Machine translation

FEDERAL SUPREME COURT

ON BEHALF OF THE PEOPLE JUDGMENT

X ZR 119/20

Published on:
November 15, 2022
Schönthal
Clerk of the Court
as Clerk of the
Registry

in the patent nullity case

ECLI:EN:BGH:2022:151122UXZR119.20.0

The X. Civil Senate of the Federal Supreme Court, at the oral hearing on November 15, 2022, by the Presiding Judge Dr. Bacher, Judges Hoffmann and Dr. Deichfuß, and Judges Dr. Kober-Dehm and Dr. Marx,

found in favor of the following:

> The appeal against the judgment of the 6th Senate (Nullity Senate) of the Federal Patent Court of November 18, 2020, is dismissed at the expense of the plaintiff.

By law

Facts:

1 The defendant is the owner of European patent 954 909 (patent in suit), granted with effect for the Federal Republic of Germany, which was filed on March 13, 1998, claiming a German priority of July 14, 1997, and relates to methods for encoding and decoding an audio signal. Patent claim 2, to which six further claims are referred back, reads:

> Method for decoding an encoded audio signal, comprising the following steps:
> Receiving (212) an encoded audio signal;
> Acquiring (214) information in the page information relating to noise substitution and noise ranges of spectral residuals;
> Generating (312) residual spectral noise values based on the acquired information in the noise regions;
> Performing inverse prediction (900) over frequency to obtain spectral values from the noise substituted spectral noise residuals; and
> Transform (218) the spectral values into the time domain to obtain a decoded audio signal.

2 Claim 1, to which the same six claims are referred back, relates to a corresponding coding method.

3 The plaintiff, who is being sued by the defendant under claim 2, has claimed that the subject matter of the patent in suit is not patentable. The defendant has defended the property right as granted. After the expiration of the patent in suit, the parties unanimously declared the legal dispute resolved with respect to claims 1 and 3 to 8.

4 The Patent Court dismissed the complaint. The plaintiff appeals against this decision and continues to seek a declaration of invalidity in respect of claim 2. The defendant opposes the appeal and defends the patent in suit as granted and with an auxiliary request.

Reasons for Decision:

5 The admissible appeal is unfounded.

6 I. The patent in suit concerns the coding and decoding of audio signals.

7 1. According to the description of the patent in suit, such methods have been developed, for example, by the standardization organization ISO/IEC JTC1/SC29/WG11, which is also known as the Moving Pictures Expert Group (MPEG).

8 In the known method, a continuous-time audio signal is sampled to obtain a discrete-time signal (par. 5). This is processed with a function to obtain successive blocks or frames with a certain number of windowed discrete-time samples. Each of these blocks is transformed into the frequency domain, for instance by means of a modified discrete cosine transform (para. 6).

9 The spectral values obtained are quantized in such a way that the quantization noise is superimposed (masked) by the quantized signals and thus becomes inaudible. For this purpose, a psychoacoustic model is used which takes into account the properties of the human auditory system (par. 6). For quantization, spectral values are grouped into scale factor bands that can be scaled by certain factors. The information about this is transmitted to the decoder as side information (para. 7).

10 After quantization, the spectral values would be redundancy coded. For this purpose, they are divided into sections to obtain areas with the same signal statistics. Section boundaries are provided only at scale factor band boundaries. This makes it possible to code a section using a single coding table, such as a Huffman coding table. From the twelve coding tables available, the one that yields

the greatest coding gain is selected. The length of the section in scale factor bands and the number of the coding table used would be transmitted to the decoder as page information (par. 10).

11 The technique of "Temporal Noise Shaping" (TNS), which allows the temporal shaping of the fine structure of the quantization noise by means of a predictive coding of the spectral values, is also known (para. 12). The quantization noise could be put temporally under the actual signal and thus masked. Problems of temporal masking of transient signals or speech signals could be avoided in this way (para. 20 lines 29-38). In the TNS method, an input signal is transformed into its spectral representation by means of a high-resolution analysis filter bank. Subsequently, an linear Prediction is performed in the frequency domain, between frequency-neighboring spectral values. The original spectral values would be replaced by the prediction errors (so-called spectral residuals). These residual values would be quantized and transmitted to the decoder entropy- or redundancy-co-doped just like usual spectral values, so that the values could be decoded again, inverse quantized and inverse predicted (par. 29).

12 The description of the patent in suit further refers to findings from psychoacoustics, according to which the perceptual impression of noise signals is primarily determined not by their actual signal form, but by their spectral composition. This allows the use of a noise substitution technique for data reduction of audio signals (par. 31). One of Donald Schulz's (Improving Audio Codecs by Noise Substitution, Journal of the Audio Eng. Soc. Vol. 44 [1996], No. 7/8, pp. 593-598, D2) is based on the fact that the human ear is not able to detect the exact time course of noisy signals. Coding the waveform requires high bit rates for information that is not audible. If it is possible to detect noise components of signals, one can be content with coding information about the noise level, the

frequency range or the temporal expansion range (par. 33). (par. 33).

13 Against this background, the task of coding an audio signal is to identify noise-like or noisy spectral values in the spectrum of the audio signal (para. 34 lines 5-7). These are defined by the fact that they can be reconstructed without audible differences for the human ear by means of a noise substitution process. Noise ranges in spectral values of the audio signal could be detected in different ways. Corresponding methods were based on the spectral values, on the discrete-time audio signal or on both the audio signal and the spectral values (paras. 49, 34, 35).

14 A group of spectral values classified as noisy is not transmitted to the receiver quantized and redundancy coded as usual. Instead, only an identifier indicating the noise substitution and a measure of the energy of the noisy group of spectral values are transmitted to the decoder as side information. There, random values with the transmitted energy would then be used for the substituted values (para. 36). Because only one energy information is transmitted instead of a group of codes, considerable data savings are possible (para. 37).

15 With the known noise substitution technique, decoding with no audible loss of quality can be achieved if the input signal has a uniform noise structure, i.e., a flat or even spectrum. This is not the case with transient signals or speech signals, so that noise substitution has to be dispensed with here or disturbing distortions have to be accepted (para. 41).

16 2 Against this background, the patent in suit concerns the technical problem of creating a method for encoding or decoding audio signals which enables a high encoding efficiency and results in as little audible signal distortion as possible (paragraph 42).

17 3. To solve this, claim 2 proposes a method whose features can be divided as follows:

> Method for decoding an encoded audio signal, comprising the following steps:
>
> a) Receive (212) the encoded audio signal;
>
> b) Acquiring (214) information in the page information relating to noise substitution and noise ranges of spectral residuals;
>
> c) Generating (312) residual spectral noise values based on the acquired information in the noise regions;
>
> d) Performing inverse prediction (900) over frequency to obtain spectral values from the noise substituted residual spectral noise values, and
>
> e) Transform (218) the spectral values into the time domain to obtain a decoded audio signal.

18 4. The claim requires further discussion.

19 a) The patent in suit is primarily concerned with the MPEG-2 AAC standard, which was under development at the relevant time. However, the subject matter of claim 2 is not limited to this.

20 Also covered are other methods for coding and decoding audio signals using transform coding, which are similarly structured with respect to the points of interest herein.

21 b) Claim 2 deals only with noisy groups of spectral values for which noise substitution is performed.

22 For such groups, according to feature b - unlike usually for non-noise groups - no signal values are encoded, but side information that allows the reconstruction of the noise signal.

23 c) According to feature c, this noise substitution does not refer to the original spectral values, but to residual spectral noise values, such as can be generated using the TNS technique.

24 d) According to the description of the patent in suit, the combination of noise substitution and TNS technology enables an increase in coding gain without audible signal distortions.

25 The patent in suit states that the spectral residual values have a significantly lower energy content than the original spectral values. The corresponding signal has a flatter course compared to the original signal. By predicting the spectral values over the frequency, the strongly fluctuating course of the envelope of transient signals is extracted, leaving a signal with a flat envelope to which noise substitution can be applied in order to achieve considerable bit savings even with transient signals (paras. 44, 52).

26 II The Patent Court gave the following reasons for its decision - insofar as they are still of interest in the appeal proceedings:

27 The subject-matter of claim 2 was not suggested by the prior art.

28 The paper by Schulz (D2) mentioned in the description of the patent in suit represented a suitable starting point for solving the problem posed because it concerned the improvement of methods for encoding and decoding audio signals (audio codecs) by noise substitution.

29 Spectral values detected as noisy would not be encoded in the method disclosed in D2, but would be replaced by information on the noise ranges, the average power of the groups of noisy spectral values and the noise envelope. Such information is understood by the expert, a graduate engineer in the field of electrical or communications engineering or a graduate computer scientist with

several years of professional experience and relevant knowledge in the field of digital signal processing, in particular the coding and decoding of audio signals, as page information.

30 D2 does not disclose the detection and substitution of noise ranges in spectral residual values. Therefore, no decoding method with the corresponding specifications of features b, c and d can be taken from D2.

31 Such measures would not be obvious on the basis of D2. None of the possibilities taught there to distinguish between noisy and tonal signal components of an audio signal uses a prediction of spectral values over frequency to obtain spectral residual values.

32 However, the expert may have had reason to improve the coding method described in D1 by the method for detecting and substituting noise presented in D2. However, this merely leads to a coding method which, according to D1, carries out a prediction of spectral values over the frequency with the aid of the TNS technique and, according to D2, additionally detects and substitutes noise-like components of the audio signal, using a prediction of the spectral values, of the audio signal or of the individual sub-bands over time. On the other hand, there is no indication or suggestion for detecting information relating to noise ranges of the spectral residual values (feature b), for generating spectral noise residual values on the basis of such information (feature c) and for performing an inverse prediction over frequency to obtain spectral values from such residual values (feature d).

33 From the publication by Herre and Johnston (Enhancing the Performance of Perceptual Audio Coders by Using Temporal Noise Shaping (TNS), In: An Audio Engineering Society preprint 4384 (N-3), Presented at the 101st Convention 1996 November 8-11 Los Angeles, California, D5/D11) and the draft standard MPEG-2 AAC (ISO/IEC 13818-7 First edition 1996(E). MPEG-2 Audio NBC Committee Draft. ISO/IEC JTC1/SC29/WG11 N1307, D9), do not

show anything different. Their content did not go beyond the content of D1 with respect to the subject matter of claim 2.

34 Whether the paper by Edler (Very Low Bit Rate Audio Coding Development, in: Proc. 14th Audio Eng. Soc. Int. Conf., June 1997, D12) was pre-published, it was not necessary to decide, since the subject-matter of claim 2 was not suggested even on the basis of this.

35 In the PARA codec described in D12, when analyzing the audio signal for frames of samples, parameters describing sinusoidal components (e.g., frequency, amplitude, and phase) and noise components (e.g., spectral envelope) would be extracted. Due to the perceptual model, phase information for the sinusoidal component would not be transmitted. Furthermore, the transmission of the spectral envelope for the noise component is sufficient and no residual signal has to be transmitted. This residual signal was not disclosed as a difference between the actual spectral value of the audio signal and its prediction by a spectral value adjacent in frequency and thus not as a spectral residual value. D12 also did not contain any other reference to the relevant teaching of the patent in suit.

36 The other publications introduced into the proceedings are still available or have been republished.

37 III This assessment stands up to scrutiny in the appeal proceedings.

38 1. The Patent Court correctly held that the subject matter of claim 2 was not suggested based on D1 and D2.

39 a) D1 does not fully anticipate the features of claim 2.

40 aa) D1 deals with temporal noise shaping in section 8.1. (Temporal Noise Shaping).

41 For transient and tonal input signals with a non-flat spectrum, achieving a masking effect in the reproduced audio signal is problematic due to the time discrepancy between masking threshold and quantization noise. This is known as the pre-echo problem.

42 The TNS technique makes it possible to control the fine temporal structure of the quantization noise even within a filter bank window. By applying predictive coding techniques to spectral data, the level of quantization noise could be effectively brought below the actual audio signal. This allows more efficient use of masking effects by matching the temporal fine structure of the quantization noise to that of the masking signal (p. 23 para. 2).

43 The predictive coding and decoding process over frequency could be realized by adding a building block to the standard structure of an encoder and decoder. After the analysis filter bank, TNS filtering would be added, as shown in Figure 8.1 reproduced below.

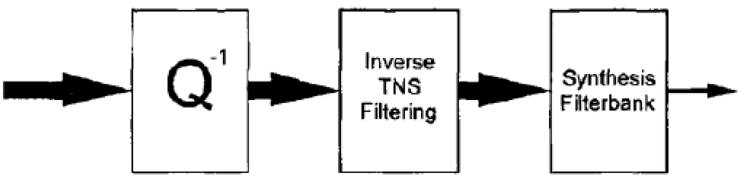

Figure 8.3 - TNS Processing for the MPEG-2 NBC Decoder.

44 There would be a prediction of spectral values of the audio signal over the frequency to replace them by their spectral residual values (section 8.1.1 first paragraph). These residual values would be quantized and transmitted to the decoder entropy- or redundance-encoded.

45 In a corresponding manner, as illustrated by Figure 8.3 below, inverse TNS filtering is provided before the synthesis filter bank during decoding.

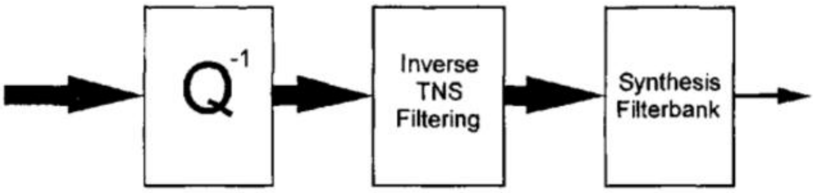

Figure 8.3 - TNS Processing for the MPEG-2 NBC Decoder.

46 bb) Thus, as the Patent Court correctly assumed, a method for decoding an encoded audio signal is disclosed in which, in the sense of feature a, an encoded audio signal is received (block Q^{-1}) and an inverse prediction is performed over frequency to obtain spectral values from spectral residual values. To obtain a decoded audio signal, these spectral values must then be transformed to the time domain according to feature e.

47 cc) Features b to d are not fully disclosed.

48 D1 does not provide for detecting noise regions in the spectral residuals during encoding, replacing the spectral residuals for these regions with information relating to the noise regions, and including this information in the side information of the encoded audio signal. Accordingly, there is also no disclosure of decoding steps in this regard.

49 b) D2 also does not anticipate all the features of claim 2.

50 aa) D2 states that the reduction of the required data in audio coding has so far been achieved mainly by masking effects, for example by not coding

components that are masked by other components. However, most audio signals also contain frequency ranges that are not perceived by the human ear.

51 For the correct coding of such ranges, high bit rates would be needed so far. If, on the other hand, the detected noisy components of the signal were coded only as information about the noise level, the frequency and the time domain, a correct description of the noisy signals could be achieved with considerable savings (p. 593 left column, section 0).

52 A signal component is noisy if it can be characterized by the level, the frequency range and the time range in such a way that a reconstruction is possible without the listener noticing a difference (p. 593 right column, section 1). There are various ways in which noisy components can be distinguished from non-noisy (tonal) components and recorded.

53 According to D2, it is advantageous to combine data compression of audio signals for noisy signals with conventional compression for non-noisy signals. As for noise substitution, it is also necessary to subdivide the entire frequency band for masking. This could be done by adaptive transform coding, such as modified discrete cosine transform (MDCT) or by using a subband filter bank, he said. With MDCT, high compression can be achieved, but it requires more computing power than subband coding (p. 595 right column, section 2).

54 As illustrated in Figure 4 below, adaptive transform coding with noise substitution involves transforming an audio signal and a prediction of that audio signal into the frequency domain.

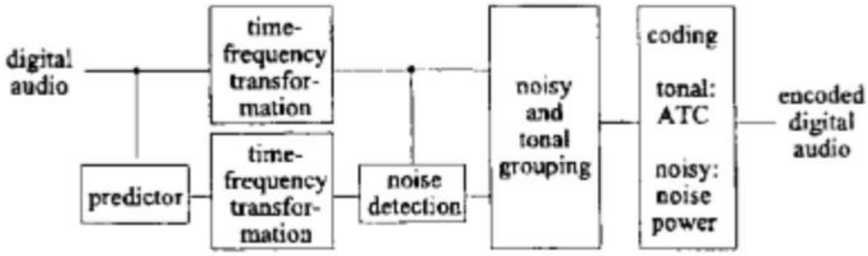

Fig. 4. Noise substitution in combination with adaptive transform coding.

55 Noise spectral values are detected by comparing the spectral values of the audio signal with those of the predicted audio signal (p. 596 left column, paragraph 1). For this purpose, a prediction could be used in a corresponding manner as shown in the following figure 2 for noise detection in subbands.

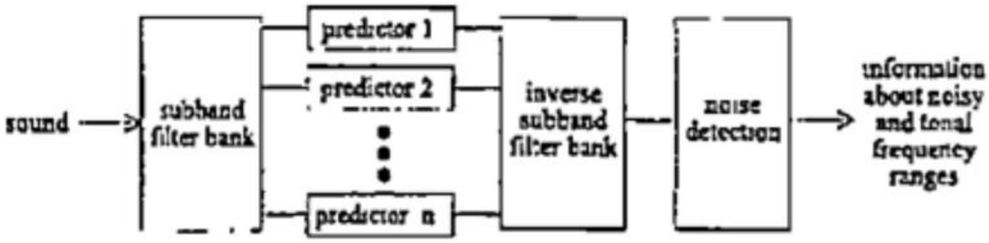

Fig. 2. Noise detection by prediction in subbands.

56 The audio signal is divided into subbands by a multiphase quadrature filter bank. In each subband, a separate prediction is performed over time. The predicted subband signals would be recombined using the inverse filter bank and transformed into the frequency domain using a fast Fourier transform (FFT). Equation (3) reproduced below could be used to detect the tonality of the spectral values and thus to distinguish tonal from noisy spectral values (p. 594 right column, section 1.4).

$$T^i(n) = \alpha \left| \frac{\hat{P}^i(n) - P^i(n)}{P^i(n)} \right| + \beta \left| \frac{\hat{\Theta}^i(n) - \Theta^i(n)}{\Theta^i(n)} \right| \quad (3)$$

57 In this respect, the Patent Court found without objection that the tonality measure T^i (n) includes the normalized difference between the spectral value ^{i} (n) of the frequency *n* in the *i-th* frame of the signal combined from the predicted subband signals and the spectral value P^i (n) of the frequency *n* in the *i-th* frame of the original audio signal.

58 To increase the compression factor, neighboring tonal and noisy frequency values are grouped together. A noisy spectral value (noisy frequency value) that does not have any noisy neighbors is considered tonal. Since the auditory system perceives time delays between the right and left channels in the frequency range up to 5 kHz, noise substitution should be limited to the range above 5 kHz. If a transformation with a frequency resolution of 40 Hz is used, an envelope of the time structure of noisy signal components is required. Then the average energy level of the remaining group, its frequency range and its envelope would be transmitted to the decoder (p. 596, section 2.1, penultimate paragraph).

59 bb) The Patent Court correctly assumed that the description of the coding method also discloses the necessary inverse method steps for decoding the signal. Consequently, D2 discloses features a and e.

60 cc) Features b to d are not disclosed.

61 To this extent, D2 does disclose capturing information in the page information relating to a noise substitution and to noise ranges of the spectral

values to produce noise spectral values in those noise ranges.

62 On the other hand, it is not disclosed that the detected information relates to noise ranges of the spectral residual values, as provided by features b to d.

63 c) Based on D1 and D2, the subject matter of claim 2 was not suggested. The prior art did not convey any suggestion to apply the noise substitution also to the spectral residual values obtained by using the TNS technique.

64 aa) As the Patent Court rightly assumed, however, there was reason to examine, on the basis of D1, whether the noise substitution proposed in D2 could be incorporated into the coding method provided for the standard.

65 (1) The starting point of the TNS technique described in D1 is to perform a prediction of spectral values of the audio signal over the frequency, for example in the case of transient signals, in order to shape the temporal fine structure of the quantization noise in such a way that it can be masked by the audio signal and, in particular, that no audible pre-echoes occur. Only in this context the spectral values are replaced by their spectral residual values (pre-diction residual) and transmitted to the decoder quantized and entropy or redundancy coded.

66 In contrast, the method proposed in D2 is based on the approach of not encoding noisy spectral values in the usual way, but merely transmitting values that allow the spectral values of noisy groups to be replaced by spectral values randomly generated on the decoding side. In this context, a prediction of the audio signal is also carried out. However, this is not used for the formation or analysis of residual values, but for the detection of noisy components of the audio signal.

67 In order to arrive at the teaching of the patent in suit, it was necessary to recognize that the noise substitution proposed in D2 offers advantages not only with respect to noisy regions of the original signal, but can also be applied with comparable or even better effect to spectral residual values determined by means of prediction. As the Patent Court correctly recognized, no suggestion for this is apparent from D1 and D2.

68 (2) Contrary to the opinion of the appeal, it was not sufficient to simply connect the procedures according to D1 and D2 in series. Such a combination would indeed be possible with regard to the individual process steps. However, in order to arrive at the teaching of the patent in suit, the process steps provided for in D2 must be carried out on a different object, namely on the spectral residual values determined according to D1. This kind of combination was not obvious without suggestion.

69 (3) Contrary to the view of the appeal, a corresponding suggestion did not arise for the expert from the reference of the D2 to the time resolution of the human auditory system and the resulting problem of jointly encodable groups of no more than three subband samples (p. 596, section 2.3, first paragraph).

70 These explanations relating only to subband coding do not provide any indication of transformation coding according to D1. Since subband coding has a higher temporal resolution and the causes described in D1 for the occurrence of pre-echoes are not present, it is also not apparent why the person skilled in the art should establish a connection with the temporal noise transformation technique according to D1 in the case of subband coding. Furthermore, there is no evidence that a prediction via the frequency in subband coding would lead to any relevant improvement in the temporal resolution at all.

71 (4) Nothing else follows from the disclosure in D2 that noise is detected by comparing the frequency values of the predicted sound transformed into the frequency domain with the original sound (p. 596 left column, first paragraph; Fig. 4).

72 The Patent Court has found that the proposed pre-diction according to Figure 2 is also carried out over time. The appeal does not show any circumstances which give rise to concrete doubts as to the completeness or correctness of these findings.

73 bb) No further suggestions result from a combination of D2 with D1, D5 or D11.

74 The fact that D2 describes the prediction shown in figure 2 as ideal (p. 596, left column, first paragraph) does not exclude alternatives. However, it does not suggest a fundamentally different prediction model.

75 A suggestion to search for prediction methods outside the disclosure of D2 can also not be derived from the fact that, according to D2, equation (3), which works well for true stationary signals, fails for the sinusoidal curves with frequencies changing over time that frequently occur in audio signals and that this method can therefore not be used together with noise substitution (p. 594, last paragraph before section 1.2). This is because D2 presents its own solution approaches for this in the subsequent sections.

76 Even if it is assumed that there was reason to consider other prediction methods than those presented in D2, there is in any case no suggestion to make a prediction over frequency instead of a prediction over time.

77 Thus, contrary to the view of the appeal, there was also no reason to consider the prediction model disclosed in D1, D5, or D11 for the TNS technique, in which spectral residuals are generated in connection with the shaping of the

temporal fine structure of the quantization noise, and to examine it for usability in connection with the noise substitution taught in D2.

78 cc) D5, D11 or D9 do not, according to the unchallenged findings of the Patent Court, go beyond the disclosure of D1 in a relevant way with regard to the description of spectral residual values in connection with the shaping of the temporal fine structure of the quantization noise according to the TNS technique. Thus, no further suggestions result from these publications.

79 2: Finally, the Patent Court correctly held that the subject matter of claim 2 was also not suggested by a combination of the prior art TNS process, such as disclosed in D11, with D12.

80 a) According to D12, the insensitivity of the human ear to changes in the waveform of audio signals allows the development of coding techniques based on an efficient representation of spectral properties without generating an approximation for the temporal waveform of the input signal. The underlying model for these coding techniques was based on the assumption that the input signal consisted of sinusoidal signal components and noise-like components with relatively constant properties for certain time intervals.

81 The corresponding analysis of the audio signal would have to extract parameters describing sinusoidal components (e.g. frequency, amplitude and phase) and noise components (e.g. spectral envelope) for frames of samples. Phase information for the sinusoidal component would not need to be transmitted. Furthermore, the transmission of the spectral envelope for the noise component

would be sufficient. A residual signal does not have to be transmitted. The model can be refined with regard to speech and music content of an input signal.

82 D12 refers to an approach based on this technique as PARA codec because it is based on a parametric representation of the signals rather than an approximation of the waveform (p. 3 Chapter 3, first paragraph).

83 b) The Patent Court correctly assumed that this does not disclose spectral residual values describing the difference between the spectral value of the audio signal and the prediction by a spectral value neighboring in frequency, and thus a fortiori not the use of noise ranges of spectral residual values or spectral noise residual values in the sense of features b to d.

84 The objection of the appeal that the term noise substitution can be defined as replacing a waveform representation of the noise signal by its para-metric representation does not lead to any other result. For even such a consideration does not necessarily lead to the assumption that spectral residual values are generated in the sense mentioned and are detected and processed according to the claimed method.

85 IV The decision on costs is based on Sec. 121 (2) Patent Law and Sec. 97 (1) Code of Civil Procedure (ZPO).

Bacher Hoffmann Deichfuß

Kober-Dehm Marx

Lower court:

Federal Patent Court, decision of 18.11.2020 - 6 Ni 30/17 (EP) -