7.6.2 The spectrum of decaying partials

An ideal, un-damped vibration of a string has a harmonic **spectrum** and may be represented by a Fourier-series without much effort. In a **real string**, however, several damping mechanisms are at work at the same time: the string itself radiates sound, heat is generated in its interior, and in the bearings, active power is withdrawn from the vibration of the string. The amplitudes of the partials (constants in the Fourier-series) become time-dependent, and the string vibration looses its periodicity. The standard tool for analyzing non-periodic vibrations is the **Fourier integral** comprising a special window-weighing. Choosing the parameters of the analysis, though, we run into the classic conflict of goals that cannot always be resolved satisfactorily: using a short window duration, the leakage effect broadens the frequency lines, but a long window duration will deteriorate the time resolution too much. If all partials were regularly spread out (Chapter 1.3), we would possibly be able to find an acceptable compromise, but the allpass-driven generation of additional tones requires an analysis with bands as narrow as possible (compare to Fig. 7.39).

As its amplitude becomes time-variant, the spectral line of a continuous tone (represented as a Dirac in the density spectrum) turns into an infinitely broad, continuous spectrum [6]. For the exponential decay process, the Dirac line needs to be convolved with the Fourier-transform of the *e*-function; the parameter of the latter is the **time constant** τ:

$$u(t) = e^{-t/\tau} \cdot \cos(\omega_0 t); \qquad \underline{U}(j\omega) = \left(j\omega + 1/\tau\right) \big/ \left[(j\omega + 1/\tau)^2 + \omega_0^2\right]$$

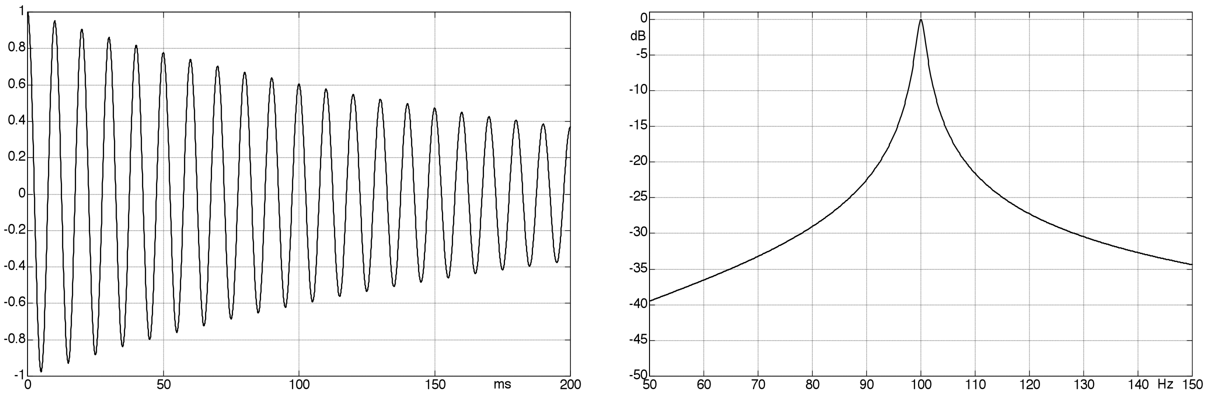


Fig. 7.57: Time function and spectrum of an exponentially decaying cosine-tone; f = 100 Hz, τ = 200ms.

The faster the tone decays (i.e. the smaller τ is), the broader the spectrum will be in the relevant level range. **Fig. 7.57** depicts time function and spectrum of a decaying 100-Hz-tone. In support of a purposeful presentation, the time constant is – at 200 ms – chosen to be relatively short; in guitar strings, values of in excess of 5 s are possible. The widening of the spectrum in Fig. 5.57 must not be confused with your typical DFT-leakage; rather, it results purely from the decrease of the amplitude over time. If we cut a DFT-**frame** (block) starting at $t = 0$ from the infinitely long time function, and transform it into the spectral domain (short-term spectrum), this weighing over time results in an additional convolution in the frequency domain. This is the **leakage** – an additional diffluence of the spectrum particularly noticeable in the area of the peak. The shorter the DFT-frame, the stronger the spectral diffluence is – with the shape of the weighing function over time (window function) to be considered as a further parameter.

Instead of convolving the spectral line twice, it is usually easier to interpret the DFT as a filtering analysis. The transfer function of the analysis filter is the Fourier-transform of the window function. **Fig. 7.58** presents the spectrum of a decaying two-tone-signal, and the frequency response caused by the filter-effect of a Kaiser-window♣: on the left for a 4-k-DFT, on the right for a 16-k-DFT. We can clearly see that the 4-k-DFT will not be able to separate the closely adjoining lines, and even a 16-k-DFT will not be able to deliver the perfect result. There are two reasons for this: the window-lobe is still relatively wide, and moreover the spectra of the two decaying tones will overlap. The smaller the distance in frequency, and the faster the decay process, the more difficult the analysis becomes. From the figure, we can also clearly see a further issue specifically present for the Kaiser window used here: the side lobes. While the latter may be reduced in height by choosing a larger window-parameter (β), this change will, however, further widen the window-lobe even more.

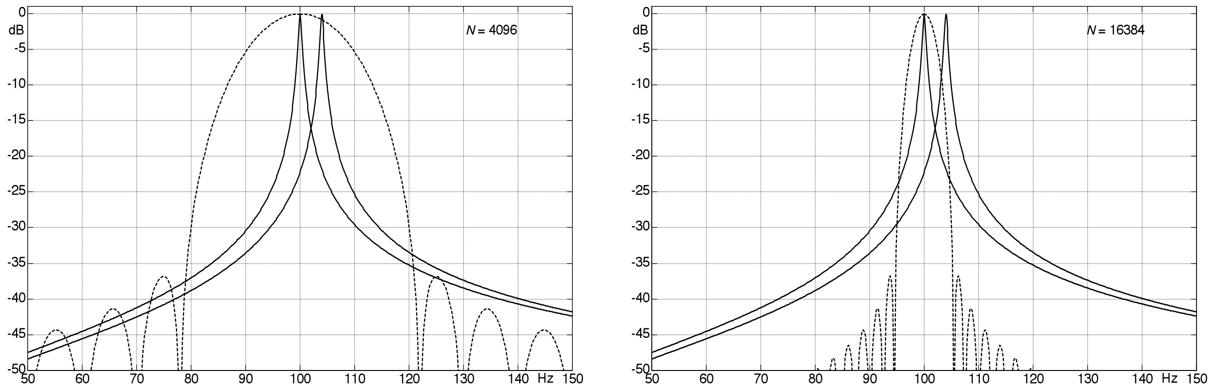


Fig. 7.58: Two-tone signal (100 Hz, 104 Hz, τ = 0.5s), Kaiser-window (β = 5), sampling frequency 48 kHz.

A decaying partial may be described by four parameters: frequency, level, phase and time constant. The **phase** of real guitar tones is of minor importance for the sound (compare to Chapter 8.2.5) – however, all three other parameters should be identifiable with high accuracy via a spectral analysis. If it nevertheless does not work out that way, the fault does usually not lie with the analyzer but with the measuring principle.

Each DFT shows special **level errors** that may well amount to 1.4 dB for a Hanning window: this is known and for the most part acceptable. If not: there are many alternatives with a level error smaller than 1 dB. This is for continuous tones, though! Because for all other signals much larger level errors may result – otherwise we would work exclusively with the flat-top window. Every window has its specific advantages and disadvantages, the awareness of which singles out the expert. Tried and tested are the Blackman-Harris windows, and the Kaiser windows – the parameter of the latter is not consistently specified, though. Not all windows have such a **parameter**. If it exists, it is useful to search for a compromise between strong attenuation of the side lobes, and small bandwidth of the widow. Enlarging the parameter increases the dynamic range of the measurements but deteriorates the spectral selectivity. The parameter should therefore be chosen such that a dynamic range of about 40 – 60 dB is obtained.

The following **table** offers a short overview regarding some important window parameters. More extensive details may be found in specialist literature.

♣ Parameter-definition as customary in MATLAB

Table: Data of common-place DFT-windows [from M. Zollner: Signalverarbeitung, HS-Regensburg, 2010].

	SLA dB	*MLW* lines	Ripple dB	*BW* dB	*FM*	*PM*
Rectangle	13.26	1.62	3.92	0	1.000	1.000
Triangle	26.52	3.24	1.82	1.25	0.500	0.333
Exponential $\alpha = 1$	12.6	1.72	3.65	0.34	0.632	0.432
2	10.8	2.16	3.03	1.18	0.432	0.246
3	–	–	2.35	2.19	0.317	0.166
4	–	–	1.77	3.17	0.246	0.125
Hanning	31.47	3.37	1.42	1.76	0.500	0.375
Hamming	42.68	3.83	1.75	1.34	0.540	0.397
Rosenfeld	48.4	5.78	0.90	2.81	0.381	0.277
Gauss $\alpha = 2.50$	(40)	(5.9)	1.58	1.60	0.495	0.354
3.16	60	7.1	1.06	2.53	0.396	0.281
3.76	80	10.4	0.76	3.27	0.333	0.235
4.32	100	13.9	0.57	3.87	0.290	0.205
Blackman exact	68.24	5.87	1.15	2.29	0.427	0.308
" approx	58.11	5.64	1.10	2.37	0.420	0.305
Blackm.-H. 3/62	62.05	5.38	1.27	2.07	0.450	0.326
3/71	70.83	5.91	1.13	2.33	0.423	0.306
4/74	74.39	6.43	1.03	2.54	0.402	0.290
4/92	92.01	7.88	0.83	3.02	0.359	0.258
Nuttall 3/47	46.74	5.78	0.86	2.89	0.375	0.273
3/64	64.19	5.88	1.05	2.49	0.409	0.296
4/61	60.95	7.79	0.62	3.64	0.313	0.226
4/83	82.60	7.88	0.73	3.27	0.339	0.244
4/93	93.22	7.92	0.81	3.06	0.356	0.256
4/98	98.17	7.33	0.85	2.96	0.364	0.261
Kaiser-Bessel 1.74	40	3.84	1.63	1.49	0.533	0.385
2.60	60	5.45	1.16	2.26	0.431	0.313
3.42	80	7.03	0.91	2.81	0.378	0.272
4.22	100	8.60	0.75	3.25	0.340	0.245
Flat-Top 40	40	5.34	0.05	4.44	0.299	0.247
60	60	7.01	0.05	4.97	0.260	0.212
80	80	8.78	0.05	5.40	0.233	0.188
100	100	10.29	0.05	5.69	0.216	0.173
Dolph-Tsch. 2.4	40	3.80	1.78	1.55	0.537	0.412
3.4	60	5.26	1.29	2.07	0.450	0.326
4.4	80	6.68	1.03	2.61	0.395	0.285
5.4	100	8.06	0.88	3.04	0.356	0.256
Barcilon-T. 2.21	40	3.71	1.74	1.38	0.536	0.395
3.26	60	5.18	1.27	2.12	0.446	0.324
4.27	80	6.60	1.01	2.65	0.391	0.282
5.30	100	7.98	0.85	3.08	0.353	0.254

SLA = Sidelobe Attenuation
MLW = Mainlobe Width
Ripple = Level error
BW = Power bandwidth in dB
FM = window mean value
PM = power mean value

The window parameters may be calculated either from the time function of the window $w[\vartheta]$, or from the spectral window function $W[\nu]$. The time functions of the window are always **amplitude-normalized**, i.e. their maximum value is 1. The zero-point for time is located in the middle of the window for all symmetric windows; for unsymmetrical windows it is at the beginning of the window. The term **polynomial window** characterizes symmetrical windows the time function of which may be described as a superposition of cosine functions:

$$w[\vartheta] = a_0 + \sum_{k=1}^{n} a_k \cdot \cos[k 2\pi\vartheta] \qquad -0.5 \le \vartheta < +0.5$$

Formulae for polynomial window ⇓

Mean value of window *FM:*

$$FM = \frac{1}{N}\sum_{\vartheta=1}^{N} w[\vartheta] \qquad\qquad FM = a_0$$

FM is the linear mean value. Max*(w)/FM* is termed *coherent gain* in English language literature

Power mean value *PM:*

$$PM = \frac{1}{N}\sum_{\vartheta=1}^{N} w^2[\vartheta] \qquad\qquad PM = a_0^2 + \frac{1}{2}\sum_{k=1}^{n} a_k^2$$

PM is the mean value across the squares of the weighing function.

Effective bandwidth B_{eff}:

$$B_{eff} = \frac{\sum w^2}{\left(\sum w\right)^2} \cdot f_a = \frac{PM \cdot \Delta f}{FM^2} \qquad\qquad B_{eff} = \left(1 + \frac{1}{2a_0^2}\sum_{k=1}^{n} a_k^2\right) \cdot \Delta f$$

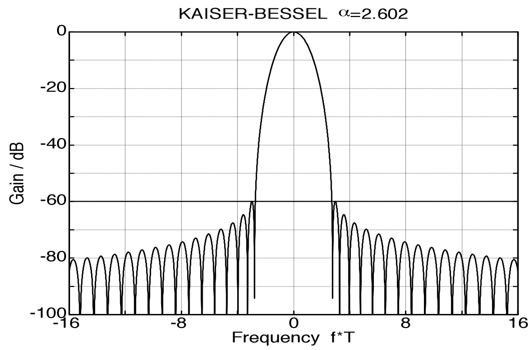
Often, B_{eff} is referenced to the line-width Δf: $BW_L = B_{eff} / \Delta f$; $BW_{dB} = 10\lg(BW_L)$dB.

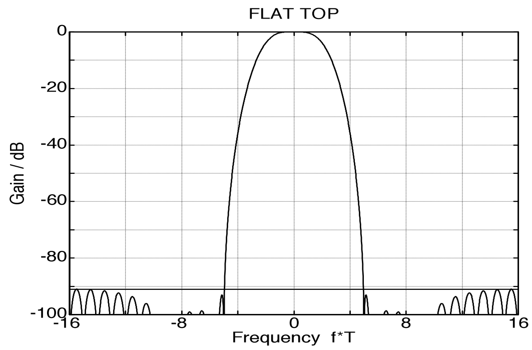
Effective duration T_{eff}:

$$T_{eff} = \frac{\sum w^2 \cdot \Delta t}{\left(\mathrm{Max}(w)\right)^2} = PM \cdot T \quad [\text{for } \mathrm{Max}(w)=1] \qquad\qquad T_{eff} = \frac{a_0^2 + \frac{1}{2}\sum_{k=1}^{n} a_k^2}{\left(\sum_{k=0}^{n} a_k\right)^2} \cdot T$$

Often, T_{eff} is referenced to the duration T of the window: $T_{\%} = T_{eff} \cdot 100\% / T$.

Side lobe attenuation *SLA:*



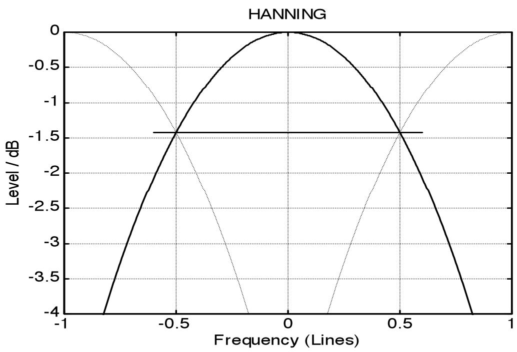


For the Kaiser-Bessel-window shown here, SLA is 60dB; for the Flat-Top-window it is 91 dB.

Main lobe width *MLW:*

MLW is specified as multiple of the line distance; in the example: $MLW = 2\cdot 1.87 = 3.74$

Level error ΔL (Ripple, Scallop loss):



The level error is determined at half the line-distance; in the example: ΔL = 1,4 dB

Further reading:

Brigham E.:	FFT - Schnelle Fourier Transformation, Oldenbourg 1985
Gade S.:	Use of weighting functions in DFT-Analysis, B&K T. Rev. 387
Harris F.:	Use of windows for harmonic analysis, Proc. IEEE, Vol.66, 1/1978[Δ]
Papoulis A.:	The Fourier Integral and its Applications, McGraw-Hill, 1962
Zollner M.:	Signalverarbeitung, Hochschule Regensburg, 2010
Zollner M.:	Frequenzanalyse, Hochschule Regensburg, 2010.

[Δ] This (actually very good) publication contains a number of typing and drawing errors!

If you trust the DFT, you will expect the frequency of a partial to coincide with the local maxima of the magnitude spectrum. **Fig. 7.59** clarifies that this is not necessarily the case: the spectrum of the decaying two-tone signal indeed contains two maxima but these are not located exactly at the frequencies of the two partials. Nor is it evident that, at t = 0, both partials have the same level. Looking at the theory, it does of course work out: since the higher-frequency partial decays faster, its energy within the DFT-frame is lower. However, this cannot be gathered from the figure without knowledge about the signal – and as a rule you will want to analyze *unknown* signals.

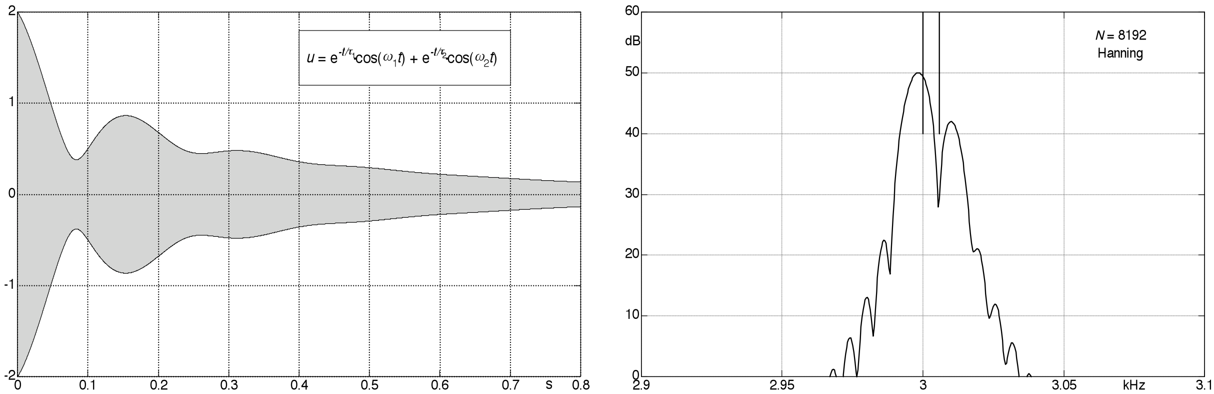


Fig. 7.59: Decaying two-tone signal: f_1 = 3000 Hz, τ_1 = 0.4 s, f_2 = 3006 Hz, τ_2 = 0.1 s; $u = u_1 + u_2$. On the left, the time-envelope is shown; on the right we see the level-spectrum. On the right, two vertical lines mark the frequencies of the two partials.

It is not possible to gather the decay time-constant from *one single* spectrum although it must somehow be contained in the latter. Since that is not the case in an obvious way, we put together a spectral **array** in which the level is registered as a function of time and frequency. To assemble the array, the window is shifted along the time axis (possibly with overlaps) with the shift yielding the abscissa-value for the representation of the level decay. We do hit one snag, though: a purposeful spectrum cannot be obtained for a *point* in time (the spectrum of a Dirac is of "white" characteristic) – spectral analysis is meaningful only for a time *range* (the frame). To derive the decay of a partial from the DFT, we therefore need to know whether the zero on the time axis is allocated to the start, or the middle, or the end of the frame. In the following, we always use the middle of the window as a reference; consequently the level decay cannot be represented from t = 0. Given a sampling frequency of 48 kHz, 8192 sample points result a window duration of already 170 ms, and the first spectrum therefore needs to be assigned to t = 85 ms. This exposes a fundamental issue of every short-term DFT: in order to obtain a high resolution with respect to time, the frame needs to be short (e.g. N = 256), but this leaves the spectral resolution lacking. A longer frame (for example a 32-k-DFT) yields good spectral resolution … but now the time-resolution leaves much to be desired.

So as not to dwell too extensively on synthetic signals, we shall now analyze recordings of a **real guitar string** (∅ = 0,7 mm, M = 68 cm, f = 152 Hz). It was stretched across a stone table with steel cylinders (∅ = 3 mm) serving as bearings (i.e. representing nut and bridge). The bend angle was 17°. A laser-vibrometer took measurements of the transversal vibration at a position of 9 mm from the "nut". The string was struck in impulse-fashion very close to the "bridge". String, strike direction, and laser beam all were in the same plane. The laser signal was sampled and recorded with 48 kHz; the subsequent evaluations were done using MATLAB. The low partial showed regular decay – irregularities start from about 3 kHz.

Fig. 7.60 shows, on the upper left, an excerpt from a level spectrum, and next to it the decay of the 3633-Hz-partial derived from the DFT-array. This spectrum alone would not compellingly reveal any irregular decay of the tone; the short-term DFT, however, shows intense beating. If we change the DFT parameters (only the number of the points shall be varied here), entirely different decay curves result. These have again a different shape if the type of window is changed. With N = 8192, we see not one partial at 3633 Hz, but two – the corresponding levels do still not decay in a linear fashion, though. They show a beating, and thus more partials must exist. The latter can, however, not be isolated even with N = 32768. The window duration amounted already to 0.68 s, so that not much room remained for any further time-shifting within the signal duration of 1 s. By the way: it should not bother us that the minima cannot be reproduced with precision: they result from an interference that is highly sensitive to variations in attenuation. Moreover, it is actually a characteristic of a window to attenuate partials.

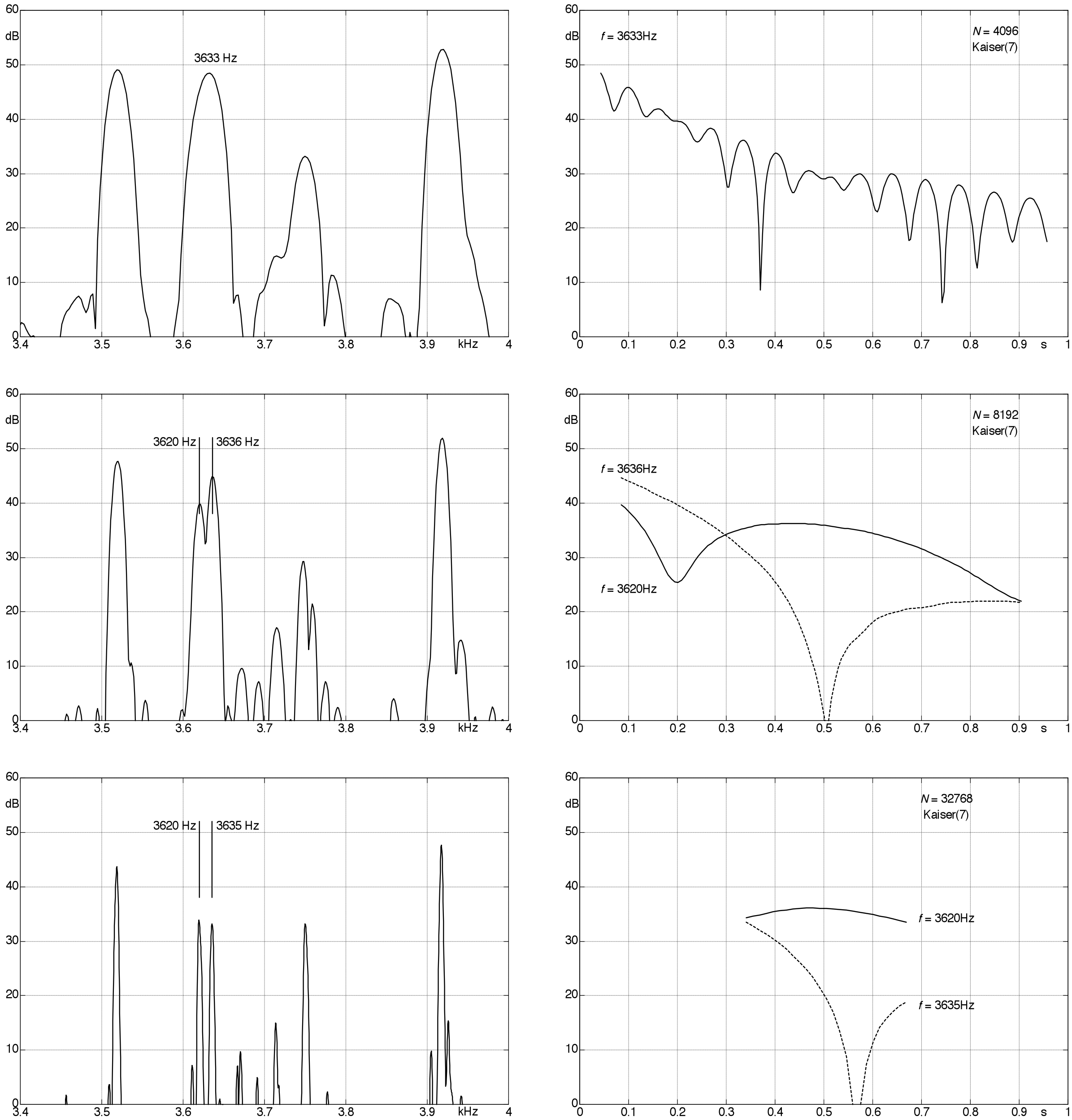


Fig. 7.60: Spectrum (left). Level decay (right) of individual DFT lines. Kaiser-window (MATLAB).

Now, which spectrum is the correct one? Of course, they are all correct – there is an infinite number of (correct) short-term spectra. Which one is purposeful? That question is much more difficult to answer. Counter-question: purposeful for what? Investigating causes, or effects, or mechanics, or psychoacoustics? To determine Eigen-frequencies, a DFT with a high number of points is usually chosen; for the perceived sound, however, the exact frequency of a tone is only of indirect importance. Two tones of small frequency distance are not perceived as detuned but as beating i.e. as a temporal rather than a spectral effect. Therefore, the structure in time will be more interesting for investigations into sound, and short DFT-lengths will be preferable – knowing full well that **psychoacoustics** still have a hard time with complex sounds. On the market for psychoacoustics-analyzers we see devices competing with each other that have very different analysis filters, we see gamma-tone filters put next to 1/3rd-octave-like critical-band filters as if they were equivalent, we see specific loudness calculated via 6-pole reference filters or via true critical-band filters, we see no importance given to the filter phase at all. A rough indication may be determined that way – but not subtle differences in sound. The cochlea is a time-variant, non-linear system the transmission characteristic of which (i.e. frequency response of phase and amplitude) is influenced by the sound signal. In contrast, customary analyzers use time-invariant filters, and if they at all calculate the non-linear fanning out of the upper masking flank (as it is found in the hearing system), they do so after the fact as a correction into the signal flow. This approach works for relatively simple signals but gives merely an orientation for complex sounds.

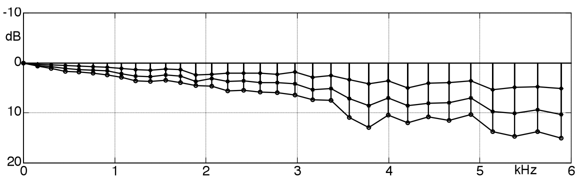
Since evidently it is not possible to determine all partials of a guitar tone metrologically, it is not recommended to try and achieve an ever better frequency resolution via an ever increasing number of DFT-points. *Rather, we could consider whether it is at all purposeful to seek to expand a signal that consists – in the model – of a series of decaying inharmonic partials, into a series of non-decaying partials*! In fact, this is what the DFT-algorithm will do: the process of the Fourier-integral seeks to find steady tones, and it determines the corresponding amplitudes, frequencies and phases. There is no mention of any decay constant. For causal time-functions – the Fourier-transform of which does not include any poles on the $j\omega$-axis – the Laplace-transform may be specified as alternative to the Fourier-transform. This theoretically opens up the possibility to search for residuals, and derived from that a description via poles in the complex p-plane. Among other aspects, MATLAB offers – with the **Prony-algorithm** – a possibility to determine from a signal directly the poles and zeroes of an ARMA-Model (IIR/FIR-filter), and thus to find the Eigen-frequencies and attenuation factors of individual partials. In order not to stress this algorithm too much, it appears purposeful to feed the signal through a band-pass filter first of all, such that only few partials (difficult to separate) remain. Still, it must not be expected that now all signals can be analyzed as desired: each process includes system-immanent artifacts, and with increasing complexity of the signal, these artifacts become more complex, as well♣.

In order not to drown completely in the mist of the speculative, here are two recommendations: for the analysis of low-frequency partials, the DFT is well suited. It may be deployed for the bass strings up to about 1 kHz and for the treble strings up to about 2 kHz. In the higher frequency ranges, it may be purposeful to additionally carry out a 1/3rd-octave analysis that is better adaptable to time effects. Not of any purpose is a loudness analysis of the pickup signal: it never reaches the hearing system in this form!

♣ This highly general statement even leaves room for a seemingly philosophical question: is a square signal very complex because it is composed from an infinite number of sine tones, as is generally known, or is the sine-tone the more complex signal because it may be summed up from an infinite number of square signals?

The bandwidths of **1/3rd-octave filters** (23%) approximately correspond to the bandwidths of the filters found in the hearing system for frequencies above 500 Hz. Therefore, combining neighboring partials into one analysis channel rests on a similar basis. Also, a 1/3rd-octave filter will share with the critical-band filter the characteristic that a very low priority is given to phase responses.

The **Volagramm**♣ gives a clear (yet somewhat arbitrary) representation. It shows the decay of individual partials (in fact: DFT-lines) as a difference spectrum: $L(f, t + \Delta t) - L(f, t)$. **Fig. 7.61** conveys an idea of this approach: level differences were calculated for 4 DFT-spectra (determined for 0 / 170 / 340 / 510 ms) and outlined as polylines. On the left, we see a rather regular decay of the partials, as time passes, the polylines fan out downwards because higher-frequency partials decay more quickly that the lower-frequency ones. On the right, more pronounced irregularities can be seen – caused by fluctuating envelopes of the partials. This representation is not unambiguous because both the type of window and the time-spacing are chosen arbitrarily – but it does provide a quick orientation across frequency ranges of interest.



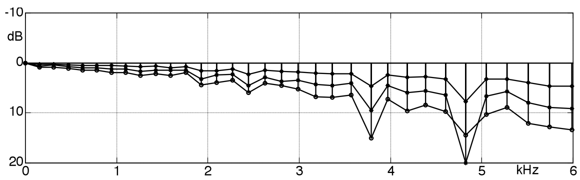


Fig. 7.61: Volagrams: string mounted on the stone table; ends of the string clamped (left) and supported (right). 0,7-mm-string, f_G = 150 Hz, Δt = 170 / 340 / 510 ms, N = 4096, Matlab-Kaiser-window (β = 12). As ordinate, the attenuation is shown; as time progresses, the polyline fans out downward.

7.6.3 The decay time T_{30}

There are several possibilities to quantitatively describe the attenuation in a resonance circuit: degree of damping, time constant, loss factor, logarithmic decrement, measure of decay, or Q-factor. The vibration of a spring-mass-system damped by Stokes-friction will decay exponentially after excitation by an impulse:

$$v(t) = e^{-t/\tau} \cdot \cos(\omega_0 t)\,; \qquad \tau = \text{time-constant}$$

The full designation for the time-constant used in this formula is amplitude-time-constant because it describes the decay of the amplitude. The power also decays exponentially for this vibration, but because power has a square-dependency on the amplitude, the time constant for the power decay will be different: this so-called power-time-constant is half the other time-constant. Standardized sound measurements use e.g. a power-time-constant of 125 ms in the “Fast”-mode of averaging; the corresponding amplitude-time-constant is 250 ms. A time-constant specifies the period of time during which the quantity characterized with it decays to $1/e = 0.368$. Alternatively, a decay to other specific values may be given – such as is practice e.g. with the **reverberation-time** T_N used in room acoustics. During T_N, the signal level drops by 60 dB (i.e. the amplitude drops to 1/1000). Since such a drastic drop is lacking in practical relevance for musical tones, Fleischer [9] has proposed 30 dB as **decay-time** T_{30}.

♣ volare = to disappear, to be volatile, to decay (latin); graphein = to draw (greek).

The decay-time must, however, not be understood such that we pluck the string and then wait until the level has dropped off by 30 dB. Rather, we have to form a **smoothed straight line** in the $L(t)$-diagram, the gradient of which results in the decay-time. On the left in **Fig. 7.62**, we see a perfectly linear level decay. With an exponential decay of the amplitude over time, the level (i.e. the logarithm of the amplitude) will decrease linearly over time. The **decay-rate** – the negative gradient of the curve – is 8.7 dB/s in this example; the time-constant is 1 s and the decay-time is 3.45 s.

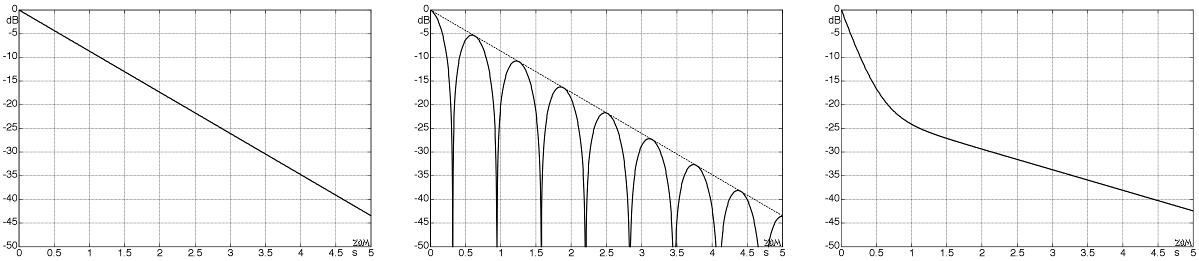


Fig. 7.62: Various decay processes.

The centre curve shows the level decay of a beating signal: after 0.31 s, the level has dropped (relative to the initial value) by 30 dB for the first time. This is, however, not the decay-time – that amounts to 3.45 s just as in the example on the left and is calculated via the (dashed) envelope. Such **beats** occur if two partials of the same initial amplitude and the same damping, but with slightly different frequencies, decay jointly. In this example it is not difficult to find an envelope for the maxima of the curve, and to determine its gradient. The process become more difficult if the periodicity of the beat is much longer, e.g. if the first minimum is only reached after 5 s. It may be impossible to determine the level values of subsequent maxima because the signal has already become too small and disappears in the ever-present noise.

The analysis of the decay becomes even more problematic if partials of very different time-constants decay (Fig. 7.62, right). We could determine T_{30} from the initial slope (as it would be done in room acoustics for the early-decay-time), or from the final slope, or we could – after all – take the point in time when L passes through -30 dB. In the case of a combination of beats and different time-constants this could easily lead to an unusable T_{30}-value, though. In most cases, the decay-time is a highly useful measure to describe decay processes or attenuation. Still, in some special scenarios it may not be purposeful. Therefore, caution is advised especially when using programs that automatically calculate d T_{30}.