chromatic · Major · Minor · Dorian · Mixolydian · Pentatonic · Lydian · Phrygian · Locrian · Minor Pentatonic · Egyptian · Blues · Major Blues · Blues Dorian.

The blues sets

The three blues scales are hexatonic — six notes, not seven — and they exist because a seven-note scale cannot hold a blue note honestly.

SCALE	SEMITONES FROM THE ROOT	THE POINT
Blues	0 · 3 · 5 · **6** · 7 · 10	Minor pentatonic with the b5 restored between the 4th and 5th.
Major Blues	0 · 2 · **3** · 4 · 7 · 9	Major pentatonic with the b3 sitting next to the natural 3rd.
Blues Dorian	0 · 3 · 5 · **6** · 7 · 9	The b5 kept, with a natural 6th. Minor that does not sound sad.

The bolded degree is the reason each set exists, and in every case it is a note a conventional scale would have **corrected away**. Under Minor, a blue b5 is a semitone from two legal targets and gets pulled to whichever is nearer — the correction removes the note the phrase was built on. Under Major Blues, the b3 and the natural 3rd are adjacent and **both legal**, which is what lets a singer sit between them where the blues actually lives.

If a take sounds correct and lifeless, and the singer is bending thirds and fifths, this is usually why.

> **Engine note.** The three blues sets are in the native engine. The legacy preview build carries the original eleven; its DSP has no mask for a hexatonic scale and cannot be given one.

CHROMATIC is a first-class choice, not a failure state. It snaps to the nearest semitone with no key at all. When you don't trust the key, use it. A wrong 7-note scale hands the corrector targets the singer never sang, and those forced corrections are what people hear as doubling or phase weirdness.

Auto-key detection

On import, the engine analyses the **whole file** — about 350 ms for a 4-minute stereo track, no playback required — and proposes a key and scale.

- It scores nine candidate scales across all 12 roots. The engine can *target* fourteen; nine are proposable, and Locrian is excluded deliberately — it shares its seven pitch classes with the natural minor a tone below and roots on the ninth, which is where a great many singers sit. Offered as a candidate it wins arguments it should lose.
- Modes must earn their promotion. Since a mode differs from its parent by a single note, the engine measures that characteristic tone against the note it displaces rather than trusting a hairline correlation difference.
- **It always commits a reading, and tells you how sure it is.** The status line gives you the key, the scale, the evidence or the confidence, and the time taken — `F# MAJOR · HIGH · 82ms`, or `G MIXOLYDIAN · b7 over nat7 (8.4% vs 0.9%) · 82ms` when a mode was promoted on a measured note, or `A MINOR · LOW · 82ms` when the margin was thin.

A low-confidence reading is still set. Earlier builds left the scale on Chromatic when the call was close, and that was the wrong instinct: Chromatic is not a neutral answer, it disables the scale mask altogether. A take loaded and left alone would get no correction to a key at all — where the low-confidence proposal is right most of the time and is labelled a guess either way. A reading you can disagree with beats a measurement thrown away.

Chromatic is reserved for the one case where it is the truth: no tonal content to rank.

Auto-key is a suggestion, not a warden. Set KEY or SCALE by hand and it is yours — the detector will not overwrite it, on this take or on a re-read of it. The status line says so when it stands down: `AUTO-KEY · A# · SCALE HELD AT PENTATONIC MINOR`.

To hand the decision back, double-click HELIX. That re-runs the whole analysis and releases your scale, so the engine is free to propose again. It is the one gesture that overrules you, and only because you asked it to.

KEY LOCK — root by hand, mode by measurement

Press a key on the tuning keyboard to **lock the root**. Press the same key again to release it. The status line confirms both: `KEY LOCK · D# · SCALE STILL TRACKING`, and on release, `KEY LOCK · RELEASED · HELIX FREE`.

What makes this worth a section is what locking the key does NOT do: it does not switch the detector off.

That distinction is the whole idea. An earlier version *did* stop auto-key on the reasoning that the operator had overruled the machine. That was half right. You overruled it about the **root** — which is a thing you know, because you wrote the song or you are looking at the session. You did not overrule it about **which seven notes are actually being sung**, and that is a question no menu can answer, because it depends on the performance rather than the intention.

So the two halves are split along the line where each party is genuinely better informed:

	DECIDED BY	BECAUSE
Root	You, by hand	You know what the song is in
Mode / scale	HELIX, by measurement	It depends on notes actually sung, not intended

With the lock engaged, HELIX goes on ranking every scale **within your root** and commits the one the material supports. Its confidence margin is recomputed against only those candidates — comparing against a root the lock already excluded would describe a choice that was never on the table.

In practice. You know it is in D# but you are not sure whether the take lands minor, Dorian or Phrygian — and the three differ by one note apiece. Lock D#, play the take, and let the engine measure the note that separates them. You have supplied the half you know and kept the half you would only be guessing at.

Re-import does not release it. Loading the next take leaves your root pinned. Pinning a key by hand is a deliberate act; silently undoing it because a new file arrived would be the machine overruling you, which is the exact failure the whole design is built to avoid.

The keybed — ON, BYPASS and REMOVE

The twelve keys under the gear are not decoration and they are not only a key picker. They do three jobs.

They report. The key you are in lights green. The pitch class AURA is hearing right now lights cyan, and it moves as you sing — when the two agree, you are in key, and that is the readout.

They set the key. Left-click any key. Clicking the same key again releases the KEY LOCK — see **KEY LOCK** above.

And they set what happens to each note. Right-click a key to cycle it:

STATE	COLOUR	WHAT IT DOES
ON	normal	Notes nearest this one are corrected to it.
BYPASS	amber	Notes nearest this one are **passed through with no correction at all**.
REMOVE	grey	This note is not a target. Notes nearest it are corrected to the next allowed note instead.

Ctrl-click any key to put it straight back to ON.

Once any key is off ON, the label under the keybed reads **MODIFIED**, because the printed scale name is no longer the whole truth.

When to use BYPASS

When the performance is doing something around one note that you meant.

A scoop up into the tonic, a fall off the end of a phrase, a blue note sitting deliberately between two degrees — all of these are gestures a corrector will flatten, because from the outside they look exactly like being out of tune. Bypass that note and everything nearest it is left alone, including the approach and the exit, while the rest of the take is corrected normally.

It is also the surgical option in reverse: bypass eleven notes and only the twelfth gets corrected.

When to use REMOVE

When the singer is landing nearer a note than the one they meant.

The classic case: the line wants **F**, the singer is under it and lands closer to **E**. The corrector does exactly what it was told and nails them to E — a note that is in the scale, perfectly in tune, and wrong. Remove E, and the nearest available target becomes F, so the same performance is pulled **up** to where it was aiming.

Bypass and remove are not two strengths of the same idea. Bypass preserves an intention; remove overrules a miss. One says *leave this alone*, the other says *this note should not exist*. Reaching for the wrong one is the difference between keeping a scoop and keeping a flat note.

11. DRIVE RATE — the input and output shafts

A speed reducer is defined by two rates and the number between them. For most of this instrument's life the face showed one: **RATIO 3.5:1**, stated against a blank. The drive rate supplies the missing half.

On import the engine measures the tempo of the whole file. That measurement is the **input shaft rate — IR**, in r/min. The **output shaft rate — OR** is that figure divided by the ratio the lever is currently set to. Move the lever and both the ratio and the output rate change together, because they are the same arithmetic.

Where to read it. The input rate is the large figure under the rosette, captioned INPUT SHAFT. Both rates appear together on the curved scale below it, following the lower edge of the glass. The large figure is for reading across a room; the curved pair is for watching the gap between input and output open and close as you gear up and down. They are the same quantity at two reading distances, in the way a tachometer carries both a needle and a digit.

How it is measured

Whole file, on import, in well under a second — the same approach as auto-key (§10) and for a stronger reason. **Tempo is a property of a span, not of an instant.** You cannot know a period without having heard several of them, so a counter that reports in real time necessarily spends its first bars wrong in front of you. The engine waits until it holds the entire file, then states one figure it can defend.

The chain is an onset-detection function — how much the spectrum *rises*, frame to frame, since energy leaving a note is a note ending rather than a beat — then autocorrelation across 60–200 r/min, then a harmonic sum to identify which period actually organises the bar.

What it will not do

It reports dashes rather than a guess. Three cases produce no reading, by design:

- material shorter than about six seconds, which cannot establish a period;
- material with no beat in it at all;
- material where the measurement is not clear enough to defend.

The octave is genuinely ambiguous, and we will not pretend otherwise. 80 and 160 r/min are the same signal, and a half-time feel really is both. The engine prefers the reading a person would tap, but if it shows half or double what you count, it is not malfunctioning — it has resolved a real ambiguity differently than you did.

One known limit, stated plainly. On a *solo, unaccompanied* voice with strong vibrato, the counter can lock onto a subharmonic of the vibrato rate rather than a tempo — a 5.5 Hz vibrato is itself a periodicity, and periodicity is what the measurement is looking for. With any percussion or accompaniment present it reads correctly. Treat a drive rate taken from a bare vocal as advisory.

11.4 TRANSPOSE, MAGIC GEAR, and breaking the chain

The GEARBOX lever moves speed and pitch **together**, because that is what a varispeed does — see §1. TRANSPOSE and MAGIC GEAR are how you take them apart.

TRANSPOSE is the sixth rotary on the panel. ±12 semitones, in whole semitones, and it moves the singer's PITCH without moving the tempo.

It does nothing until MAGIC GEAR is engaged. That is not a limitation, it is the design: with the chain intact the lever already owns pitch, and a second control for the same axis would make the two fight. The sublabel under the knob says `pitch axis · MAGIC GEAR` for this reason.

MAGIC GEAR is the button under the knobs. It breaks the link between the speed axis and the pitch axis.

	TEMPO	PITCH	THROAT
MAGIC GEAR off	× ratio	× ratio	× ratio
MAGIC GEAR on	× ratio	**+ TRANSPOSE**	× ratio

Read the right-hand column carefully, because it is the whole instrument in one table: **the throat follows the drive ratio in both modes.** Breaking the chain frees the pitch, not the formants.

That gives you three distinct sounds from two controls:

- **Chain intact, ratio up.** Everything scales together — the singer is physically smaller. This is the 1958 tape trick, and it is ALVIN.
- **Chain broken, ratio at unity, TRANSPOSE up.** Pitch rises, tempo does not, the throat stays the length it was. A normal-sized person singing high — which is a squeezed adult, and it is HEY MICKEY.
- **Chain broken, ratio up, TRANSPOSE at zero.** Tempo changes, pitch returns to where it started, the throat stays scaled. That is helium: a small throat over an unmoved voice.

TRANSPOSE steps in whole semitones on purpose. This axis moves a *performance* — a third of a semitone of transposition is not a musical idea, it is a detune, and the throat lever is the continuous control.

12. The ADVANCED face

SIMPLE and **ADVANCED**, top right of the ribbon. SIMPLE is the gear, the cards and the panel — the instrument you play. ADVANCED is the instrument you *inspect*: a spectrum analyser, the microtuning grid, and the key/scale parameters gathered in one place.

Nothing on the ADVANCED face is a different engine. It is the same correction, shown at a level of detail that would be noise on the playing surface.

The two faces share their state. Switching back and forth changes nothing you have set.

12.1 REAL-TIME SPECTRUM

The frequency content of what is leaving the plugin, updated live.

Frequency runs left to right, log by default — so the octaves are evenly spaced the way you hear them, rather than the bottom four octaves being squeezed into the first inch. Level runs bottom to top, from the floor up to 0 dB.

CONTROL	WHAT IT DOES
FFT SIZE	1024 … 32768. How finely the analysis divides the frequency axis.
SCALE	LINEAR or LOG frequency axis.
SMOOTH	0…0.99. How much of the previous frame carries into this one.
FLOOR	The dB value at the bottom of the graph.
WAVE	Overlays the waveform itself, drawn through the middle of the graph.
PEAK HOLD	Draws the falling peak envelope over the live trace.
FREEZE	Stops the trace where it is.

FFT SIZE is a real trade and not a quality setting. A larger size resolves frequency more finely and time more coarsely — 32768 will separate two close harmonics and will smear the moment a consonant happens. 2048 is the default because a voice is mostly the second problem.

SMOOTH is what makes the trace readable. At 0 it flickers too fast to follow; at 0.8 it moves like a meter. It changes nothing about the audio — only what you can see.

FREEZE stops the data, not the drawing. The trace it holds is the one that was there when you pressed it, so you can hold a vowel, freeze it, and then compare a second one against it.

If there is no source, the graph says so rather than drawing a flat line, because an empty grid reads as a broken analyser rather than an idle one.

12.2 MICROTUNING

Twelve columns, one per pitch class, each an offset in cents from equal temperament. This is where the instrument stops being a corrector and becomes a *tuning*.

Drag a column to move that note. Double-click a column to return it to 0 ¢. **RESET** clears all twelve.

The **TEMPERAMENT** selector above the grid loads a historical or modern tuning into all twelve columns at once — 12-TET, Vallotti, Young, the meantones, the EDOs. Choosing one fills the grid; you can then edit any column by hand, which leaves the bank and becomes your own.

The bank holds two different kinds of thing, and both are live.

- **The alternate tunings** — 5-EDO, 19-EDO, 31-EDO, Bohlen-Pierce. Choosing one *replaces* the target degrees outright: a "major scale" has no meaning in Bohlen-Pierce, whose period is not even an octave, so the tuning becomes the available notes and the scale mask steps aside. KEY still applies — it is the root the period is measured from.
- **The twelve-note temperaments** — 12-TET, Just, Pythagorean, ¼-comma meantone, Werckmeister III, Kirnberger III, Vallotti, Young. These *bend* the twelve notes you already have, by a cent offset per note name.

The offsets are per NOTE NAME, not per scale degree, and that is the whole point of a well temperament. Werckmeister makes C sweet and F# harsh *as notes*. Index it off the key instead and the character would travel with the key, every key would sound identical, and you would have reinvented equal temperament with extra steps.

The badge on the MICROTUNING header names the temperament in force, and reads **EDITED** once you have dragged a column — because a hand-edited grid has left the bank and is yours, not Werckmeister's.

The snap happens before the bend. A note forty cents flat of C still chooses C, and *then* takes C's offset. The temperament decides where a degree sits, not which degree gets chosen — otherwise a sharpened neighbour could steal a note that was never near it.

The cyan column is the pitch class AURA is hearing right now. It moves as you sing.

12.3 KEY / SCALE / PARAMS

The same KEY and SCALE as the ribbon — §10 — with the six panel rotaries alongside them, so a session can be set from one screen without moving between faces.

Nothing here is unique to ADVANCED. It is the same parameters; changing one changes it everywhere.

12.4 What belongs on this face, and what does not

The rule the instrument follows: **SIMPLE is for deciding, ADVANCED is for checking.**

You set a throat length by ear on the FORMANT · THROAT tab and then come here to see where F1, F2 and F3 actually landed. You hard-tune a take on SIMPLE and then come here to see whether the corrector is snapping to the degrees you meant. The spectrum will not tell you what to do; it will tell you what you did.

A control that changes the sound belongs on SIMPLE. A display that explains the sound belongs here. The microtuning grid is the one deliberate exception, because a temperament is a decision that needs to be *seen* to be made.

13. VOCAL PRESET bank

Eighteen factory voices. Each drops retune, tolerance, humanize, smoothing and formant into the engine in one selection. The menu shows each preset's character line and its retune time.

PRESET	RETUNE	CHARACTER
BELIEVER '98	0 ms	Patient-zero snap. The original hard-tune sound.
TEDDY P	4 ms	Glossy, every note nailed.
HEARTBREAK '08	6 ms	Heavy tune, moody, chest-cavity weight.
MOON HUM	22 ms	Dreamy melodic sing.
CANDY GLOSS	15 ms	Trap-pop wet gloss.
RODEO NIGHT	10 ms	Dark psychedelic reverb-tune.
PURPLE MIC	30 ms	Sing-rap pocket.
CABIN	60 ms	Art vocal, formant-lifted, ethereal.
HARMONY '05	12 ms	Vocoder-adjacent harmony voice.
AUSTIN NIGHTS	45 ms	Country-pop: corrected but organic.
PYRAMID	55 ms	Atmospheric pitched-formant.
DISCOVERY	8 ms	Hard vocoder-tune hybrid.
STRATO GLOSS	25 ms	Radio-bright pop polish.
BAJAN SNAP	18 ms	Vocal-comp accuracy, no drama.
ORANGE CRUSH	2 ms	Frozen glass-pop. Locked hard to the grid, but cold and glassy rather than wobbly.
PATIENT ZERO	400 ms	Dry through. Hear it before we touched it.
DAFT DUCK	400 ms	The helium trick by geometry. Throat halved, pitch untouched.

PRESET	RETUNE	CHARACTER
PINSTRIPES	70 ms	Uptown ballad. Widest tolerance in the bank — vibrato left intact.

Selecting a preset that is already showing re-applies it — use that to snap back after you have moved knobs away.

PATIENT ZERO is your A/B reference. It is the sound of the material essentially untouched.

DAFT DUCK is the only preset here whose effect is the throat chamber rather than the corrector. It is the helium trick, performed by geometry instead of by gas: breathing helium raises the speed of sound in the vocal tract, which lifts the **formants** while the vocal folds keep vibrating at the same rate — the resonances move and the **pitch does not**. Formant +12 halves the tract, 175 mm nominal down to 87.5 mm. The slow 400 ms retune is what makes it rubbery rather than merely small: the correction arrives late while the formants are already shifted. See §2 for the physics.

PINSTRIPES is the opposite instinct, and the reason it exists is worth stating: **restraint is a real setting, not an absence of one.** Every other voice in this bank corrects a singer who needs it. This one assumes a singer whose pitch is already where they meant it, so its job is to catch genuine drift on long sustained notes and touch nothing else. Its 25 ¢ tolerance is the widest here on purpose — a wide, slow vibrato swings further than most dead-zones, and correcting *inside* that swing flattens the one thing that identifies the voice. Turned up, the same singer sounds embalmed.

Preset names describe sonic character and era. They are not affiliated with, endorsed by, or derived from any artist or their recordings.

14. Materials and wear

Two controls that look decorative and are not. **MATERIAL** is a voice preset wearing a finish. **WEAR PROFILE** is the shape of the imperfection the instrument puts back into a perfectly corrected note.

14.1 HUMANIZE — what the wear profile is a profile *of*

Take this first, because the three cards mean nothing without it.

A perfectly corrected note is *exactly* on its degree, forever, until the next note. No human being has ever done that, and the ear knows. **HUMANIZE** puts a controlled amount of wrongness back — a slow drift around the target rather than a fixed offset.

Range	0–100 %, where 100 % is **±25 cents** of drift
Rate	**1.6 Hz** — under two cycles a second
Gated on	voiced frames only

The rate is the part that is chosen rather than tuned. Human vibrato lives between 4 and 7 Hz. Anything the instrument added in that band would be heard as *vibrato* — as the plugin performing — and it would fight the singer's own. 1.6 Hz is well below it: slow enough to read as a machine that is not perfectly rigid, which is exactly what it is meant to be. This is drift, not vibrato.

It only applies while you are singing. Offsetting the target through consonants and breath would wobble material that has no pitch to be wrong about. The oscillator keeps running underneath so its phase stays continuous across a syllable — only its *output* is gated. Turn HUMANIZE up and listen to a consonant: it sits still.

14.2 WEAR PROFILE — three shapes, one amount

The three cards on the right of the panel choose the **shape** of that drift. Same amount, three different motions.

PROFILE	SHAPE	MOTION	SOUNDS LIKE
NO_BACKLASH	sine	Smooth, continuous, symmetrical either side of the target.	A singer breathing. The transparent one — with HUMANIZE low it is nearly undetectable, and it is the default for that reason.
HARD_LASH	square	Sits sharp, jumps, sits flat. No travel between the two.	A gear train taking up its play in one motion. Two distinct pitch centres rather than a wander. Reads as *character* rather than as drift.
SLIP_CATCH	sawtooth	Creeps steadily one way, snaps back, creeps again.	A worn thread slipping under load. The asymmetric one — the ear hears a direction, not just movement.

Nothing is discontinuous in the output, even though two of these shapes are. A square wave stepping the pitch target would make the shifter jump its read pointer, and a jumped read pointer is an audible click. Every profile is therefore slewed over **12 ms** on the way out — under a syllable, over the click. That slew is not a softening for taste; it is what makes the square and the sawtooth usable at all.

Amount and character are independent, and that is deliberate. How much play there is, and what the play sounds like, are two different questions. You can run HARD_LASH at 10 % for a barely-there edge, or NO_BACKLASH at 80 % for a lot of very smooth wander.

14.3 MATERIAL — three finishes, three voices

Click the **pinion hub** — the small gear at the top right of the face — to cycle the alloy. It recolours the gear train, and it loads a complete voice.

The three are the Billboard certification levels, which is why there are exactly three and never a fourth.

	BRASS	GOLD	PLATINUM
MECHANICAL_LAG	40 ms	8 ms	15 ms
INERTIAL_SPOOL	30 ¢	10 ¢	5 ¢
HUMANIZE	70 %	0 %	50 %
SURFACE_FINISH	15 %	25 %	55 %
Wear profile	**HARD_LASH**	**NO_BACKLASH**	**SLIP_CATCH**

BRASS — journeyman raw. Slow correction, a wide dead zone and a lot of humanize. It catches genuinely wrong notes and leaves everything else, including the character. The finish is almost off, so the top end is untouched. This is the setting for a take you like.

GOLD — radio-certified. Fast, tight, and **zero humanize** — the only one of the three that adds nothing back. Every note nailed and held there, RIAA-bright. Note that it pins NO_BACKLASH even though there is no backlash to shape: the profile is what *would* happen if you turned HUMANIZE up, and it stays honest about which one is selected.

PLATINUM — the album cut. Surgical and warm at once, which is why singers reach for it: a very tight dead zone at 5 ¢ so nothing is left uncorrected, a moderate 15 ms glide so it does not click, half humanize on SLIP_CATCH so the result still moves, and the heaviest finish of the three taking the hardness off the top.

Each pins its wear profile independently of its humanize amount, and BRASS is the clearest case: high humanize, and it still lights HARD_LASH. The card is telling you the *character* of the play, not how much of it there is.

14.4 Using them

MATERIAL writes five values at once. It is a starting point, not a lock — everything it sets stays live, so the normal workflow is: pick the alloy that is closest, then move TENSION.

IF YOU WANT…	START AT
The performance, barely touched	**BRASS**
Modern pop, locked and bright	**GOLD**
Corrected but still breathing	**PLATINUM**

And if a take sounds *too* perfect — technically right and somehow dead — the control you want is HUMANIZE at 20–30 % on NO_BACKLASH, not a slower retune. Slowing the correction makes it *late*. Adding drift makes it *human*. Those are different problems and they have different fixes.

15. Other controls

CONTROL	RANGE	EFFECT
MECHANICAL_LAG	ms	Retune speed. Driven by TENSION.
INERTIAL_SPOOL	¢	Pitch dead-zone. Notes inside it are left alone. Driven by TENSION.
SURFACE_FINISH	0…100 %	Metal ring and polish — a high-shelf cut above 6.5 kHz, up to -10 dB.
MIDRANGE_CUT	0…100 %	Fry-tame. Peaking cut at 3 kHz, to -12 dB.
SPEED RATIO	0.5×…2.0×	Time axis. See §1.
TRANSPOSE	±12 st	Pitch axis. See §1.
FORMANT	±12 st	Chamber axis. See §2.
MASTER OUT	dB	Output trim with L/R meters. A limiter stays engaged regardless — it is the safety net against clip-cracks.

Transport: IMPORT · MIC · scrub bar · LOOP · playback controls · MAGIC GEAR toggle · **EXPORT WAV** with selectable sample rate and bit depth (16/24/32).

15.1 Gestures every control shares

Two gestures work the same way everywhere on the instrument, so they are stated once here rather than repeated in each section.

Double-click returns a control to its rest position. Every knob, every lever and every fader. The rest position is the value the control was calibrated around, not zero:

CONTROL	RETURNS TO
Any bipolar rotary — TRANSPOSE, FORMANT	its **centre**, the point where it is doing nothing
Any other panel rotary — TENSION, MECHANICAL_LAG, INERTIAL_SPOOL, SURFACE_FINISH, MIDRANGE_CUT	its **minimum**

CONTROL	RETURNS TO
GEARBOX lever, PITCH TRACK	**1.00× — the nominal 3.5:1 drive**
THROAT lever, FORMANT · THROAT	**0 st — the natural 175 mm chamber**
MASTER OUT fader	**0 dB**
MICROTUNING — any of the twelve note offsets	**0 ¢ for that note**
MICROTUNING, whole grid	RESET clears all twelve at once

This matters most where the detent is fine. The gearbox lever's last tenth of a semitone is about a pixel of travel, and the microtuning grid's cents are finer still — a centre you can see and cannot reliably hit is decoration. Every hardware fader with a centre notch solves this mechanically; a drawn one has to solve it with a gesture, and double-click is the one every plugin already uses.

Hover for a tip. Resting the pointer over a control names it, states its range and reminds you of its reset. The sublabel under a knob is four words wide because the cell is only so many pixels; the tip is where the sentence goes. Nothing in the instrument is a gesture you are expected to already know.

16. Dropdown menus

16.1 The SYS menu

The **SYS** key at the top right of the ribbon. Everything that is about the instrument rather than about the take.

ROW	WHAT IT DOES
SKIN	Eight finishes. The dot beside each is that skin's own accent colour.
OPERATOR'S MANUAL	Opens this document.
RESTORE DEFAULTS	Returns every control to its factory position and clears the loaded preset.
TOOLTIPS	Turns the hover tips on and off.
(identity)	Model, division, version and build date. Not selectable.

The skin diamond at the far LEFT of the ribbon does the same job faster: click it to step to the next finish, right-click it to open this same list. It is also the only always-visible sample of the current accent colour.

Your choice of skin persists between sessions.

TOOLTIPS is worth knowing about. Hovering any control names it and states its gesture — which is how you find out that the play key is also the pause key, or that double-clicking a knob resets it. Indispensable on the first day and in the way on the third, so it is a switch rather than a fact. Turning it off does not disable anything else; the controls behave identically.

RESTORE DEFAULTS does exactly what it says and is not undoable. It is the way back when a session has drifted somewhere you cannot hear your way out of — TENSION to 7, the throat to 175 mm, the drive ratio to 3.5:1, MIX to 100 %.

16.2 The four ribbon dropdowns

VOCAL PRESET, **KEY**, **SCALE** and **TRACKING** open by NAME rather than stepping — clicking through twelve roots to reach B is not a control, it is a chore, and avoiding it is the whole reason these are lists.

All dropdowns are frosted-glass menus. Each row shows the entry, a character line where one exists, and its key value on the right. The colour bar on the left is an index stripe — it makes a long bank scannable, it does not encode anything.

Keyboard: **^ v** to move, **Home / End** to jump, **Enter** to commit, **Esc** to close without changing anything.

16.3 Skins

Eight finishes, chosen from the system menu. A skin changes the panel only — the instrument, its calibration and every reading are identical in all eight.

DARK · GANGSTA · VENOM · ROSE PANTERA · TANGERINE DREAM · AURUM · ICE · **PINSTRIPES**

PINSTRIPES is powder blue over navy, and the ground is pinstriped rather than merely tinted — navy hairlines on a fixed repeat, the way the cloth reads. It is the finish that matches the vocal preset of the same name.

Your choice persists between sessions.

PART III

REFERENCE

17. Troubleshooting

Grouped by what you noticed, not by what is actually wrong — because when something is wrong you do not yet know which it is.

17.1 It sounds wrong

SYMPTOM	CAUSE	FIX
"It's fighting me."	TENSION too high for the performance.	Lower it. At 0–3 only genuinely wrong notes are corrected. §9.3.
"It's chewing the vibrato."	The dead zone is tighter than the vibrato's own excursion, so the corrector pulls against it several times a second.	TENSION below about 4, or raise INERTIAL_SPOOL directly. §9.4.
"It smears between words."	MECHANICAL_LAG is longer than the syllable — the correction is still travelling when the next note starts.	Raise TENSION, or set MECHANICAL_LAG under 10 ms. §9.4.
"It sounds lazy — like it isn't doing anything."	A wide dead zone AND a slow glide together. Most notes are never corrected, and the ones that are arrive late.	**Read MECHANICAL_LAG and INERTIAL_SPOOL before touching anything.** If they are around 5 ms / 15 ¢ when you expected a hard tune, TENSION was moved after them and overwrote both. §9.6.
"Too perfect — technically right and dead."	No humanize.	HUMANIZE 20–30 % on NO_BACKLASH. Do **not** slow the retune — that makes it *late*, not human. §14.4.
Doubling or out-of-phase artefact.	Wrong scale. The corrector is chasing targets that were never sung.	SCALE to Chromatic, or set the key by hand. §10.
Metallic grit on up-shifts.	Known aliasing in the current resampler.	Reduce the up-shift, or shift down from a higher source. §1.10.
Chipmunk after transposing up.	Varispeed drags the formants with the pitch — which is correct behaviour, and is the whole ALVIN sound.	Negative FORMANT to push the chamber back out. §11.4.

17.2 It isn't tracking

Open the GATE card first. It answers this whole section — see §5.1.

SYMPTOM	CAUSE	FIX
Lamp never goes blue.	Voicing below threshold.	If VOICE sits under 0.32 on notes you *are* singing, raise the input level before the plugin.
PITCHED reads NO on sung notes.	AURA cannot find a period.	Almost always the source: two voices on the track, a double, or reverb/bleed. Correct dry, then add the space. §5.1.
PITCHED flickers off between words.	Consonants and breath are unpitched.	**Normal, and deliberate.** They have no pitch to be wrong about and are passed through untouched.
F1/F2/F3 read as dashes.	No confident spectral peak — silence, breath, or an unvoiced consonant.	Normal. They return on voiced material. §2.5.
Auto-key picked the wrong scale.	Ambiguous material.	Override KEY/SCALE by hand — your choice always wins and is never overwritten. §10.
Detector ignores a test tone.	Pure sines do not trip voicing.	Test with real vocal material. A sine has no glottal structure to detect.

17.3 A control does not do what I expect

SYMPTOM	CAUSE	FIX
FORMANT does nothing.	HELIX is in BYPASS — the formant stage is gated by the master engage.	Engage HELIX. Green is enough; blue is not required. §2.7.
TRANSPOSE does nothing.	**By design.** It is inert until MAGIC GEAR is engaged.	Engage MAGIC GEAR. The knob's own sublabel says so. §11.4.
The speed lever still changes pitch.	MAGIC GEAR is OFF, so the lever is plain varispeed — pitch and tempo locked together.	Engage MAGIC GEAR to break the chain. §11.4.
TRACKING will not click.	It is a **readout**, not a control. It reports MECHANICAL_LAG.	Change it with TENSION or the MESH FIT chips. §7.
A preset sounds different each time I load it.	Should not happen. Every preset states every value it sets.	If you can reproduce it, report it — that is a bug, not a setting.

SYMPTOM	CAUSE	FIX
Loading a preset did not change the drive ratio back.	Only ALVIN and HEY MICKEY set the ratio; the rest restore the default if a preset had moved it.	Expected. A voice does not imply a tempo. §13.
A knob is somewhere I did not put it.	TENSION overwrites MECHANICAL_LAG and INERTIAL_SPOOL whenever it moves.	Set TENSION first, then the two knobs. §9.6.
I cannot get a control back to normal.	—	Double-click any knob or lever to reset it. SYS › RESTORE DEFAULTS resets everything. §15.1, §16.1.

17.4 The transport and export

SYMPTOM	CAUSE	FIX
There is no pause button.	**PLAY is pause.** Pressing it while running pauses and keeps your place.	The glyph changes to ■ while playing. STOP is the one that returns to 0:00. §6.2.
MIC does nothing.	Not wired in this build — there is no capture path yet.	Use IMPORT, or run the plugin in a host. §6.1.
EXPORT WAV is greyed out.	Nothing loaded to render.	IMPORT a file first. §6.6.
The exported file is a different length.	The drive ratio changes duration. At 1.75:1 a four-minute take renders in two minutes.	Expected. The time readout shows the rendered length. §6.5.
The export does not match the preview.	Should not happen — the render runs the same chain in the same order.	If it does, report it. §6.6.
Export sounds clipped.	The render peaked above full scale and 16- or 24-bit flattened it.	Export at **32-bit float**, which cannot clip — or lower GAIN. §6.6.
Brief pause when I release the lever.	The stretch renders offline on release, by design.	Expected — one to three seconds on a long file. §1.9.

17.5 Installation

SYMPTOM	CAUSE	FIX
The plugin does not appear in my DAW.	Not installed where the host scans, or the host has not rescanned.	Confirm the bundle is in `C:\Program Files\Common Files\VST3`, then rescan — in FL Studio, *Options › Manage plugins › Find more plugins* with **Verify plugins** ticked.
The plugin looks like an older version.	**A second copy is shadowing the new one.** Hosts scan several folders and a stale bundle in any of them wins silently.	Search the machine for every `WormGear.vst3`, compare their dates, and delete all but the newest. This has no error message and looks exactly like a broken update.
It worked yesterday and now looks different.	Same cause as above.	Same fix.
The standalone opens larger than my screen.	It opens at its design size, 1600 × 1050.	Drag any corner. It scales from 0.5625× to 1.6× and the whole layout follows.

18. Specifications

18.1 Correction

Pitch detector	AURA — two-state Kalman tracker, on its own thread
Detector range	60–1000 Hz
Retune speed (MECHANICAL_LAG)	0…400 ms
Tolerance (INERTIAL_SPOOL)	0…100 ¢
TENSION	0…10, drives both of the above
MESH FIT	6 named steps — 0 / 5 / 10 / 20 / 60 / 400 ms
Humanize	0…100 %, ±25 ¢ at full, 1.6 Hz, gated on voiced frames
Wear profiles	3 — HARD_LASH, NO_BACKLASH, SLIP_CATCH, each slewed 12 ms
Scales	14, across 12 roots
Shifter	PSOLA, pitch-synchronous, period-locked

18.2 The throat chamber

Formant shift	±12 semitones, 0.1 st steps
Nominal vocal tract	175.00 mm
Chamber range	87.50…350.00 mm
Formant analysis	LPC order 12, at 12 kHz, 512-sample frame, 128-sample hop

Reported	F1, F2, F3 in Hz, plus F2/F1 and F3-F2

18.3 Time and pitch

Transpose	±12 semitones, whole steps, active with MAGIC GEAR
Drive ratio	0.5×…2.0× — displayed 7:1…1.75:1, nominal 3.5:1
Gear train	42-tooth worm wheel, 12-tooth pinion, matched module
Drive rate detection	Offline, whole file, 60–200 r/min

18.4 Analysis and output

Spectrum	FFT 1024…32768, log or linear, floor -70/-90/-110 dB
MIDRANGE_CUT	Peaking, 3 kHz, Q 1.4, to -12 dB
SURFACE_FINISH	High shelf, 6.5 kHz, to -10 dB
MASTER OUT	±24 dB, with peak metering

18.5 Files

Import	WAV, AIFF/AIFC, MP3. PCM 8/16/24/32-bit and IEEE float 32/64-bit, any channel count, any sample rate. **Format detected from the file, not the extension**
Take metadata	Title and artist read on import — WAV `LIST`/`INFO`, AIFF `NAME`/`AUTH`, MP3 ID3v2
Export	WAV — 44.1 / 48 / 96 kHz, 16-bit, 24-bit or 32-bit float
Export render	Whole file from the top, through the full chain at current settings

18.6 Plugin

Formats	VST3, CLAP, and a full standalone application
Platform	Windows 64-bit
Channels	Mono and stereo
Latency	**2112 samples** — 44.0 ms at 48 kHz, 47.9 ms at 44.1 kHz. Reported to the host for automatic compensation
Parameters	18, all host-automatable
Factory voices	22
Skins	8
Window	1600 × 1050 design size, resizable 0.5625×…1.6×

18.7 Not in this version

Stated here so the list is in one place rather than discovered one at a time.

AU, AAX	Planned
macOS	Planned
ARA2	**The headline of version 2**
Live microphone input	The MIC key is drawn and inert — there is no capture path yet
Per-note pitch editing	Not the shape of this instrument. Use an editor. §0.4

19. Glossary

Two halves. The first is the vocabulary of audio and tuning — the words this manual and every other pitch corrector use, in plain English. The second is the vocabulary of *this* instrument: the parts, by name.

You do not need to read either. They are here for the moment a word turns up and you would rather not go looking.

19.1 The audio words

Backlash — In a gear train, the play between two teeth before the driving gear takes up the load. Here: the small, controlled wrongness HUMANIZE puts back into a corrected note. See §14.

Bit depth — How finely each measurement of the sound is stored, and in practice the distance between the loudest and quietest thing a file can hold. 16-bit is CD. 24-bit is the studio standard. 32-bit float cannot clip. §6.6.

Cent — One hundredth of a semitone. The unit tuning is argued in, because a semitone is far too coarse: the difference between "in tune" and "slightly off" is about 10–20 cents, and no one has a name for that interval except this one.

Clipping — What happens when a signal is louder than the format can hold. The tops of the waveform are flattened, permanently, and it sounds like distortion. 32-bit float is the one export depth where it cannot happen.

dB (decibel) — A ratio, on a logarithmic scale, because hearing is logarithmic. 0 dB is "unchanged". -6 dB is about half as loud. A number like -90 dB describes something effectively silent.

dBFS — Decibels relative to *full scale* — the loudest a digital file can be. 0 dBFS is the ceiling; everything else is negative. This is why meters count down to zero rather than up.

Dead zone — A range of error the corrector deliberately ignores. Worm Gear's is INERTIAL_SPOOL, in cents, driven by TENSION. Without one, a corrector fights vibrato. See §9.1.

EDO — Equal Division of the Octave — How many equal steps a tuning cuts the octave into. Standard Western music is 12-EDO. 19 and 31 are historical alternatives that make thirds much sweeter; 5-EDO is common in gamelan. §12.2.

F0 (fundamental) — The actual pitch of a note — the rate the vocal folds are vibrating at. What AURA measures and what the corrector moves. Everything above it is harmonics.

F1, F2, F3 (formants) — The resonances of the vocal *tract* — the tube between the folds and the lips. They are what makes "ah" different from "ee", and they are set by the SHAPE of the throat and mouth, not by the pitch. This is the axis the throat chamber controls. §2.

Formant shifting — Moving those resonances without moving the pitch, or vice versa. Moving them together makes a singer sound physically smaller or larger. Moving them apart is what helium and the classic cartoon voices are. §11.4.

FFT — The transform that turns a slice of sound into a picture of its frequency content. The analyser's FFT SIZE sets how long that slice is: longer resolves frequency better and time worse. §12.1.

Latency — Delay the plugin introduces. Worm Gear reports its own to the host so the host can compensate. It is why the exported file is not 53 ms late.

Monophonic — One note at a time. What every pitch corrector needs. §5.1.

Period — How long a waveform takes to repeat itself. Pitch is the inverse of it — a period of 1/200 second is 200 Hz. AURA works by measuring the period, which is why two singers at once cannot be corrected: there are two periods and no single number describes both.

Pitch correction — Detecting the pitch of a note and moving it to the nearest allowed one. The "allowed" list is what KEY and SCALE decide.

PSOLA — *Pitch Synchronous Overlap and Add.* The technique that moves pitch by cutting the waveform into single periods and re-spacing them. Because it works in periods rather than in a spectrum, it keeps consonants crisp — and it is why this instrument needs a period before it can shift anything.

Retune speed — How long the correction takes to arrive. Fast is a robot, slow is a repair. MECHANICAL_LAG, here. §9.1.

RMS — A measure of average level, closer to how loud something *seems* than a peak reading. The scrub strip's colour rides it.

Sample rate — How many times per second the sound is measured. 44.1 kHz is CD, 48 kHz is video and most DAWs, 96 kHz is double. §6.6.

Scala (.scl) — A plain-text file format for describing a tuning. There is a large free archive of historical and experimental tunings in it, and MICROTUNING imports them straight into the grid, where they retune the corrector. §12.2.

Temperament — A *tuning system*: the decision about where the in-tune notes actually sit. Standard tuning (12-TET) makes every key equally slightly wrong so that all of them are usable. Historical temperaments — Vallotti, Young, the meantones — make some keys sweeter and others harsher, which is what a lot of old music was written for. §12.2.

Varispeed — Changing playback speed the way tape does, so pitch and tempo move together. Speeding up by 2× raises pitch an octave *and* halves the duration *and* raises the formants. §1, §11.

Vibrato — Deliberate pitch wobble, typically 4–7 Hz. A corrector with too tight a dead zone will fight it. See §9.4.

Voiced / unvoiced — Voiced sounds have a pitch (vowels). Unvoiced ones do not (most consonants, breath). Worm Gear corrects only the voiced ones, on purpose. §5.1.

19.2 The instrument's own names

Every control here is named for the mechanical part that does that job. **The names are literal rather than metaphorical** — retune speed genuinely *is* how much play there is in the mesh before the gear takes up the load, and output r/min genuinely *is* the corrected frequency times sixty. That is why they read correctly even if you have never touched a lathe.

The gear train

TERM	WHAT IT IS
Worm wheel	The big gear, 42 teeth — the one filling the glass.
Pinion	The small gear, upper right, 12 teeth. Turns 3.5× faster, and the other way.
Pinion hub	The dark disc at the pinion's centre. **This is the MATERIAL selector** — click it to cycle the alloy. §14.3.
The glass	The dark disc the gear sits in. The instrument's display.
Cycloidal drive alignment	The rosette figure in the centre of the glass. It turns with the drive ratio.

The controls

TERM	WHAT IT IS
TENSION	The macro. Relax <-> tighten, 0–10. The main control. §9.

TERM	WHAT IT IS
MECHANICAL_LAG	Retune speed, in ms. How long a correction takes to arrive.
INERTIAL_SPOOL	The dead zone, in cents. How wrong a note must be before anything happens.
MESH FIT	Six named retune speeds — PRESS FIT to FREEWHEEL. §9.5.
SURFACE_FINISH	A high shelf above 6.5 kHz, to -10 dB. Takes the hardness off the top.
MIDRANGE_CUT	A bell at 3 kHz, to -12 dB. Tames fry and harshness.
TRANSPOSE	The pitch axis, ±12 semitones. Inert until MAGIC GEAR is engaged. §11.4.
FORMANT / THROAT LENGTH	The throat chamber, in semitones and in millimetres. §2.
HUMANIZE	How much drift is put back into a corrected note. §14.1.
WEAR PROFILE	The *shape* of that drift — HARD_LASH, NO_BACKLASH, SLIP_CATCH. §14.2.
MATERIAL	BRASS, GOLD, PLATINUM. Three complete voices. §14.3.
DRIVE RATIO / GEARBOX lever	Varispeed, 0.5×–2×, displayed as a gear ratio. §11.
MAGIC GEAR	The chain break. Frees pitch from the drive ratio. §11.4.
HELIX	The correction on/off, and the auto-key detector. §8.
AURA-TONE™	The badge for the pitch engine. Lit when AURA has a confident lock.
MASTER OUT	Output trim, ±24 dB, with peak meters.

The readouts

TERM	WHAT IT IS
GATE [HELIX LOCK]	Whether the corrector is tracking, and how confident it is. §5.1.
GEAR RATIO [PITCH TRACK]	Note in, note out, and the correction in cents.
OR [OUTPUT R/MIN]	The corrected pitch expressed as revolutions per minute — frequency × 60.
VT [THROAT CHAMBER]	Vocal tract length in millimetres, and the F2/F1 ratio. §2.
RESONANT FREQUENCY LOCK	The three measured formants, in hertz. §2.5.
INPUT / OUTPUT HP	Level at each end, in "horsepower" — the gear train's own unit for signal level.

The GATE card, row by row

Six rows, and between them they answer *why is the lamp doing that*. §8.

TERM	WHAT IT IS
STATE	LOCKED, TRACKING or BYPASS. The one-word answer.
VOICE	How strongly a voice is being heard right now, smoothed. The number that must rise, and stay up, before the lamp commits.
PITCHED	YES or NO — whether a pitch was found at all this frame.
HOLD	Counts **up** toward the lock while hunting, **down** from the release once locked. Either way: how far from changing state.

TERM	WHAT IT IS
SOURCE	**AURA** when the pitch is being heard, **MIDI** when it is being told, with the frequency beside it.
KALMAN P	How settled the tracker is. It is named for the filter that produces it — a Kalman tracker carries pitch *and* the rate pitch is moving, which is why a scoop reads as a voice going somewhere rather than a run of errors. Low means it has converged.

Units on the faces

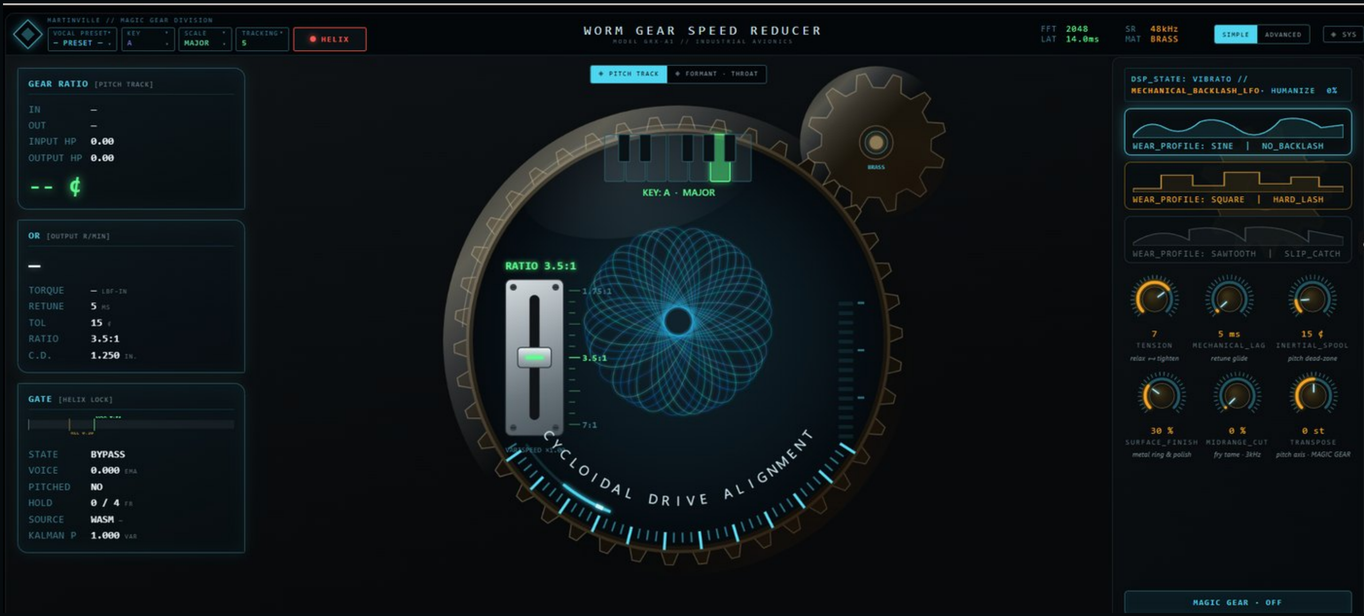
Abbreviations that appear beside a value rather than as a control.

TERM	WHAT IT IS
EMA	Exponential moving average — a smoothed reading rather than a raw one.
VAR	Variance. How much the figure is still moving about.
FR	Frames.
LAT	Latency, in milliseconds. §18.
ST	Semitones.
TOL	Tolerance.
SR / BITS	Sample rate and bit depth, for import and export. §6.6.

20. Index

PLATES

The face



SIMPLE — THE WORM GEAR SPEED REDUCER PLATE

The PITCH TRACK face: the worm wheel and its 12-tooth pinion, the steel SPEED RATIO lever with the ratio ticks reading 1.75:1 / 3.5:1 / 7:1, the CYCLOIDAL DRIVE ALIGNMENT band with its rosette and arc, and the GEAR RATIO, OR and GATE cards. Every reading here is live. The lever sits at unity, so the arc is half-lit and RATIO reads 3.5:1 — see §1.11 for how to read the band, and note GATE showing SOURCE WASM with the lamp in BYPASS, which is §4.5.

Advanced view

ADVANCED — SPECTRUM · MICROTUNING · FORMANT · THROAT

Advanced adds the real-time spectrum analyser, the microtuning grid with its temperament bank, and direct access to every parameter TENSION otherwise drives. The header reports the analyser's resolution live — FFT size and the resulting Hz per bin — so the trade described in §3.4 is stated on the face.

The finishes

Seven finishes, selectable from the **SYS** menu or by cycling the ribbon diamond. Cosmetic only — no audio effect, no parameter moves, and the choice is remembered between sessions. Each is plated below at full width.

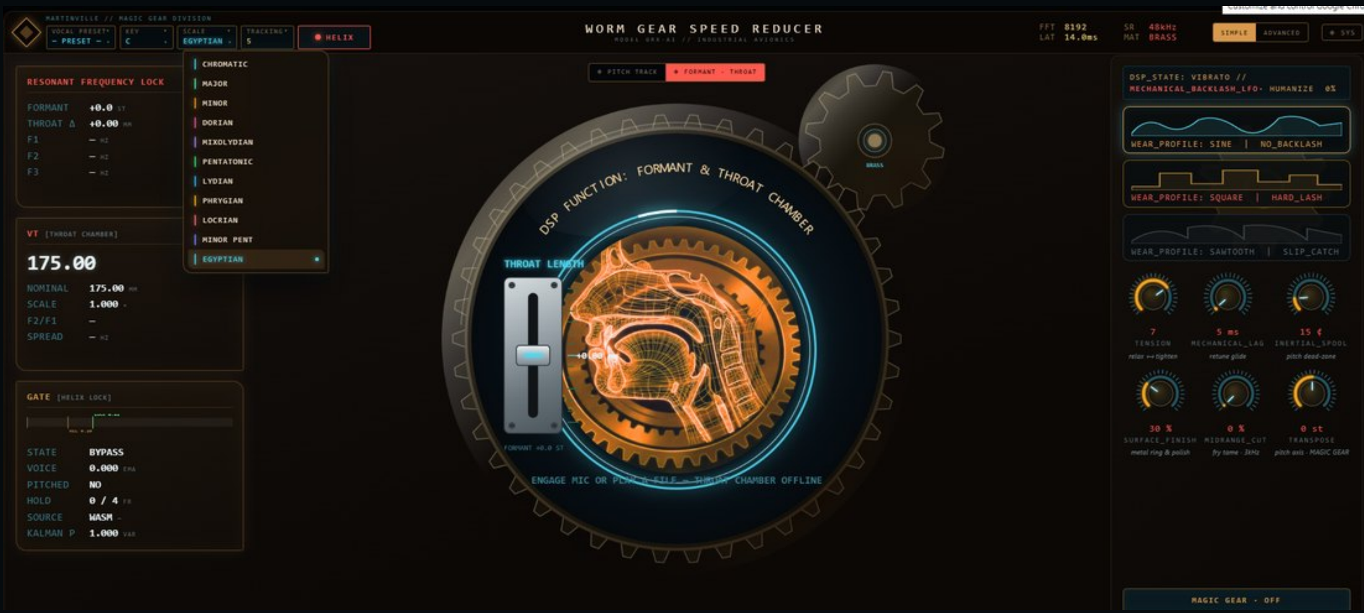
> Captures marked ONE REVISION BEHIND predate the frosted menus and the SYS button. The finish itself is accurate in every case; only the furniture around it has moved on.

DARK

DARK — ONE REVISION BEHIND

The default. Instrument-panel black with cyan reticles — the lowest-contrast finish and the one designed for long sessions in a dim room. Every colour decision in the other six is a departure from this one.

GANGSTA



GANGSTA — CURRENT CAPTURE

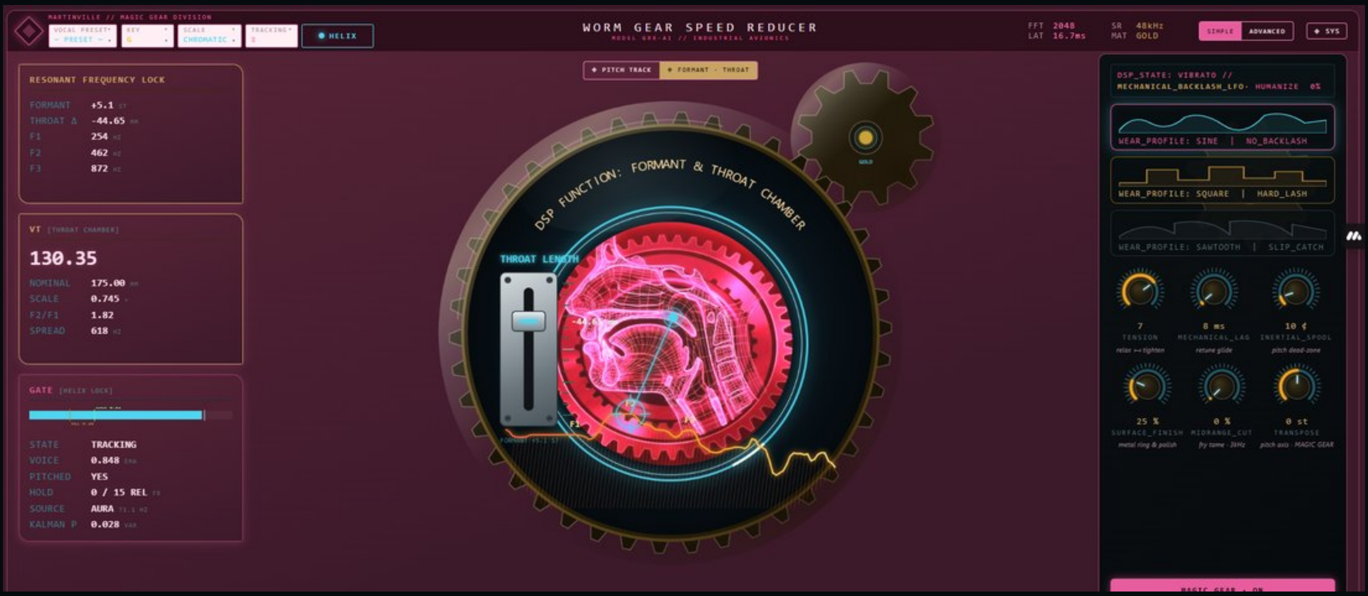
Monochrome chrome over deep charcoal. The gear train reads as bare machined metal with no tint at all, which makes the amber wear cards and the HELIX lamp the only colour on the plate.

VENOM

VENOM — CURRENT CAPTURE

Acid green on near-black. The highest-contrast finish in the set and the easiest to read the alignment arc against, at the cost of being the least restful.

ROSE PANTERA



ROSE PANTERA — CURRENT CAPTURE

Rose gold over oxblood. Warm metal, soft rims, the gentlest treatment of the gear teeth — the finish that most reads as an object rather than an instrument.

TANGERINE DREAM

TANGERINE DREAM — CURRENT CAPTURE

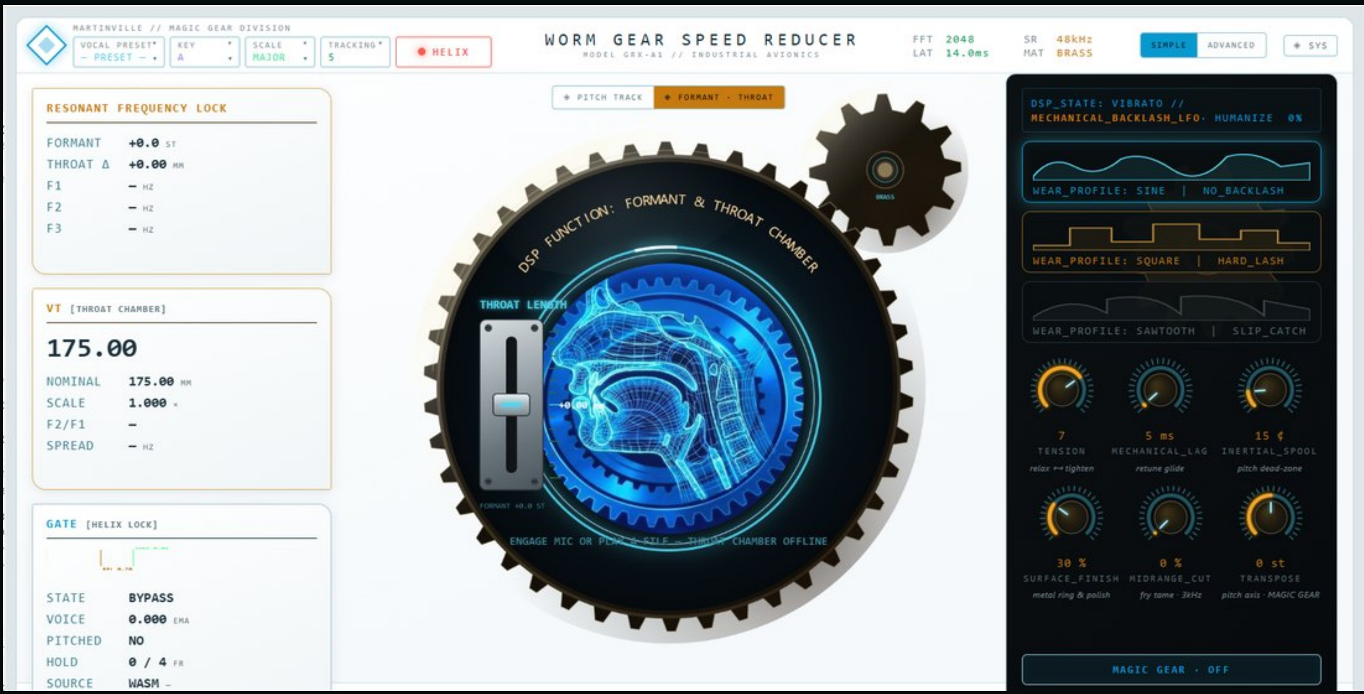
Burnt orange and amber throughout. The wear cards stop being an accent here and become the dominant temperature; the cyan readouts sit against it as the only cool element.

AURUM

AURUM — ONE REVISION BEHIND

Full gold. Polished worm wheel, brass-tinted rims, the warmest and brightest finish in the set. Pairs with the GOLD alloy for a plate where the material and the skin agree.

ICE



ICE — CURRENT CAPTURE

Pale blue-white, the lightest finish. Highest overall luminance, which makes the throat-chamber figure and the spectral envelope read most clearly of any skin.

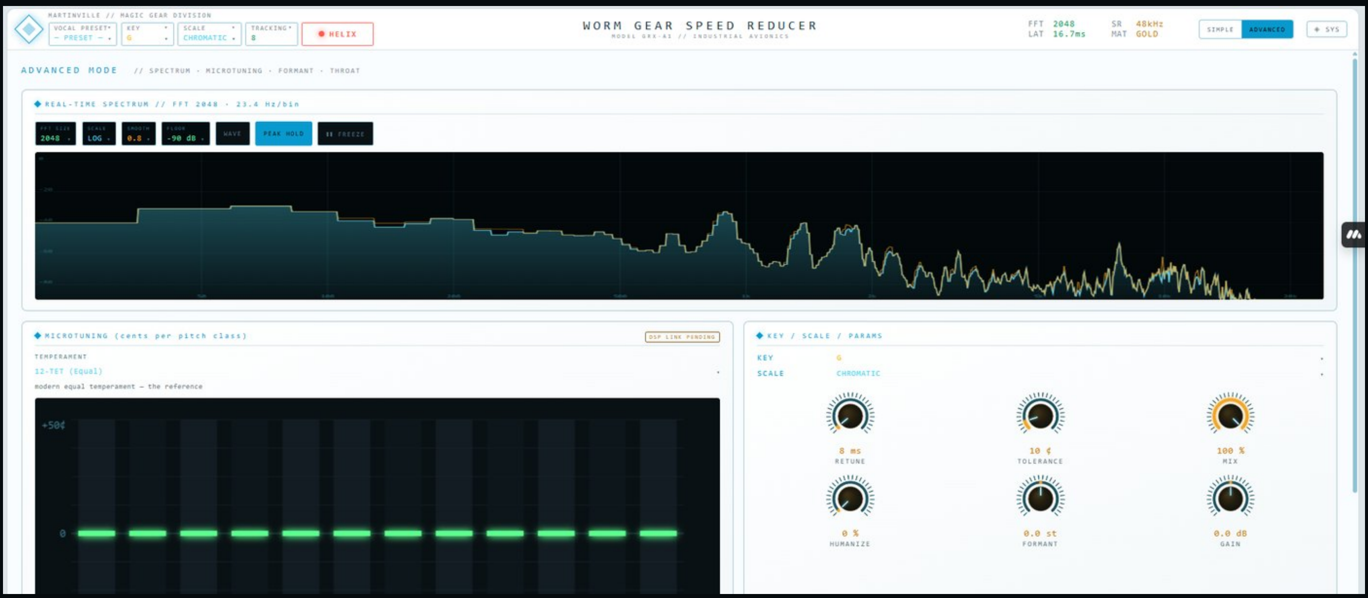
The finishes — advanced view

The same finishes carrying the spectral bench of §3. These are shot with signal running, because an analyser with nothing through it shows nothing: the trace, the peak-hold envelope and the Hz-per-bin readout are all live here. Note how much the finish changes the legibility of the trace against its own floor — this is the one place where the skin choice has a practical consequence rather than a purely cosmetic one.

DARK — ADVANCED

DARK — REAL-TIME SPECTRUM, SIGNAL RUNNING

ICE — ADVANCED



ICE — REAL-TIME SPECTRUM, SIGNAL RUNNING

ROSE PANTERA — ADVANCED

ROSE PANTERA — REAL-TIME SPECTRUM, SIGNAL RUNNING

AURUM — ADVANCED

AURUM — REAL-TIME SPECTRUM, SIGNAL RUNNING