J. Acoust. Soc. Jpn. (*E*) 7, 4 (1986)

Detection of slight defects in ball bearings by non-periodic analysis

Hiroshi Kanai,* Masato Abe,** and Ken'iti Kido*

**Research Center for Applied Information Sciences, Tohoku University, 2-1-1, Katahira, Sendai, 980 Japan*
***Education Center for Information Processing, Tohoku University, Kawauchi, Sendai, 980 Japan*

(*Received 27 September 1985*)

This paper describes a new automatic method for detecting defects in ball bearings. The detection of defects is currently carried out by inspectors who listen to vibration signals obtained by a vibration pick-up in the Anderon meter. The pick-up is attached to the outer ring of a ball bearing while the inner ring rotates at a uniform speed. Several methods have been proposed so far for the automatic detection of flaws. These methods are based on the periodicity of vibration pulses excited by flaws. However, the periodicity of vibrations is not always guaranteed when (a) there are slight flaws on the surface of the race, (b) there are flaws on the surface of balls, and (c) there is dust in the grease. However, we have developed a new method, by which the non-periodic resonant vibrations due to both slight flaws and dust are detected. We have applied this new method to the detection of these defects in small-sized ball bearings; the defects were detected with an 98% accuracy rate.

PACS number: 43. 85. Ta, 43. 60. Gk

1. INTRODUCTION

When ball bearings which have flaws on the surface of the ball or races or dust in the grease are used to support the spindle of the head or the capstan in a video tape recorder, such defects debase the quality of the played-back picture. Therefore, currently at the final step in the manufacturing process, inspectors check the bearings, one by one, by listening with headphones to the vibration signal which is the output of a pick-up attached to the Anderon meter.[1] This method, however, has its drawbacks; it requires a great deal of time to train a person to be a good inspector. In addition, an inspector in poor health, physical or otherwise, may make mistakes in the detection of defects.

Accordingly, because of these limitations, several methods have been proposed for the automatic detection and classification of the defects in ball bearings.[1-8] The defects in bearings are detected by these present methods when the flaws generate periodical vibration signals, since the period is calculated from the shape of a bearing. However, the periodicity of the vibration is not always guaranteed when (a) there are slight flaws on the surface of the race, and (b) when flaws are on the surface of the balls. When there are slight flaws on the surface of the race, the ball does not always come into contact with the flaws since the pathway on which the balls have contact with the race changes every moment. In the case of flaws on the surface of balls, the race does not always contact with the flaw on the ball since the balls do not always roll regularly. Another problem arises if dust is mixed in the grease: the periodicity of the resonant vibration driven by the dust is not also guaranteed. Therefore, we have developed a new diagnostic method for the automatic detection of the resonant vibration due to flaws on the surface or dust in the

grease. This method is based on the assumption that the signal to be analyzed is a non-periodic one.

In this paper, we first discuss the problems inherent in the conventional methods which are applied to the detection of the resonant vibrations due to defects. Next, we propose a new method, by which the driving pulse sequence $x(n)$ due to the defects is estimated, and it is found that the amplitude of the estimated impulse sequence $x(n)$ is in proportion to the 6/5-th power of the mechanical size of surface defects. Finally, it is proved by our experiments that the defects can be detected with the accuracy rate of 98%.

2. APPARATUS FOR THE EXPERIMENTS

Figure 1 shows a block diagram of the experimental system in which the vibration signal is picked up and recorded into a digital audio tape to carry back to the laboratory. The inner ring rotates at a constant speed of 1,800 rpm, and the outer one is fixed by imposing an axial pressure, where the rotation speed and the amount of pressure are the same as those actually used in a video tape recorder. Under such the conditions, a flaw causes a radial movement of the outer ring and the signal resulting from the movement is picked up by a vibration pick-up attached to the outer ring. The signal is amplified and filtered through a high-pass filter to eliminate the primary frequency component (=30 Hz) corresponding to the rotation of the inner ring. The filtered signal is stored in a high-fidelity digital audio tape. In the laboratory, the signal is played back and is A/D converted with a 12 bits A/D converter at a sampling period of 15 μs or 30 μs.

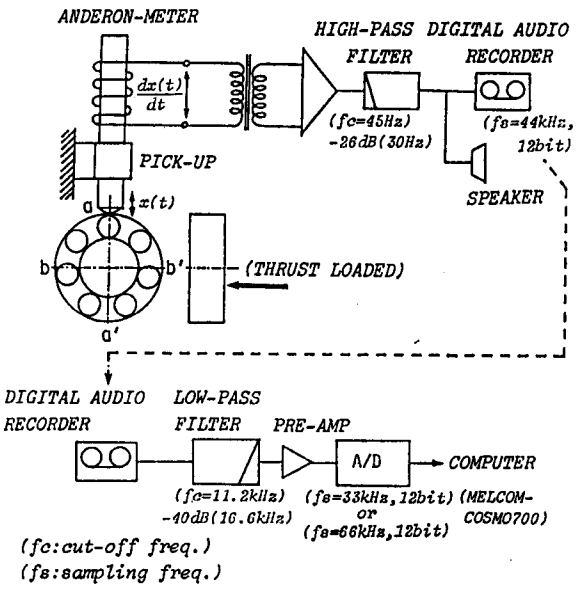

Fig. 1 A block diagram of the procedure for measuring the vibration signal of a ball bearing.

3. DIFFICULTIES IN THE CONVENTIONAL METHODS

3.1 Application of the Linear Predictive Analysis

The linear predictive analysis has been applied to the analysis of an exponentially decaying sinusoidal signal.[4] To detect the resonant vibrations due to defects by this method, the following three assumptions must be satisfied:

(a) The mechanical excitation caused by a defect is expressed by a unit impulse, where

$$\delta(n)=\begin{cases} 1 & (n=0) \\ 0 & (n\neq 0). \end{cases}$$

(b) The repetition interval of a pulse sequence $x(n)$ of mechanical excitations is so long that the response to the previous excitation impulse sufficiently decreases before the next excitation.

(c) The transfer function from the mechanical excitation to the resonant vibration, which is detected by the pick-up on the outer ring, is represented by an all pole model $1/A(z)$.

If the above three assumptions are satisfied, all pole model $1/A(z)$ can be identified by using the linear predictive analysis. In addition, the pulse sequence $x(n)$ of mechanical excitations can be obtained by using the inverse filter $A(z)$.

However, when the configuration of flaws on the surface of the race is complicated, the mechanical excitation is expressed by multi-pulses and the second assumption is not satisfied. Thus, the linear predictive analysis cannot be applied to the vibration analysis in such a case.

3.2 Use of a Band-Pass Filter

The signal-to-noise ratio can be improved by extracting only the frequency components around the central frequency of the resonant vibration using a band-pass filter. However, the use of a band-pass filter involves the following two problems:

(a) The amplitude of a band-passed signal does not correspond to the mechanical size of defects as described later in Chapter 5.

(b) It is difficult to select a single narrow-fre-

quency-band with which the defects are successfully detected. The reason is described in detail here. The vibration signal $y(t)$ obtained by the pick-up is expressed by the convolution between the resonant vibration $h(t)$ and the pulse sequence $x(t)$ of mechanical excitations as follows:

$$y(t)=h(t)*x(t), \tag{1}$$

where $*$ denotes the convolution. This relation can be expressed in the frequency domain as follows:

$$Y(\omega)=H(\omega)\cdot X(\omega), \tag{1'}$$

where the $Y(\omega)$, $H(\omega)$ and $X(\omega)$ are the Fourier transform of $y(t)$, $h(t)$ and $x(t)$, respectively. Since the characteristics of $h(t)$ and $H(\omega)$ are determined both by the size and by the material of the outer ring, the bearings with the same size and the same material show the same resonant characteristics. On the other hand, the spectrum $X(\omega)$ of the exciting signal $x(t)$ is determined by the configuration of each defect. A broad frequency band is necessary to detect the exciting signal because the exciting signal is composed of pulses of very short duration.

Figure 2 shows the characteristics of a bearing having two kinds of defects. Figures 2 (a-1) and 2 (b-1) show the vibration signals $y_1(n)$ and $y_2(n)$ due to the defects whose surface cross-sectional views are shown in Figs. 2 (a-3) and 2 (b-3), respectively. The power spectra of the vibration signals $y_1(n)$ and $y_2(n)$ are shown in Figs. 2 (a-2) and 2 (b-2), respectievly. Since the characteristics of the resonant vibration $h(n)$ is the same for both $y_1(n)$ and $y_2(n)$, the difference between $y_1(n)$ and $y_2(n)$ depends on the configuration of the flaws. It is difficult to select a single narrow-frequency-band with which a high signal-to-noise ratio is obtained because the frequency spectrum spread onto a broad frequency band as shown in Figs. 2 (a-2) and 2 (b-2). The difference between these frequency spectra is caused by the difference between the excitation pulse train generated by the flaws on the surface shown in Fig. 2 (a-3) and 2 (b-3). Therefore, an improvement in the detection of defects by an improvement in signal-to-noise ratio cannot always be expected by simply passing the signal through a band-pass filter.

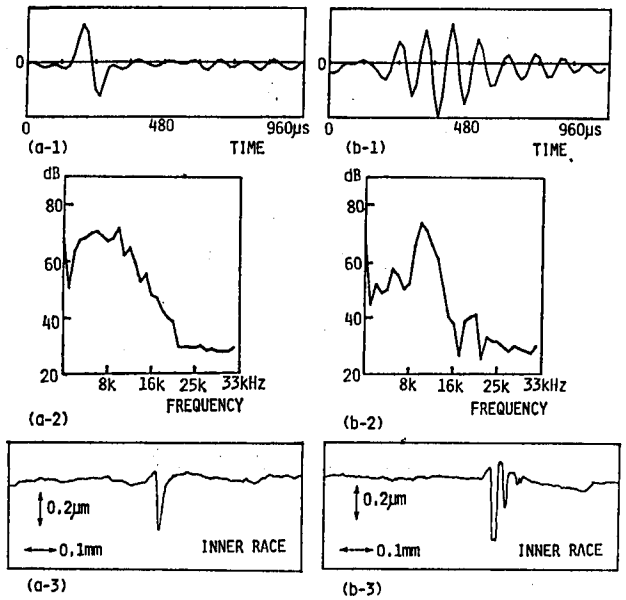

Fig. 2 Two vibration signals due to two flaws on the inner ring of the same sample. (1) vibration wave $y_i(n)$, $i=1, 2$, (2) power spectrum, (3) profile of the surface measurement (cross-sectional view).

4. ESTIMATION OF POLES BY USING TWO-PULSE MODEL

We propose a new method of estimating the pulse sequence $x(n)$ excited by the surface defects using the vibration signal $y(n)$. If the resonant vibration $h(n)$ has been already estimated, the sequence $x(n)$ can be calculated from the vibration signal $y(n)$ since $y(n)$ is expressed by the convolution of the sequence $x(n)$ and the resonant vibration $h(n)$.

Therefore, $h(n)$ should be estimated first. This chapter describes the method of identifying the resonant vibration $h(n)$. By this method, $h(n)$ is estimated from the vibration signal $y(n)$ excited by the flaw with a simple configuration as shown in Fig. 2 (a-3).

Using $h(n)$ thus obtained, the pulse sequence $x(n)$ is estimated by inverse-filtering the vibration signal $y(n)$ using $h(n)^{-1}$, where

$$h(n)^{-1}*h(n)=\delta(n). \tag{2}$$

The estimated sequence $x(n)$ is found to correspond to the mechanical size of surface defects as described later in Chapter 5.

The surface defects of the ball bearing are diagnosed using the amplitude of the sequence $x(n)$.

4.1 Estimation of the Resonant Vibration $h(n)$ Using the Two-Pulse Model

Figure 2 (a-1) shows a part of the vibration signal $y(n)$ caused by a flaw with a simple configuration on the inner ring. The cross-sectional view of the flaw is shown in Fig. 2 (a-3). This part of the vibra-

tion $y(n)$ is represented by the sum of two resonant vibrations driven by two impulses $\delta(n-\tau_1)$ and $\delta(n-\tau_2)$. The two impulses are excited when a ball comes into contact with the flaw on the inner ring and hits the inner ring and the outer ring, respectively. As reported in a previous investigation,[3] the resonant vibration $h(n)$ is expressed by an all pole model. Therefore, the modeled vibration signal $\hat{y}(n)$ driven by the two pulses is expressed as a linear combination of M complex exponentials as follows:

$$\begin{aligned}\hat{y}(n)=&\sum_{m=1}^{M}(u_{m1}\cdot z_m{}^{(n-\tau_1)}\\&+u_{m1}{}^{*}\cdot z_m{}^{*(n-\tau_1)})\cdot U(n-\tau_1)\\&+\sum_{m=1}^{M}(u_{m2}\cdot z_m{}^{(n-\tau_2)}\\&+u_{m2}{}^{*}\cdot z_m{}^{*(n-\tau_2)})\cdot U(n-\tau_2)+b,\end{aligned}\qquad(3)$$

where u_{m1} and u_{m2} are complex coefficients, $u_{mk}{}^{*}$ the complex conjugate of u_{mk}, $U(n)$ a unit step function, 'b' a real constant (representing the bias term), and z_m $(m=1,2,\ldots,M)$ the roots of the polynomial in the denominator of the all pole model as follows:

$$A(z_m)=0\,,\qquad \text{for } m=1,2,\ldots,M,\qquad(4)$$

where $1/A(z)$ is the z-transform of the resonant vibration $h(n)$. From Eq. (3), the ideal response $\tilde{y}(n)=\hat{y}(n)-b$ is expressed as follows:

$$\tilde{y}(n)=\begin{cases}\displaystyle\sum_{m=1}^{M}\{u_{m1}{}'\cdot z_m{}^{(n-\tau_1)}+u_{m1}{}'{}^{*}\cdot z_m{}^{*(n-\tau_1)}\},\\\qquad(\text{for } \tau_1\le n<\tau_2)\\\displaystyle\sum_{m=1}^{M}\{u_{m2}{}'\cdot z_m{}^{(n-\tau_2)}+u_{m2}{}'{}^{*}\cdot z_m{}^{*(n-\tau_2)}\},\\\qquad(\text{for } \tau_2\le n)\end{cases}\qquad(5)$$

where $u_{m1}{}'=u_{m1}$ and $u_{m2}{}'=u_{m1}\cdot z_m{}^{(\tau_2-\tau_1)}+u_{m2}$. The unknown parameters u_{mk} $(m=1,2,\ldots,M;\ k=1,2)$, b, and z_m $(m=1,2,\ldots,M)$ are to be determined, and the difficulty lies in the fact that the equations are nonlinear because of the z's. This difficulty can be minimized by Prony's method,[4,5] and these unknown parameters can be obtained by solving a set of simultaneous equations as described below. Since the period of the transient time of the vibration excited by a single pulse is sufficiently short, the optimum value of b is also identified by solving the same simultaneous equations.

Since z_m $(m=1,2,\ldots,M)$ are the roots of the equation $A(z)=0$ in Eq. (4), $A(z)$ is expressed as follows:

$$\begin{aligned}A(z)=&\prod_{m=1}^{M}(1-z_m z^{-1})(1-z_m{}^{*}z^{-1})\\=&\sum_{j=0}^{2M}a_j\cdot z^{-j}\qquad(a_0=1)\\=&0.\end{aligned}$$

$$(\text{for } z=z_m,\ z_m{}^{*};\ m=1,2,\ldots,M)\qquad(6)$$

In order to determine the coefficients a_j $(j=1,2,\ldots,2M)$, by multiplying $\tilde{y}(\tau_k+2M)$ in Eq. (5) by a_0, $\tilde{y}(\tau_k+2M-1)$ by a_1, $\tilde{y}(\tau_k+2M-2)$ by a_2, $\ldots$, $\tilde{y}(\tau_k+1)$ by a_{2M-1}, and $\tilde{y}(\tau_k)$ by a_{2M}, and by adding the results, the following equation is obtained:

$$\begin{aligned}&\sum_{j=0}^{2M}a_j\cdot\tilde{y}(\tau_k+2M-j)\\&\quad=\sum_{j=0}^{2M}a_j\sum_{m=1}^{M}\{u_{mk}{}'\cdot z_m{}^{(2M-j)}\\&\qquad+u_{mk}{}'{}^{*}\cdot z_m{}^{*(2M-j)}\}\\&\quad=\sum_{m=1}^{M}\left\{u_{mk}{}'\left(\sum_{j=0}^{2M}a_j\cdot z_m{}^{-j}\right)z_m{}^{2M}\right.\\&\qquad\left.+u_{mk}{}'{}^{*}\left(\sum_{j=0}^{2M}a_j\cdot z_m{}^{*-j}\right)z_m{}^{*2M}\right\}.\end{aligned}$$

$$(\text{for } k=1,2)\qquad(7)$$

Since z_m $(m=1,2,\ldots,2M)$ satisfy Eq. (6), the right hand side of Eq. (7) is equal to zero and then for the period $n\ge\tau_k+2M$, $(k=1,2)$, $\tilde{y}(n)$ and $\hat{y}(n)=\tilde{y}(n)+b$ are represented as follows:

$$\tilde{y}(n)=-\sum_{j=1}^{2M}a_j\cdot\tilde{y}(n-j),$$

and

$$\hat{y}(n)-b=-\sum_{j=1}^{2M}a_j\cdot\{\hat{y}(n-j)-b\}.$$

$$(\text{for } n\ge\tau_k+2M;\ k=1,2)\qquad(8)$$

In a practical case, the observed signal $y(n)$ must be used instead of $\hat{y}(n)$ in Eq. (8). Then, $y(n)$ is represented as follows:

$$y(n)-b=-\sum_{j=1}^{2M}a_j\cdot\{y(n-j)-b\}+e(n),$$

$$(\text{for } n\ge\tau_k+2M;\ k=1,2)\qquad(9)$$

where $e(n)$ is the residual error. If the above model shown in Eq. (3) just represents the vibration signal $y(n)$, the error $e(n)$ are equal to zero for the periods $\tau_1+2M\le n<\tau_2$ and $\tau_2+2M\le n$. Therefore, the coefficients $\{a_j\}$, $(j=1,2,\ldots,M)$ are obtained by solving Eq. (9) so that the following squared error $\alpha(\tau_1,\tau_2)$

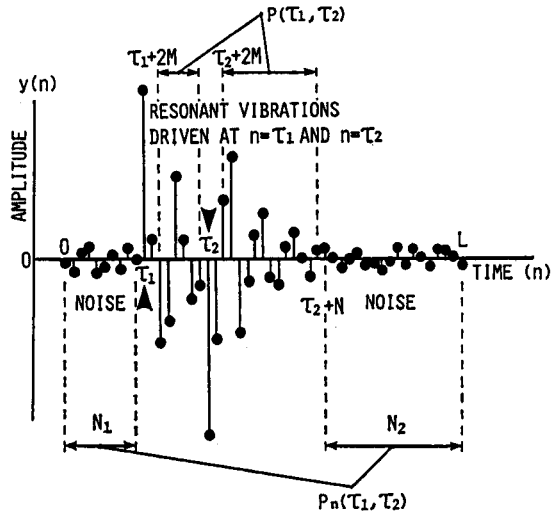

Fig. 3 Illustration showing samples used in the squared error for various lags τ_1 and τ_2 in Eqs. (10) and (17).

takes the minimum value (See Fig. 3).

$$\begin{aligned}\alpha(\tau_1,\tau_2)&=\sum_{n\in P(\tau_1,\tau_2)} e(n)^2\\&=\sum_{n\in P(\tau_1,\tau_2)}\left[\sum_{j=0}^{2M} a_j\cdot\{y(n-j)-b\}\right]^2\\&=\sum_{n\in P(\tau_1,\tau_2)}\left[\sum_{j=0}^{2M}\sum_{k=0}^{2M} a_j a_k\cdot y(n-j)y(n-k)\right.\\&\quad -2\cdot b\cdot\sum_{j=0}^{2M}\sum_{k=0}^{2M} a_j a_k\cdot y(n-j)\\&\quad \left.+b^2\cdot\sum_{j=0}^{2M}\sum_{k=0}^{2M} a_j a_k\right],\end{aligned}\tag{10}$$

where

$$P(\tau_1,\tau_2)=\{n|\ (\tau_1+2M)\le n<\min(\tau_2,\tau_1+N)\ \text{and}\ (\tau_2+2M)\le n<(\tau_2+N)\},$$

N is the length of the time window in which the squared error is calculated, and $\min(p,q)$ represents the minimum value, p or q. N is set so that the following conditions are satisfied: $N>2M$ and $N<L$, where L is the interval from one mechanical excitation to the next one. Here, the following three functions are introduced:

$$C_{jk}(\tau_1,\tau_2)=\sum_{n\in P(\tau_1,\tau_2)} y(n-j)\cdot y(n-k),$$
$$(\text{for } j=0,1,\ldots,2M;\ k=0,1,\ldots,2M)$$

$$D_j(\tau_1,\tau_2)=\sum_{n\in P(\tau_1,\tau_2)} y(n-j),$$
$$(\text{for } j=0,1,\ldots,2M)$$

and

$$E_0(\tau_1,\tau_2)=\sum_{n\in P(\tau_1,\tau_2)} 1.\tag{11}$$

Using these functions, Eq. (10) is expressed as follows:

$$\begin{aligned}\alpha(\tau_1,\tau_2)&=\sum_{j=0}^{2M}\sum_{k=0}^{2M} a_j\cdot a_k\cdot C_{jk}(\tau_1,\tau_2)\\&\quad -2\cdot b\cdot\sum_{j=0}^{2M}\sum_{k=0}^{2M} a_j a_k\cdot D_j(\tau_1,\tau_2)\\&\quad +b^2\cdot\sum_{j=0}^{2M}\sum_{k=0}^{2M} a_j a_k\cdot E_0(\tau_1,\tau_2).\end{aligned}\tag{12}$$

Then, minimization of $\alpha(\tau_1,\tau_2)$ is obtained by setting the partial derivatives of $\alpha(\tau_1,\tau_2)$ to zero[4] with respect to both $\{a_j\}$ $(j=1,2,\ldots,2M)$ and b as follows:

$$\begin{aligned}\frac{\partial\alpha(\tau_1,\tau_2)}{\partial a_j}&=0\qquad(\text{for } j=1,2,\ldots,2M)\\&=2\cdot\sum_{k=0}^{2M} a_k\cdot C_{jk}(\tau_1,\tau_2)\\&\quad -2\cdot b\cdot\left(\sum_{k=0}^{2M} a_k\right)\cdot D_j(\tau_1,\tau_2)\\&\quad +2\cdot b^2\cdot\left(\sum_{k=0}^{2M} a_k\right)\cdot E_0(\tau_1,\tau_2).\end{aligned}\tag{13}$$

$$\begin{aligned}\frac{\partial\alpha(\tau_1,\tau_2)}{\partial b}&=0\\&=\left(\sum_{k=0}^{2M} a_k\right)\cdot\left\{-2\cdot\sum_{m=0}^{2M} a_m\cdot D_m(\tau_1,\tau_2)\right.\\&\quad \left.+2\cdot b\cdot\left(\sum_{m=0}^{2M} a_m\right)\cdot E_0(\tau_1,\tau_2)\right\}.\end{aligned}\tag{14}$$

Here, since bias is very small in real applications, b^2 can be negligible in Eq. (13). Let the variable d:

$$d=-b\cdot\left(\sum_{m=0}^{2M} a_m\right),\tag{15}$$

and since $a_0=1$, the set of $(2M+1)$ linear simultaneous equations are obtained as:

$$\begin{aligned}&\sum_{k=1}^{2M} a_k\cdot C_{jk}(\tau_1,\tau_2)+d\cdot D_j(\tau_1,\tau_2)\\&\quad=-C_{j0}(\tau_1,\tau_2),\quad(\text{for } j=1,2,\ldots,2M)\end{aligned}$$

and

$$\begin{aligned}&\sum_{m=1}^{2M} a_m\cdot D_m(\tau_1,\tau_2)+d\cdot E_0(\tau_1,\tau_2)\\&\quad=-D_0(\tau_1,\tau_2).\end{aligned}\tag{16}$$

The unknown prediction coefficients $\{a_j\}$ and d (or b) are obtained by solving this set of simultaneous

equations. When the vibration over the period $0 \leq n \leq L$ is approximated by the modeled vibration signal driven by the two pulses at $n=\tau_1$ and $n=\tau_2$ in Eq. (3), the following total squared error $\alpha_T(\tau_1, \tau_2)$ is calculated for various lags of τ_1 and τ_2 within the period $(0 \leq n \leq L)$, where there is a resonant vibration due to the presence of a flaw:

$$\alpha_T(\tau_1, \tau_2)=\alpha(\tau_1, \tau_2)+\sum_{n \in P_n(\tau_1, \tau_2)} y(n)^2, \qquad (17)$$

where

$$P_n(\tau_1, \tau_2)=\{n|\ 0 \leq n<\tau_1 \text{ and } (\tau_2+N) \leq n \leq L\},$$

(See Fig. 3). If N is so large that the modeled vibration $y(n)$ sufficiently decrease at $n=\tau_2+N$, the second term of the right hand side of Eq. (17) denotes the power of the background noise over the periods N_1 and N_2 in Fig. 3. In an ideal case, this term is also equal to zero. Optimum lags of $\tau_{1\,\min}$ and $\tau_{2\,\min}$ are those which are obtained when the squared error $\alpha_T(\tau_{1\,\min}, \tau_{2\,\min})$ takes the minimum value.

By using the predictive coefficients $\{a_{j\,\min}\}$, which are calculated by Eq. (16) in the case where the lags (τ_1, τ_2) is equal to $(\tau_{1\,\min}, \tau_{2\,\min})$, the following polynomial $A(z)$ is defined:

$$A(z)=\sum_{j=0}^{2M} a_{j\,\min} \cdot z^{-j}. \qquad (a_{0\,\min}=1) \qquad (18)$$

The poles $\{z_m, z_m^*\}$, $(m=1, 2, \ldots, M)$ of the all pole model are obtained by solving the polynomial $A(z)$. By using $\{z_m\}$, b, $\tau_{1\,\min}$ and $\tau_{2\,\min}$, the vibration signal $y(n)$ actually observed is described as follows:

$$\begin{aligned} y(n)=&\sum_{m=1}^{M} (u_{m1} \cdot z_m^{(n-\tau_{1\,\min})} \\ &+u_{m1}^* \cdot z_m^{*(n-\tau_{1\,\min})}) \cdot U(n-\tau_{1\,\min}) \\ &+\sum_{m=1}^{M} (u_{m2} \cdot z_m^{(n-\tau_{2\,\min})} \\ &+u_{m2}^* \cdot z_m^{*(n-\tau_{2\,\min})}) \cdot U(n-\tau_{2\,\min}) \\ &+b+\eta(n), \quad (\text{for } \tau_1 \leq n<\tau_2+N) \end{aligned} \qquad (19)$$

where $\eta(n)$, $(\tau_1 \leq n<\tau_2+N)$ represents the residual term. Since $\{z_m\}$, b, $\tau_{1\,\min}$ and $\tau_{2\,\min}$ have already been obtained by Eq. (16), the modeled vibration signal $\hat{y}(n)$ can be obtained from Eq. (3). Here, the error function:

$$\begin{aligned} \beta &= \sum_{n \in P'(\tau_{1\,\min}, \tau_{2\,\min})} \eta(n)^2 \\ &= \sum_{n \in P'(\tau_{1\,\min}, \tau_{2\,\min})} \{y(n)-\hat{y}(n)\}^2, \end{aligned} \qquad (20)$$

where

$$P'(\tau_1, \tau_2)=\{n|\ \tau_1 \leq n<\min(\tau_2, \tau_1+N) \text{ and } \tau_2 \leq n<(\tau_2+N)\},$$

is introduced. The estimation of the complex coefficients $\{u_{m1}, u_{m2}\}$, $(m=1, 2, \ldots, M)$ are those which are obtained when the error function β takes the minimum value.[4,5]

4.2 Results Obtained by Using the Two-Pulse Model

Figure 4 (b-1) shows the wave form of the modeled vibration signal $\hat{y}(n)$, which is obtained by expressing the original vibration signal $y(n)$ shown in Fig. 4 (a) by the two-pulse model where the number $2M$ of poles is set to 2. Figures 4 (b-2) and 4 (b-3) show the positions of the identified poles: z_0 and z_0^* in the z-plane and the complex coefficients: (u_1, u_2), respectively. Figure 4 (c) shows the modeled vibration signal $\hat{y}'(n)$ obtained by expressing the original

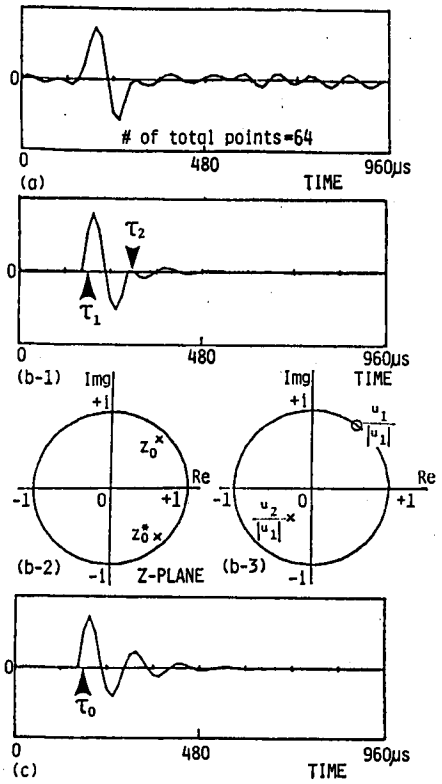

Fig. 4 Experimental results using the two-pulse model. (a) the vibration signal $y(n)$ due to a flaw (see also Fig. 2(a-1)), (b-1) predicted vibration signal $\hat{y}(n)$ by the two-pulse model, (b-2) poles z_0 and z_0^* in the z-plane, (b-3) initial phase coefficients u_1 and u_2 in the z-plane, (c) predicted vibration signal $\hat{y}'(n)$ by the one-pulse model.

vibration signal $y(n)$ by a one-pulse model:

$$\hat{y}'(n)=(u\cdot z_0{}^{(n-\tau_0)} + u^*\cdot z_0{}^{*(n-\tau_0)})\cdot U(n-\tau_0)+b, \quad (21)$$

where u is the complex coefficient, and τ_0 the time when the vibration is excited. It is found from these experimental results that the original vibration signal $y(n)$ is well matched with the modeled signal $\hat{y}(n)$ expressed by the two-pulse model. The position of the two complex coefficients u_1 and u_2 are different by about π radian in phase. That is, the directions of the driving pulsive forces are opposite each other. We consider that the first pulsive force is excited when a ball collides into the inner ring after the ball rotates over the crack on the inner race, and the second force is excited when the ball collides into the outer ring (See Fig. 9 in Chap. 5). That is, the first pulse is positive and the second pulse is negative in this case.

Twenty-seven vibration signals $y(n)$ ($n=0, 1, \ldots, 63$) are obtained from a bearing sample so that every vibration signal $y(n)$ involves a response to a pulse excitation due to a flaw. Figures 5 (a) and 5 (b) show the distribution of the position of the estimated poles and the complex coefficients obtained from these 27 vibration signals. Since the position of these poles are almost the same as shown in Fig. 5 (a), it is clear that the parameters are consistently obtained accurately by using the two-pulse model. Figure 6 shows the position of the average pole $\tilde{z}_0$ and the prediction coefficients $\{a_j\}$ calculated from the average pole $\tilde{z}_0$, where each vibration signal is A/D converted at a sampling period of 15 μs.

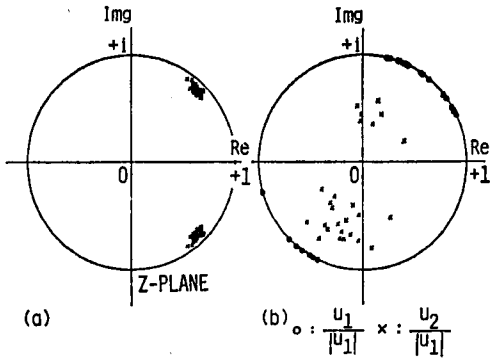

Fig. 5 Experimental results obtained by the two-pulse model using 27 vibration signals $y(n)$ obtained from a bearing sample. (a) the distribution of the estimated poles z_0 and z_0^*, (b) the distribution of the initial phase coefficients u_1 and u_2.

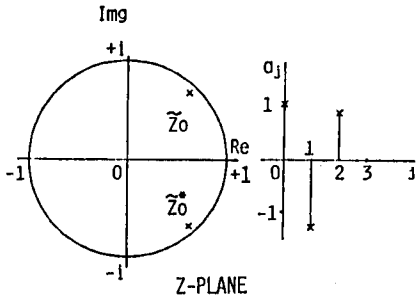

Fig. 6 The average pole $\tilde{z}_0$ and its predictive coefficients $\{a_j\}$.

4.3 Residual Power of the Two-Pulse Model

The residual error $\eta(n)$ defined in Eq. (19) is expressed by the sum of two uncorrelated components: the noise $n(n)$ and the error $\varepsilon(n; 2M)$ which occurs by using the $2M$-order two-pulse model as follows:

$$\eta(n)=\varepsilon(n; 2M)+n(n). \quad (22)$$

Figure 7 (a) shows the power distribution of the same 27 vibration signals $y(n)$ used in Fig. 5. Figure 7 (b) shows the power distribution of the residual errors $\eta(n)$ obtained from the same vibration signals $y(n)$ used in Fig. 7 (a). Other 27 vibration signals are obtained from the same bearing sample used in Fig. 5 so that these signals do not involve any responses to pulse excitations due to a flaw. Figure 7 (c) shows the power distribution of these 27 signals, which shows almost the same distribution as that of the residual errors $\eta(n)$ shown in Fig. 7 (b). This means that the original vibration signals are well matched by the two-pulse model (order: $2M=2$).

4.4 Inverse Filtering

Since the z-transform of the resonant vibration $h(n)$ is expressed by an all pole model, z-transform of the ideal response $\tilde{y}(n)=\hat{y}(n)-b$ is expressed as follows:

$$\tilde{Y}(z)=X(z)/A(z), \quad (23)$$

where $\tilde{Y}(z)$, $X(z)$ and $1/A(z)$ represent the z-transform of $\tilde{y}(n)$, $x(n)$ and $h(n)$, respectively. By multiplying both sides of Eq. (23) by $A(z)$, the following equation is obtained:

$$X(z)=\tilde{Y}(z)A(z). \quad (24)$$

In the time domain, Eq. (24) can be equivalently

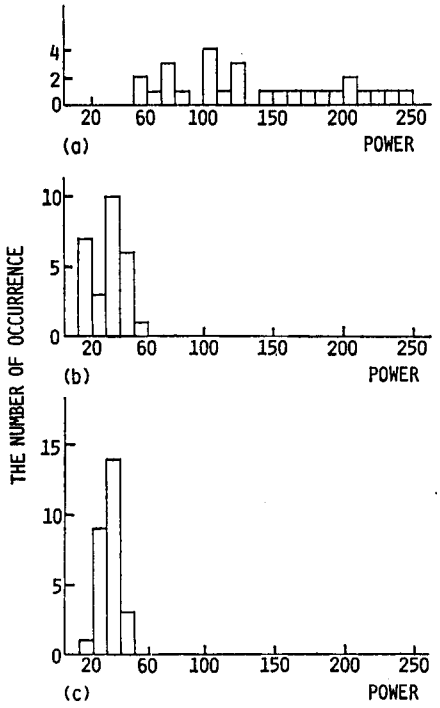

Fig. 7 Residual powers of the two-pulse model. (a) the power distribution of the 27 vibration signals $y(n)$ involving the responses to a flaw, (b) the power distribution of residual errors $\eta(n)$ for the same vibration signals used in (a), (c) the power distribution of other 27 vibration signals which do not involve any responses to a flaw.

written as follows:

$$x(n)=\sum_{j=0}^{2M} a_j \cdot \tilde{y}(n-j)$$
$$=\sum_{j=0}^{2M} a_j \cdot (\hat{y}(n-j)-b). \qquad (25)$$

Using the predictive coefficients $\{a_j\}$ calculated by the average poles $\tilde{z}_0$ as shown in Fig. 6, an inverse filter $h(n)^{-1}$ is obtained by letting $h(j)^{-1}$ be equal to a_j. By convoluting $h(n)^{-1}$ with the vibration signal $y(n)$, the driving pulse sequence $x(n)$ is estimated. Since the resonant vibration is determined only by the size and the material of the bearing parts, the inverse filter $h(n)^{-1}$ obtained above is the same as those of the other bearings with the same sizes and the same materials.

Figure 8 shows an estimated driving pulse sequence $x(n)$ obtained by the convolution between the inverse-filter $h(n)^{-1}$ and the vibration signal $y(n)$ shown in Fig. 4 (a). Two main impulses of opposite signs are observed in this figure.

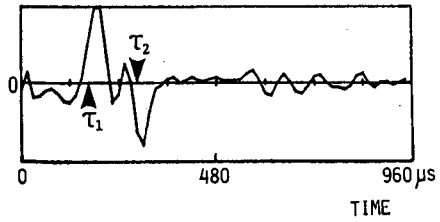

Fig. 8 Estimated driving pulse sequence $x(n)$ obtained by inverse-filtering. The amplitude of $x(n)$ is normalized by the maximum value.

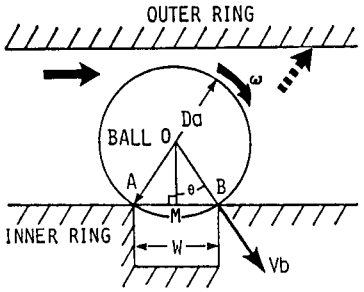

Fig. 9 The collision model of a ball with a flaw on the inner race.

5. ESTIMATION OF THE FLAW SIZE

In this chapter, we describe the relation between the size of a flaw and the estimated impulse sequence $x(n)$. The collision of a ball with the flaws on the race is described by the model[6] shown in Fig. 9. When a revolving ball arrives at the edge A of the flaw, the ball rotates around the fulcrum A, and the ball hits the other edge B of the flaw. The velocity V_b of the collision at the edge B is represented as follows[6]:

$$V_b=\omega \cdot W \cos\theta, \qquad (26)$$

where ω denotes the angular velocity of the ball rotation around the fulcrum A, W the width of the flaw and θ the angle between the lines $\overline{OM}$ and $\overline{OB}$. Since the width W of the flaw is usually much smaller than the diameter D_a of the ball, $\cos\theta$ is approximated to be 1, and the velocity V_b is expressed as follows:

$$V_b \simeq \omega \cdot W. \qquad (27)$$

Since the inner ring rotates at a constant speed, the angular velocity ω is constant.[6] Therefore, the velocity V_b of the collision at the edge B is in proportion to the width W of the flaw. It is known that the excit-

ing force F at the collision is in proportion to the 6/5-th power of the velocity V_b [6,7] as follows:

$$F \propto V_b^{6/5}. \tag{28}$$

The amplitude of the accelerating vibration A_b driven by the exciting force F is in proportion to the 6/5-th power of the velocity V_b.[6]

Since the vibration signal $y(t)$ used in this paper is picked-up by a velocity-type sensor, the following equation is obtained:

$$y(t)=\int F\,dt \propto W^{6/5}. \tag{29}$$

Thus, the amplitude of the estimated impulse sequence $x(n)$ is in proportion to the 6/5-th power of the width W of the flaw as well as the vibration signal $y(n)$. Therefore, the size of a flaw can be evaluated by the amplitude of the vibration $x(n)$. Hence, the ball bearing is sorted by the width of a flaw.

6. SORTING OF A BALL BEARING

There are the following two problems in sorting a bearing by our proposed method: (a) Since the inverse filtering makes the power spectrum of the estimated impulse sequence $x(n)$ flat, the signal to the noise ratio decreases when all of the frequency components of the vibration signal $y(n)$ is used to estimate the impulse sequence $x(n)$; (b) There are some fluctuations in the gain of the amplifier.

To solve the above two problems, the following processes are carried out in the following sequence:

(1) The impulse sequence $x(n)$ is estimated by the proposed method described previously in Chapter 4.

(2) By using N_B filters, N_B narrow-band signals are calculated from the impulse sequence $x(n)$, where N_B is the number of band pass filters. The N_B filters are realized in our experiment by $2N_B$ $(=16)$ point fast Fourier transform. The band widths of all filters are 4 kHz. By calculating the squared values of the narrow band signals, N_B power sequences $z(n;i)$, $(i=1, 2, \ldots, N_B)$ are obtained.[3]

(3) The average powers r_i, $(i=1,2,\ldots,N_B)$ of the narrow band signals $z(n;i)$ are calculated, and the normalized narrow band power signals $z'(n;i) = z(n;i)/r_i$ are calculated.

(4) Normalized signals $z'(n;i)$ are compared with thresholds $\{T_i\}$, $(i=1, 2, \ldots, N_B)$ to sort the ball bearing.

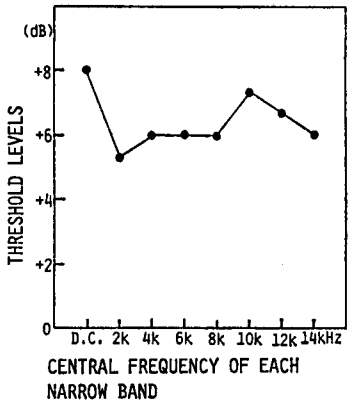

Fig. 10 Thresholds $\{T_i\}$ for detecting the flaws or dust.

The thresholds $\{T_i\}$ are determined as follows: 343 bearings were sorted aurally by a highly qualified inspector into 6 categories shown in Table 1: normal, flaw on the inner ring, flaw on the outer ring, flaw on the ball, dust, and borderline cases where he could not determine if there was a defect. The vibration signal was A/D converted at the sampling period of 30 μs, and the data length used for the sorting was 0.98 s. Thresholds $\{T_i\}$, $(i=1,2,\ldots,N_B)$ were experimentally set to be the maximum values of the narrow band powers $z'(n;i)$ of 84 normal ball bearings. Figure 10 shows the threshold level $\{T_i\}$ for various frequency bands.

Table 1 shows the results of the sorting, where the samples on the border are not counted when obtaining the recognition rate. The number of correctly sorted samples and that of incorrectly sorted samples are 314 and 5, respectively. The ball bearings are sorted at the correct rate of 98.4%. Two of the incorrectly sorted samples had dust, and their vibration signals used for the analysis did not involve the responses to the defects. The other three incorrectly sorted samples had many flaws on the outer race. Since the vibrations were driven almost always by the flaws on the rough race, the average powers $\{r_i\}$ were large, and the resonant vibration could not be detected by the procedure described above. These samples can be, however, classified by using our previously proposed method.[8]

Figure 11 shows the rate of detecting each kind of inferior samples by the eight narrow bands. Almost all flaws could be detected by using any frequency band from D.C. to 6 kHz, and the dust could be detected by using slightly higher frequency bands from D.C. to 8 kHz.

Table 1 Recognition results of ball bearings.

In	Out		Total
	The number of samples recognized to be normal	The number of samples recognized to be inferior	
Normal	84	0	84
Samples on the border	8	16	24
Dust	2	36	38
Flaw on the inner ring	0	126	126
Flaw on the outer ring	3	38	41
Flaw on the ball	0	30	30

Total number of used samples: 343, recognition rate: 98.4%.

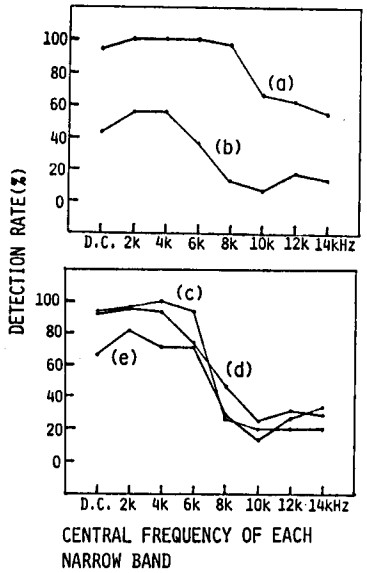

Fig. 11 The rate of detecting the following kinds of inferior samples for the narrow bands. (a) dust, (b) samples on the boarder, (c) flaw on the ball, (d) flaw on the inner ring, (e) flaw on the outer ring.

7. CONCLUSIONS

A new method is proposed for the automatic detection of slight flaws or dust in ball bearings using vibration signals. These defects, which are not detected by periodic analysis or by ordinary signal processing techniques, can be easily detected by using our new method with a 98.4% accuracy rate.

REFERENCES

1) T. Igarashi, "Sound of rolling bearings," Lubrication **22**, 751–756 (1977) (in Japanese).
2) S. Braun and B. Datner, "Analysis of roller/ball bearing vibrations," Trans. Am. Soc. Mech. Eng., J. Mech. Des. **101**, 118–125 (1979).
3) H. Kanai, M. Abe, and K. Kido, "Detection and discrimination of flaws in ball bearings by vibration analysis," J. Acoust. Soc. Jpn. (E) **7**, 121–131 (1986).
4) J. D. Markel and A. H. Gray, Jr., *Linear Prediction of Speech* (Springer-Verlag, Berlin, 1976), Chaps. 1–3.
5) F. B. Hildebrand, *Introduction to Numerical Analysis* (McGraw-Hill, New York, 1956), Chap. 9.
6) M. Noda, "A study of the inferior foresight in roller/ball bearing," Proc. Spring Meet. Lubr. Soc. Jpn. **25**, 125–128 (1981) (in Japanese).
7) G. Nishimura and K. Takahashi, "Ball bearing noise," J. Jpn. Soc. Precis. Eng. **30**, 475–489 (1964) (in Japanese).
8) H. Kanai, M. Abe, and K. Kido, "Detection and discrimination of roughness of race in ball bearings by vibration analysis," Proc. Autumn Meet. Acoust. Soc. Jpn., 287–288 (1984) (in Japanese).
9) A. V. Oppenheim and R. W. Schafer, *Digital Signal Processing* (Prentice-Hall Englewood Cliffs, 1975).