Dynamic Diagnosis of Railway Tracks by Means of the Karhunen–Loève Transformation *

UTE FELDMANN, EDWIN KREUZER, and FERNANDO PINTO
Ocean Engineering Section II & Mechanics,
Technical University of Hamburg–Harburg, Germany

Abstract. Maintenance of the quality of the railway track is crucial for the safety and the comfort of travelers on high speed trains. We believe it is possible to detect deterioration of the track by means of a constant monitoring of the dynamic vehicle response to track features. Our project is worked out in close cooperation with Deutsche Bahn AG (German Rail).

Keywords: Dynamic diagnosis, vehicle - track interaction, signal processing, proper orthogonal decomposition.

1. Introduction

High speed traffic on the Deutsche Bahn AG (German Rail) network has continuously increased in recent years. This additional use of the network raises questions about the maintenance of the quality of the trackage, which is essential for the safety and comfort of high speed travelers.

It is known that the more frequently applied dynamic loads due to the interaction between railway vehicles, the track and the sub–grade can lead to degradation of track quality and, sometimes, premature track failure. To prevent such failures or degradation in ride quality, the Deutsche Bahn AG performs expensive, frequent inspections of the railway tracks by means of special trains designed to perform measurements on the trackage.

Track geometry is assessed by bi–monthly measurements with an especially designed car filled with measurement systems at 200 km/h. In addition, vehicle reactions are measured with the ICE–S (Intercity Express–S), at high speed, every four months, Figure 1.

Decisions about what maintenance operations are necessary, and when, are based on empirically established threshold values of some parameters of track geometry. So far, however, there are no reliable means to predict actual rail failure based on variations of track geometry or the actions of excessive forces on the track by the train. Thus, it is difficult to schedule maintenance because one cannot predict where a failure will occur. Furthermore, there is no continuous assessment of small changes of track features, in an attempt to predict when the relevant parameters will reach limiting values.

We propose a program to monitor track condition constantly, at low cost, on regularly scheduled trains. It is our belief that we can correlate the dynamic response of a vehicle with changes in track and support geometry, and thus we can assess track conditions by regularly measuring vehicle behavior.

We start with a description of the problem from a dynamic system point of view and present first results of measurements from a real railway track. Then we turn to a signal processing point of view. We present the Karhunen–Loève–transformation as a tool for information compression for the expected huge amount of data resulting from the constant monitoring. Finally, we demonstrate the successful application of the Karhunen–Loève–transformation to data obtained in a small scale laboratory experiment.

* Dedicated to Ali H. Nayfeh on the occasion of his 65th birthday.

Figure 1. Special measurement train ICE–V, predecessor of ICE–S.

2. Dynamic System

The model we have used is shown schematically in Figure 2. As has been shown in [2], with such a model we can expect fairly good correlation between measured loads on a train and simulations based on this model.

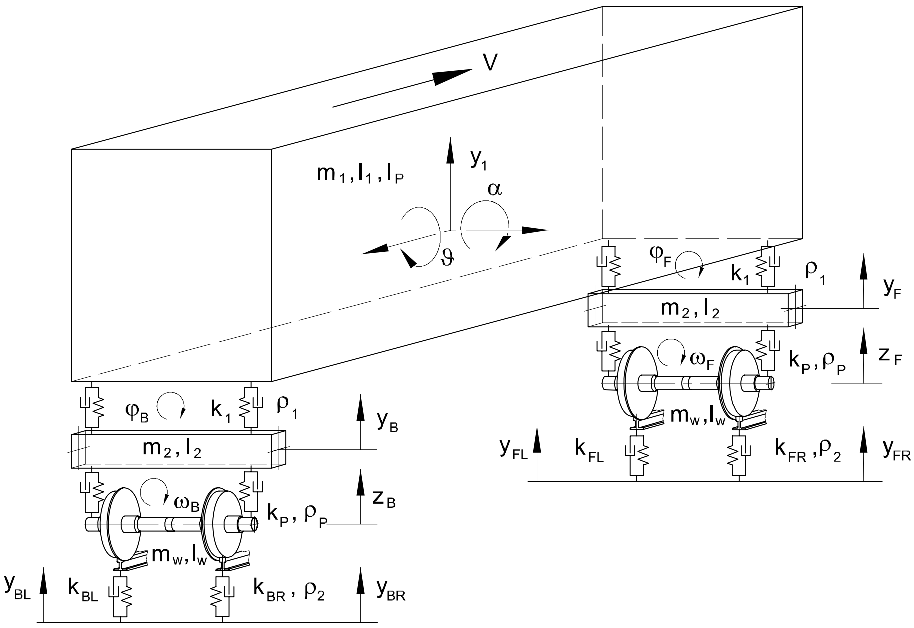

Figure 2. Multibody system: coach, bogie, wheelset, primary and secondary suspension and track.

Although this model neglects the contact dynamics between the wheels and the rails, and compresses the whole track structure (rail, rail pad, sleeper, ballast, sub–ballast, sub–grade) to a single spring damper system, the simulation based on this model allowed a fairly good prediction of track settlement due to accumulated dynamic loads. Therefore, we choose this model for our consideration.

In [2] it was found that spatial stiffness variations along the track play an important role for differential track settlement (which results in a rough vertical profile or geometry). Thus, it would be highly desirable to reconstruct the track stiffness from the dynamic response of the vehicle along the track.

Let us consider the vertical profile and the varying stiffness along the track as input u to the dynamic system and the measured signals within the vehicle as output y. The task described above requires an inversion of the dynamic system, Figure 3. The output of the inverse system u' asymptotically copies the original input signal u if and only if the inverse system has unique asymptotic behavior [1].

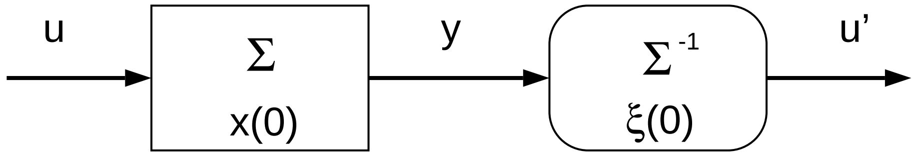

Figure 3. Inversion of a dynamic system: the output y of the system becomes the input of the inverse system. $x(0), \xi(0)$ are the initial states of the system and its inverse.

For nonlinear systems, inversion is solved in case they are *input–linear* [1], i.e., that the input u appears only as a linear term in the state equations $\dot{x} = f(x) + g(x) \cdot u$. Our system is, however, not input–linear, as products of the varying stiffnesses and the vertical profile occur in the state equations,

$$\begin{bmatrix} m_1\ddot{y}_1 \\ I_1\ddot{\vartheta} \\ I_p\ddot{\alpha} \\ m_2\ddot{y}_F \\ I_2\ddot{\varphi}_F \\ m_2\ddot{y}_B \\ I_2\ddot{\varphi}_B \\ m_w\ddot{z}_F \\ I_w\ddot{\omega}_F \\ m_w\ddot{z}_B \\ I_w\ddot{\omega}_B \end{bmatrix} = \mathbf{f}(\mathbf{x},\dot{\mathbf{x}}) + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ k_{FL}(t)y_{FL}(t) + k_{FR}(t)y_{FR}(t) \\ k_{FL}(t)ly_{FL}(t) - k_{FR}(t)ly_{FR}(t) \\ k_{BL}(t)y_{BL}(t) + k_{BR}(t)y_{BR}(t) \\ lk_{BL}(t)y_{BL}(t) - lk_{BR}(t)y_{BR}(t) \end{bmatrix} \tag{1}$$

with the state vector $\mathbf{x} = \begin{bmatrix} y_1 & \vartheta & \alpha & y_F & \varphi_F & y_B & \varphi_B & z_F & \omega_F & z_B & \omega_B \end{bmatrix}$. The state space variables together with parameters of the model are shown in Figure 2, l is the distance between the rails. The input of the vehicle multibody system can be viewed as an excitation by means a track which moves underneath the nonmoving vehicle. This results in time dependent functions of the track stiffness k_{ij} and the vertical profile y_{ij}, $i = F, B$ $j = R, L$. They are related to their spatial variation via the velocity of the vehicle. We do not present the vector $\mathbf{f}(\mathbf{x}, \dot{\mathbf{x}})$ explicitly which summarizes the forces and moments of the multibody system imposed by the springs and dampers which are either linear or nonlinear.

It is an open question whether such *input–quadratic* systems are dynamically invertible.

Since we cannot infer from system inversion the track features (stiffness and vertical profile) explicitly, we want to detect the variation or deterioration of the track implicitly via the variation of the vehicle response. That means that we expect the vehicle responses at subsequent journeys under the same conditions (track features) to be the same. Thus, we expect variations of the vehicle response to result from actual variations of track features.

This is not necessarily the case because in a chaotic system, e.g., departing from different initial states would produce diverging responses to the same input signal. Thus, we require that the dynamic system have *unique asymptotic behavior* (UAB), i.e., the system response to the *same* input signal, but possibly departing from different initial states, converge to each other.

The UAB can be checked by considering the difference between any two solutions. Thus, one must prove the asymptotic stability of the origin of the difference system,

$$\Delta\ddot{\mathbf{x}} = \ddot{\mathbf{x}}_1 - \ddot{\mathbf{x}}_2 = \mathbf{f}(\mathbf{x_1}, \dot{\mathbf{x}}_\mathbf{1}) - \mathbf{f}(\mathbf{x_2}, \dot{\mathbf{x}}_\mathbf{2}). \tag{2}$$

This nonlinear difference system can often be put into a linear time–varying form. The asymptotic stability of this system can be proven by the application of the Kalman–Yacubovitch Lemma [5]. The proof of UAB of our system is difficult because of the high dimension of the state space. It is the subject of ongoing research.

Next we present some measurement results in order to check UAB with real world data.

3. Measurements

We collected bi–weekly measurements with a specific vehicle on a selected piece of track in order to evaluate continuous monitoring. We installed an inertial measurement unit (IMU) from IMAR Company, Germany, on the floor of the coach. The IMU measures three translational accelerations and three angular velocities. In this way all six degrees of freedom are determined.

As we are especially interested in the effects on the vertical dynamics of the track, we concentrate on measurements of the pitch angle α and the vertical displacement y_1. See Figure 4 for a representative sample of the pitch angular velocity $\dot{\alpha}$ of two subsequent journeys along the same track at the same speed (200 km/h) and under presumably the same track conditions. The signal–to–distance correspondence along the track was determined by means of an additionally recorded signal (50 pulses/m) and the reproducibly unique yaw–signal. As the two signals in Figure 4 are hardly distinguishable, the dynamic system actually seems to have the UAB property.

Figure 5 shows the measured pitch anglular velocity $\dot{\alpha}$ and vertical acceleration $\ddot{y}_1$ of the coach in higher resolution with respect to the distance. Again the measurements of both journeys nearly coincide. The signals represent the response of the coach to the onset of a curve at km 21.2 of the selected piece of the track. Apparently, an impact excites some oscillations, which are clearly visible. They have a frequency of about 1 Hz (speed was 160 km/h), which presumably corresponds to an eigenfrequency of the system (in case it is linear).

It is our belief that changes of the track, e.g., changes of the magnitude of the impact, will cause noticeable changes of the response of the coach.

Our bi–weekly measurements cover so far only a short period of time and we do not have enough data to estimate whether small deteriorations of the track are actually detectable in the coach response. Nevertheless, we performed small scale laboratory experiments, whereby it is easy to collect as much data as desired and to simulate deteriorations of the track by appropriate manipulations.

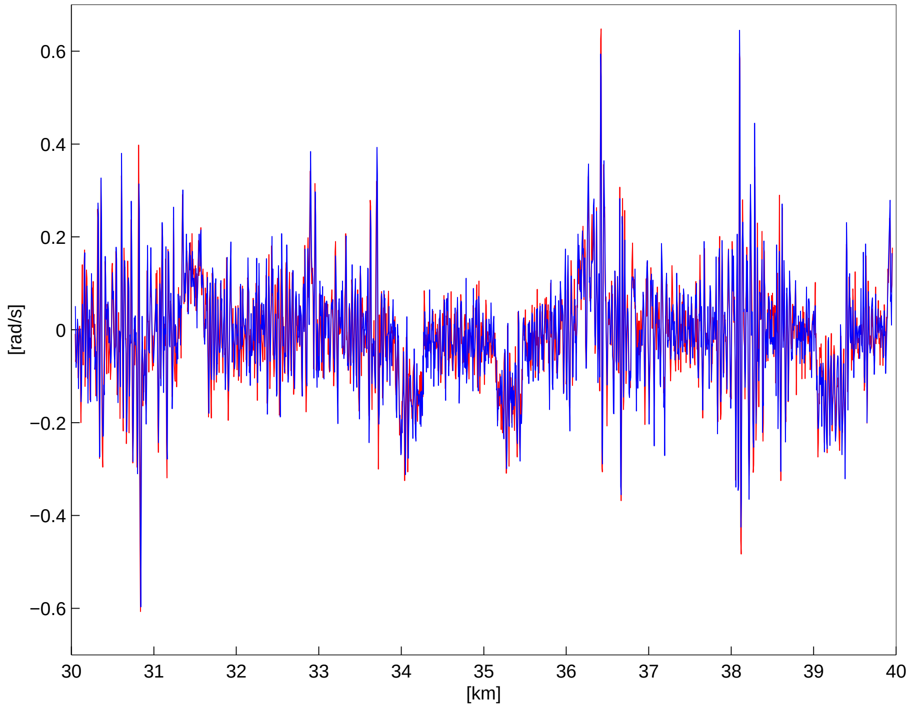

Figure 4. Two measurements of pitch angular velocity $\dot{\alpha}$ along the same track.

It is clear that, however, constant monitoring will result in a collection of a substantial amount of data, and will require an efficient compression and evaluation process. We propose to apply the **Karhunen–Loève–Transformation (KLT)** for this purpose. Therefore, we briefly present the basic idea of the KLT as well as its essential ingredients and then demonstrate its successful application to data of our laboratory experiments.

4. Karhunen–Loève–Transformation

The Karhunen–Loève–Transformation (KLT) or Proper Orthogonal Decomposition originally stems from the field of stochastic processes. An engineering presentation can be found in [3] and an application to a mechanical system is discussed in [4]. The KLT detects basic components of a set of stochastic signals by means of its covariance function. An aim can be to separate significant components from less significant ones (e.g. noise).

We intend to apply the KLT to signals to what are presumed to be a deterministic data. The purpose is also to separate the significant information from possibly added noise but mainly to extract from a huge set of signals its dominant components.

Given a set of signals

$$u(n,x) \text{ with } x \in [a,b] - \text{Distance and}, n - \text{No. of measurement} \tag{3}$$

the KLT provides a decomposition

$$u(n,x) = \sum_{i=1}^{\infty} a_i(n)\psi_i(x) \tag{4}$$

where the *characteristic functions* (CF), $\psi_i(x)$, represent basic patterns and the *amplitudes*, $a_i(n)$, are their contribution in each signal.

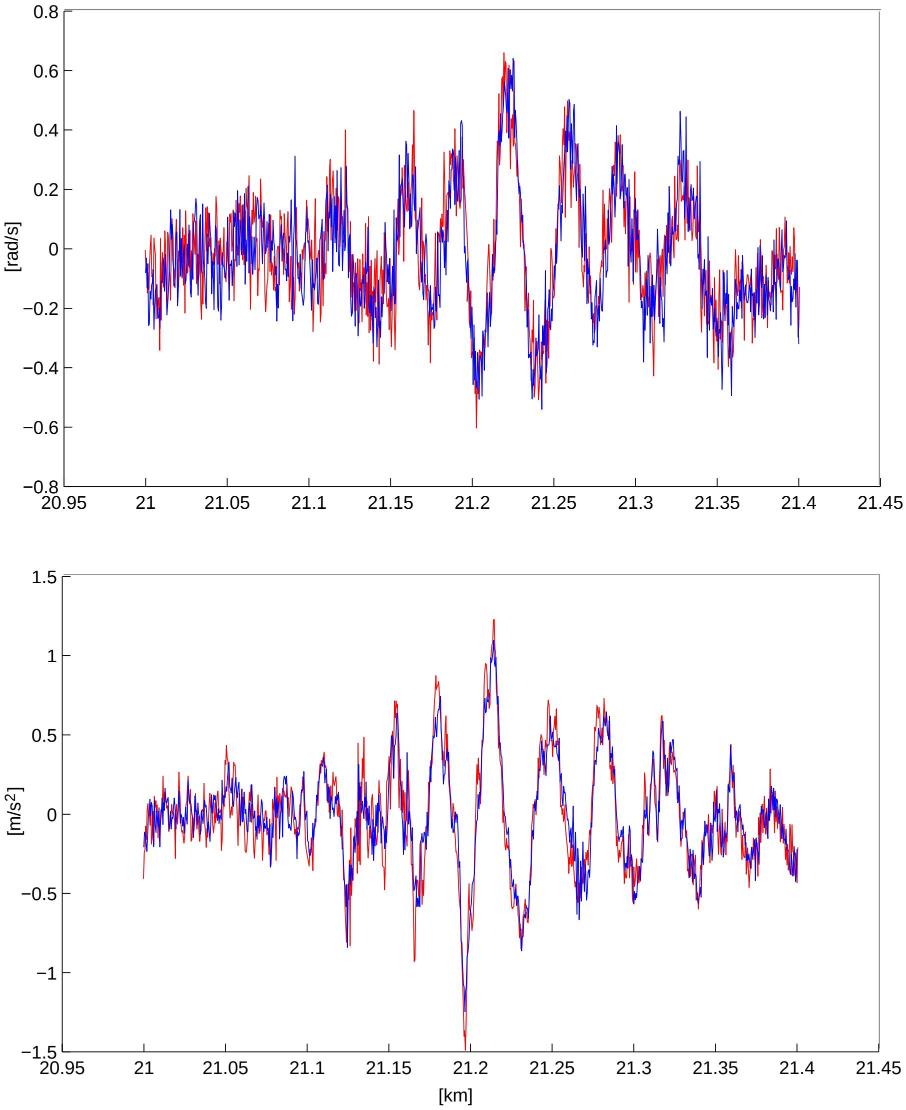

Figure 5. Above pitch anglular velocity $\dot{\alpha}$ and below vertical acceleration $\ddot{y}_1$ of the coach response to the onset of a curve (at km 21.2).

This resembles the well known Fourier transform (FT), where a signal is described as a superposition of basic functions ($\sin(x)$ resp. $\cos(x)$) multiplied by certain amplitudes. The important difference between KLT and FT is that the basic functions of the FT are fixed, whereas the characteristic functions of the KLT are determined by and thus are specific to the set of signals $u(n, x)$.

The two basic requirements of the KLT are:

First, the characteristic functions $\psi_i(x)$ form an orthogonal basis:

$$\langle \psi_i(x), \psi_j(x) \rangle = \int_a^b \psi_i(x)\psi_j(x)dx = \delta_{ij}. \tag{5}$$

Second, the amplitude functions $a_i(n)$, which are the projections of the signals onto this base,

$$a_i(n) = \langle u(n, x), \psi_i(x) \rangle = \int_a^b u(n, x)\psi_i(x)dx \tag{6}$$

are uncorrelated

$$\mathbf{E}_n(a_i(n)a_j(n)) = \lambda_i \delta_{ij}, \tag{7}$$

where $\mathbf{E}_n$ is the expected value over n.

These requirements are met by $\psi_i(x)$, which are the eigenfunctions of the integral operator

$$\int_a^b r_{uu}(x', x)\psi_j(x)dx = \lambda_j \psi_j(x') \tag{8}$$

where $r_{uu}(x', x)$ is the autocorrelation function of the signals

$$r_{uu}(x', x) = \mathbf{E}_n(u(n, x')u(n, x)). \tag{9}$$

Next, we consider the approximation

$$u_N(n, x) = \sum_{i=1}^{N} a_i(n)\psi_i(x). \tag{10}$$

The *energy* contents covered by the approximation $u_N(n, x)$ is

$$\mathbf{E}_n(u(n, x)u(n, x)) = \sum_{i=1}^{N} \mathbf{E}_n(a_i^2(x)) = \sum_{i=1}^{N} \lambda_i. \tag{11}$$

Thus, if one chooses the characteristic functions ψ_i belonging to the N largest λ_i for the approximation then it covers the highest possible information content of the set of signals, and the mean square approximation error is as small as possible.

The efficiency of the KLT arises from the fact that it is *optimal* in the sense that for all N the approximation with the first N characteristic functions covers the most information contents of the signals (or the mean square approximation error is a minimum) amongst all possible other bases $\phi_i(x), i = 1, ...N$. We employ this optimality feature for the purpose of best information compression.

All this carries quite directly over to the *discrete* Karhunen–Loève–transformation which we will apply, because every measurement naturally is a discrete set of data, where

$$u(n, x_i) = \mathbf{u}(n) \text{ with } i = 1, ..., d \text{ and } \mathbf{u} \in \mathrm{R}^d \tag{12}$$

and the characteristic functions, $\boldsymbol{\psi}_i \in \mathrm{R}^d$, are the eigenvectors of the covariance matrix

$$(\mathbf{C}_{uu} - \lambda_i \mathbf{I})\boldsymbol{\psi}_i = 0 \tag{13}$$

with

$$\mathbf{C}_{uu} = \mathbf{E}_n(\mathbf{u}(n)\mathbf{u}^T(n)). \tag{14}$$

5. Small Scale Experiment

We performed small scale laboratory experiments, Figure 6. We installed angular velocity sensors in the coach.

We present measurements along a 2 m test section. The sleepers are spaced at 20 cm and we varied the height of the sleepers no. 6 and 7 at 1 m and 1.2 m, Figure 7.

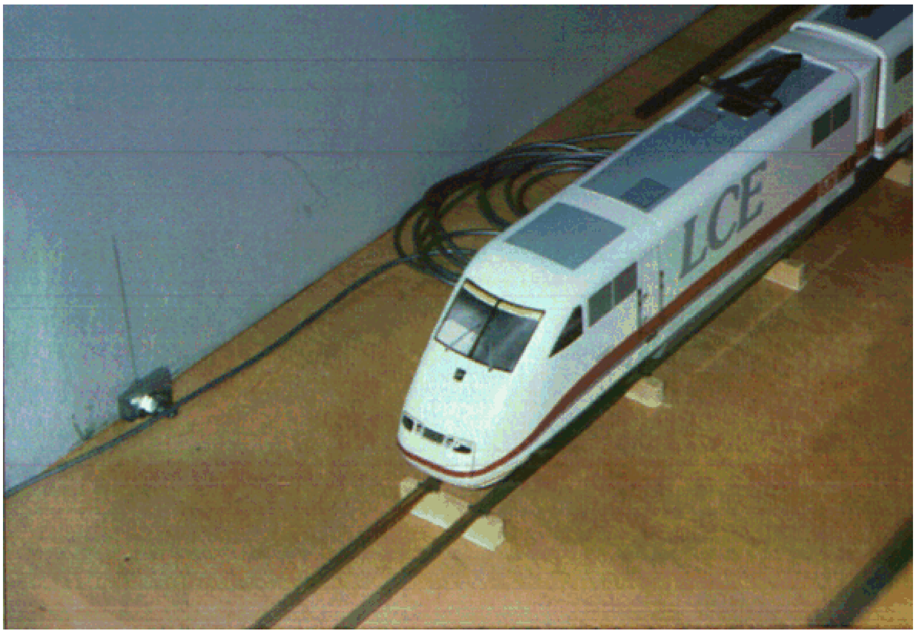

Figure 6. Small scale laboratory experiment.

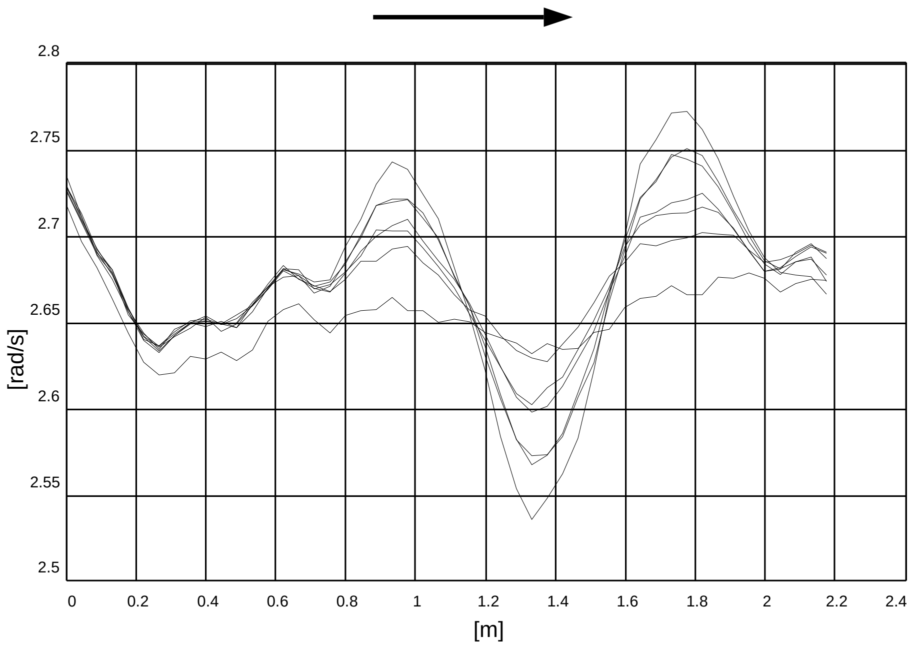

Figure 7. Measured angular velocity (pitch) of the coach under varying track conditions, journey there. The arrow indicates the direction of the train velocity along the track.

The measured response is of course a dynamic reaction of the system to the impact of track failure. This is also obvious by comparison of the responses of the journey there and the return journey, Figure 8, to the same impact.

The information contained in this set of signals has to be compressed and evaluated efficiently. We apply the Karhunen–Loève–transformation in order to detect the basic components of this set of signals and their contribution to every single signal.

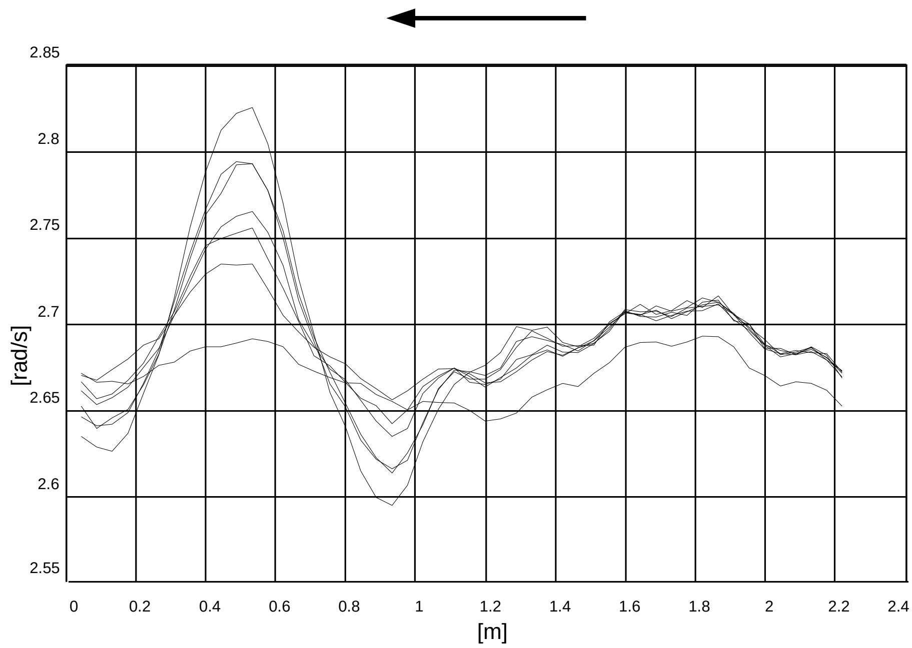

Figure 8. Measured angular velocity (pitch) of the coach under varying track conditions, return journey.

Often the KLT is applied to the mean–value free signal $\mathbf{E}_n(\mathbf{u}(n)) = 0$ in order to pronounce the actual variations in the set. We are especially interested, however, in variations with respect to the initial signal. Therefore, we apply the KLT to the difference signal with respect to the first signal:

$$\mathbf{u}(n) = \dot{\boldsymbol{\alpha}}(n) - \dot{\boldsymbol{\alpha}}(1). \quad (15)$$

First we calculate the covariance matrix, $\mathbf{C}_{uu}$, of the signal set, Figure 9. The eigenvec-

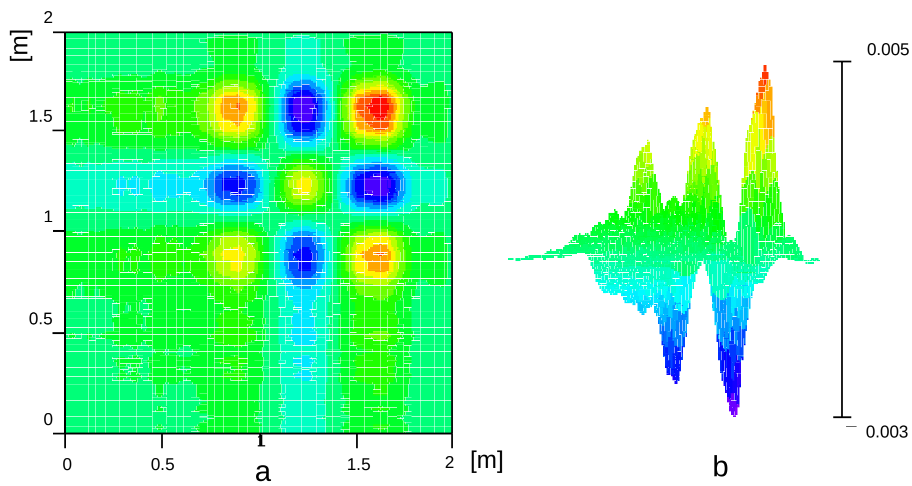

Figure 9. Covariance matrix $\mathbf{C}_{uu}$: a) contour–plot the x, y–axes represent the matrix indices corresponding to the piece of track, the value of the matrix entry is color–coded, b) perspective view on $\mathbf{C}_{uu}$ onto its diagonal.

tors of this covariance matrix are the basic components of the set $u(n, x)$; the eigenvalues corresponding to the eigenvectors are a measure of their significance (energy contribution).

The KLT reveals the major features of what we see in Figure 7. The set of signals seems to consist of only one basic pattern, which is contained in each signal, with different 'amplitudes'. Analysis of the eigenvalues of $\mathbf{C}_{uu}$ confirms that only one eigenvector is significant. This eigenvector covers already 96% of the energy of the set. The second resp. third eigenvector covers merely 3.4% resp. 0.6% of the energy. Figure 10 shows the first eigenvector, i.e., the first characteristic function, $\boldsymbol{\psi}_1$. Thus, the whole set of signals is well approximated by the first eigenvector, and its information compressed to a single characteristic function and its amplitude function.

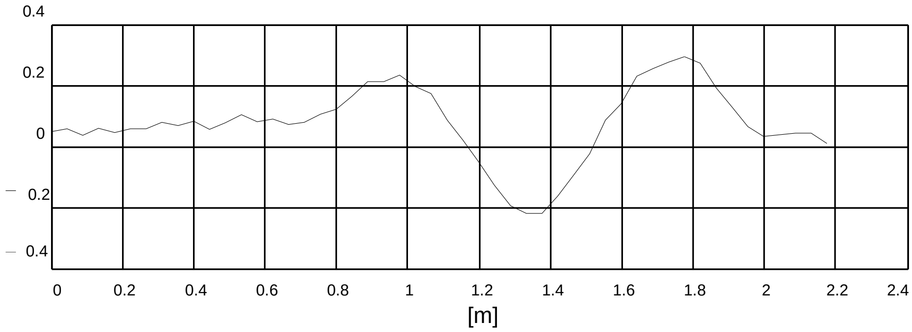

Figure 10. First characteristic function $\boldsymbol{\psi}_1$ of the set of signals in Figure 7.

The height of the sleepers no. 6 and 7 was decreased until journey 10 and then it was increased again. We conducted two measurement journeys per height. Figure 11 shows the first amplitude function, i.e., the contribution of the first CF, $\boldsymbol{\psi}_1$, to each signal.

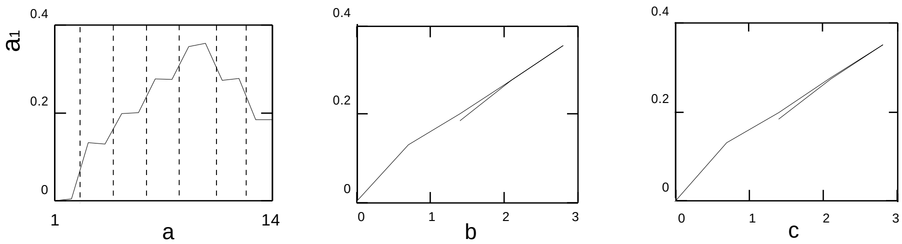

Figure 11. First amplitude function: a) $a_1(n)$ contribution of the first CF to each journey there, b) a_1 as a function of the settling of the sleepers in mm, c) same as b) but for the return journey.

The first amplitude function reveals: First, it is nearly identical for identical track conditions (for the two journeys per height). Second, a_1 is nearly independent of the history of the height of the sleepers, i.e., there is no hysteresis in Figure 11 b,c. Third, a_1 is nearly a *linear function* of the settling of the sleepers in mm. Fourth, although the first characteristic function of the return journey is completely different from the one of the journey there, its amplitude reflects the severeness of the failure in the same way.

The application of the KLT to the data of our small scale laboratory experiment shows that the Karhunen–Loève–transformation is an efficient tool for information compression. It is also capable of displaying the severity of a track failure.

6. Conclusion

This article describes an approach for a continuous track survey in the Deutsche Bahn AG railway network. With this approach a systematic and comparative inspection and assessment of railway tracks will be possible. Instead of expensive special measurement trains, we need only relatively inexpensive equipment installed in the coach of a wagon.

The first idea is to detect the track settlement due to accumulated dynamic loads by means of an inverse dynamic system that reconstructs the original input, the track features, from measurements in the coach. Due to the input-quadratic character of the system, however, this approach is not applicable.

We present real world measurements. They indicate that the dynamic vehicle-track system has unique asymptotic behaviour, such that we can expect variations of the vehicle response to stem actually from variations of the track or the vehicle.

It is understood that the dynamic response of the vehicle, i.e. the measured signals, are influenced by track irregularities and vehicle features like worn wheels and damged bogies. It is our belief that the variation of the vehicle response contains significant information about the variation of the vehicle-track system. The specific response (e.g. a certain periodicity for worn wheels) should allow to separate different kinds of failures from each other.

In order to handle the enormous amount of data obtained from the measurements we apply Karhunen–Loève–transformation (KLT) for data compression. The successful application of KLT to a small scale experiment shows that efficient information compression is possible and it enables us to measure the severeness of a failure.

7. Acknowledgment

The authors are indebted to the DFG (Deutsche Forschungsgemeinschaft/German Research Foundation) for funding the work under the contract Kr 752/18-1,2. Moreover, we would like to thank Deutsche Bahn AG for the opportunity to perform the measurements.

References

1. Feldmann, U., Hasler, M., and Schwarz, W.: 1996, 'Communication by chaotic signals: the inverse system approach'. *International Journal of Circuit Theory and Applications* **24**(5), 551–579.
2. Fröhling, R.: 1997, 'Deterioration of railway track due to dynamic vehicle loading and spatially varying track stiffness'. Ph.D. thesis, University of Pretoria, Pretoria.
3. Holmes, P., Lumley, J., and Berkooz, G.: 1996, *Turbulence, coherent structures, dynamical systems and symmetry.* Cambridge: Cambridge University Press.
4. Kreuzer, E. and Kust, O.: 1996, 'Analysis of long torsional strings by proper orthogonal decomposition'. *Archive of Applied Mechanics* **67**, 68–80.
5. Vidyasagar, M.: 1993, *Nonlinear Systems Analysis.* Englewood Cliffs N. J.: Prentice-Hall, 2nd edition.

Address for Offprints:
Dr. Ute Feldmann and Prof. Edwin Kreuzer,
TUHH, AB 3-04, 21073 Hamburg, Germany,
e-mail: u.feldmann@tu-harburg.de
Fax: +49 40 42878 2028