APPLYING A PRINCIPAL COMPONENT ANALYSIS TO MOVEMENT COORDINATION IN SPORT

K. Witte[1], N. Ganter[1], Ch. Baumgart[2], Ch. Peham[3]
Otto-von-Guericke-Universität Magdeburg, Magdeburg, Germany; [2]Bergische Universität Wuppertal, Wuppertal, Germany, [3] University of Veterinary Medicine, Vienna, Austria

Corresponding Author: K. Witte, Otto-von-Guericke-Universität Magdeburg, Dept. Sport Science, Brandenburger Str. 9, 39106 Magdeburg, kerstin.witte@ovgu.de

Abstract. Because sports movements are very complex biomechanical analyses contain many kinematical or dynamical parameters and characteristic curves. PCA is a technique for simplifying a dataset by reducing multidimensional datasets to lower dimensions for analysis. The purpose of this paper is the presentation of several studies which used the PCA to solve some problems in the movement science in sport. Especially, we interpret the number of the components or also named components with relative high eigenvalues as the number of degrees of freedom. For cyclic and automated movements the first PCA-component is dominant. The PCA was successful applied to gait analyses in rehabilitation and in triathlon as well as in riding. Phase plots could be used to quantify the variability of the movement coordination.

1 Introduction: Using of Principal Component Analysis in Movement Science

Because sports movements are very complex biomechanical analyses contain many kinematical or dynamical parameters and characteristic curves. From systems point of view only a few quantities are sufficient to describe the movement coordination [1], [2]. There are many methods for linear (time invariant) component order reduction. An important class of techniques are known as projection methods attempts to find the best approximating subspace in terms of data variance upon which to project the system dynamics. The principal component analysis (PCA) is one popular method, in part because it is numerically feasible for large dimensional systems [3], [4], [5]. It is data-driven, that means the results are inherently a function of a specific data set. It can be declare, that the PCA is a technique for simplifying a dataset by reducing multidimensional datasets to lower dimensions for analysis. In [6] PCA was used for a constructive pattern analysis for gait. This technique was applied to 15 segmental angles, and they could successfully reduce the corresponding parameter set to only three components. The shapes of the phase plots of these variables allowed a discrimination of gaits under different neurophysiologic conditions [6]. [7] and [1] assume that cyclic, learned and automatic movements are governed by only one order or systems parameter. They confirmed this hypothesis by a study of the learning process of drive a pedalo. An estimation of the number of order parameters was realized by means of the Karhunen-Loève-method or PCA [1] Likewise [8] applied the linear and nonlinear PCA to the learning process of drive a pedalo to determine the degrees of freedom during the learning process. Another example for using the PCA is the classification of the short and long serves in table tennis [9].

The purpose of this paper is the presentation of various studies which used the PCA to solve some problems in the biomechanics in sport. Especially, we interpret the number of the components or also named components with relative high eigenvalues as the number of degrees of freedom.

2 Method of Principal Component Analysis

By means of the PCA it is researched how the variance of a data vector is composed of the variances of the single components. This method is a linear transformation which transforms the data to a new coordinate system in such a way that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first component), the second greatest variance on the second coordinate (called the second component), and so on. For the definition of the data vector or “pattern” vector it is assumed that it’s components are special biomechanical parameters for this movement (angles, velocities,…). So the “pattern” vector can be expressed:

$$\vec{v}(t) = (\vec{v}_1(t), \vec{v}_2(t), ..., _{\vec{v}}\ \vec{v}_N(t)) \tag{1}$$

N as the number of components of $\vec{v}(t)$

The PCA then portrays the variance of the “pattern” vector. For this decomposition this “pattern vector” is projected onto an arbitrary, time independent vector $\alpha = (\alpha_1, \alpha_2, ..., \alpha_N)$ of unit length. This time dependent projection can be written as

$$\vec{v}(t) = \alpha_1 v_1(t)\vec{e}_1 + ... + \alpha_n v_n(t)\vec{e}_N\ , \tag{2}$$

where $\alpha = (\alpha_1, \alpha_2, ..., \alpha_N)$ denote the unit vector. The components of the vector α are estimated in such a manner that the variance of the projection reaches a maximum:

$$\sum_t \left| v(t) - \bar{v} \right|^2 = Max. \tag{3}$$

with

$$\bar{v}_j = \frac{1}{T} \int_0^T v_j(t) dt, \tag{4}$$

j = 1,…, N. That means with other words, that the variance of the difference between “pattern” vector and projection is a minimum.

The “pattern vector” is then decomposed into two time dependent series, one time series lying in the direction of the projection of the greatest variance, and the other time series lying orthogonal to it. This procedure is repeated with this second orthogonal time series. This new direction of maximum variance logically is orthogonal to that of that 1st maximum variance. Consequently iterating the entire algorithm yields further directions of relative maximum variance and leads to a series of monotonously decreasing values for the remaining variance of the pattern vector at the intersection of the orthogonal complements. The directions of maximal variance paired orthogonal to each other are called components. Each component contains a part of the total variance, which is defined by the belonging eigenvalue. For the purpose of clarity the sum of the eigenvalues is normalised to 100 %. So each eigenvalue indicates the percentage in the total variance of the related component. Generally components with eigenvalues less than 5 % are disregarded.

The PCA can also be realized by determination of the covariance-matrix of the time series of the components (angles) of the “pattern vector”. The calculated components are called eigenvectors, the variances are the eigenvalues of the covariance-matrix.

3 Results and discussion

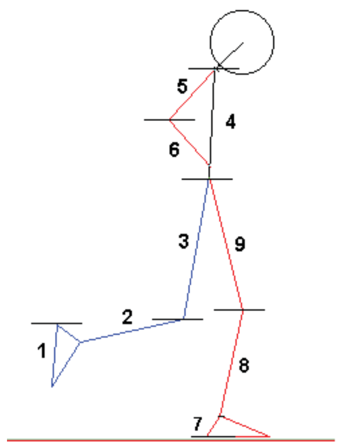
3.1 Study 1: Transition from walking to running

The aim of this study is the characterization of the walking and running movement by means of an order parameter and analysing its changes under the condition of varying velocity. It is known that slight variations of external factors can lead to considerable effects in the movement. In this way a gradual increasing of the treadmill velocity causes the transition from walking to running. Referring to thesis empirical and theoretical researches were carried out (e.g. [10], [11], [12], [13]). The investigations were carried out with two male students of sport science (CK and R), who were familiar with running on a treadmill. Several individual treadmill velocities were set, with duration of 3 minutes each. The velocities were chosen in such a way that the subject can walk comfortably at 2 – 3 velocities before the transition to run. After enough accommodation in the third minute, seven or more step cycles were recorded by a high-speed-camera (200 Hz). The two-dimensional kinematical analysis was carried out by means of the SIMI-Motion software. Further on, the angles of the body, which are given in tab. 1, were used for the PCA.

It can be discovered that the first component is distinguished by the greatest eigenvalues for both subjects and for all velocity steps (s. fig.1). Besides, small differences between the velocities and between walk and run can be observed. But these distinctions are inter-individual, so general evidence is not possible. With this result confirm the thesis that the movement coordination of walking and running can be described by only one systems parameter. To determine the systems parameter ξ and the resulted phase plot (ξ vs. dξ/dt) to characterise its dynamics ξ was calculated by means of the Haken-Friedrich-Uhl-order parameter analysis [1]. Fig. 2 shows examples for phase plots in dependence on the velocity for the subject CK with marked time points of the movement phases within the movement cycle. The closed curves refer to cyclic movements. In the phase plots for walking addition “spikes“ could be found. These irregularities can be observed in particular in the phase of double support.

Tab. 1: Definitions of the angles of body in relation to the transversal plane (xz-plane), (study 1)

Nr.	Definition of the angle	Designation
1.	Angle in relation to the connection toe - heel and xz-plane (left)	*w*1-l
2.	Angle in relation to the connection ankle – knee and xz-plane (left))	*w*2-l
3.	Angle in relation to the connection knee-hip and xz-plane (left)	*w*3-l
4.	Angle in relation to the connection shoulder-hip and xz-plane (left)	*w*4-l
5.	Angle in relation to the connection elbow-shoulder and xz-plane (left)	*w*5-l
6.	Angle in relation to the connection joint of hand-elbow and xz-plane (left)	*w*6-l
7.	Angle in relation to the connection toe - heel and xz-plane (right)	*w*7-r
8.	Angle in relation to the connection ankle – knee and xz-plane (right)	*w*8-r
9.	Angle in relation to the connection knee-hip and xz-plane (right)	*w*9-r



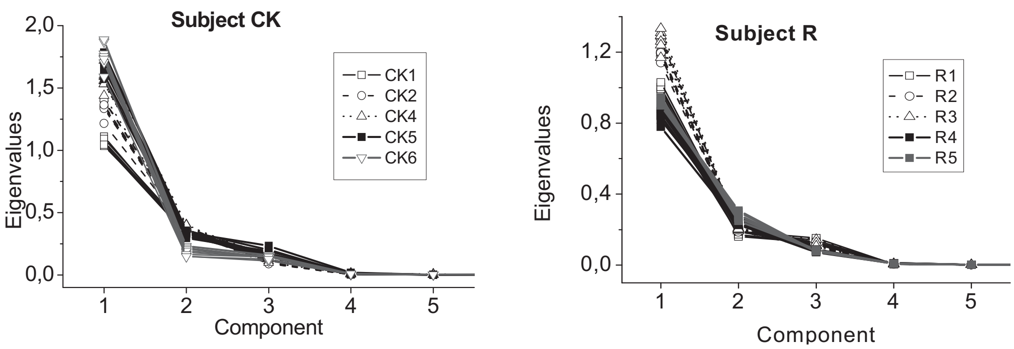


Fig. 1: Eigenvalues of the PCA-components, calculated for single cycles in each velocity step. Left: subject CK, right: subject R, (study 1)

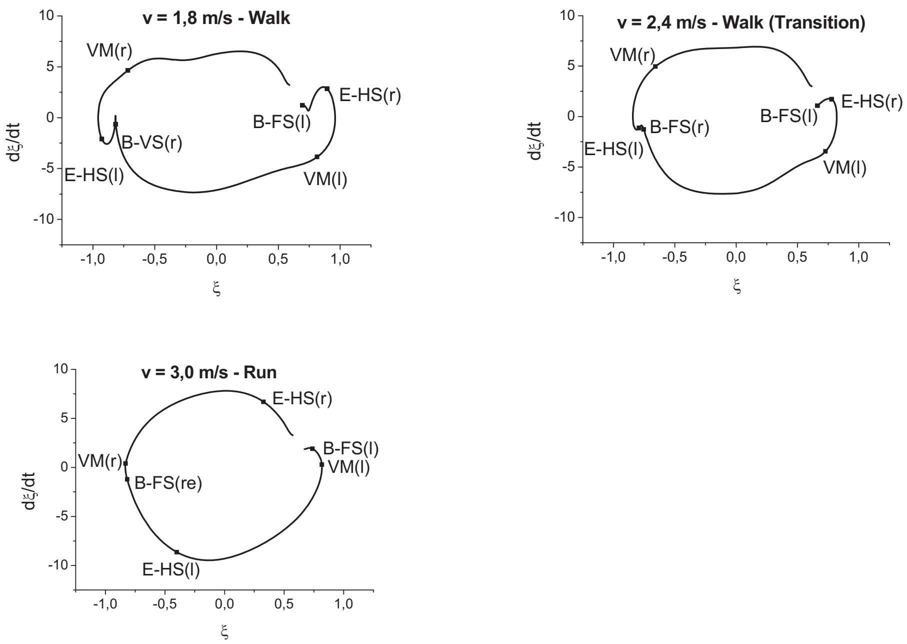


Fig. 2: Examples for phase portraits or phase plots in dependence on the velocity for the subject CK. In the movement phases of walk or run are drawn: VM – vertical moment, B-FS – begin of the front supporting phase, E-HS – end of the behind supporting phase, l – left leg, r – right leg, (study 1)

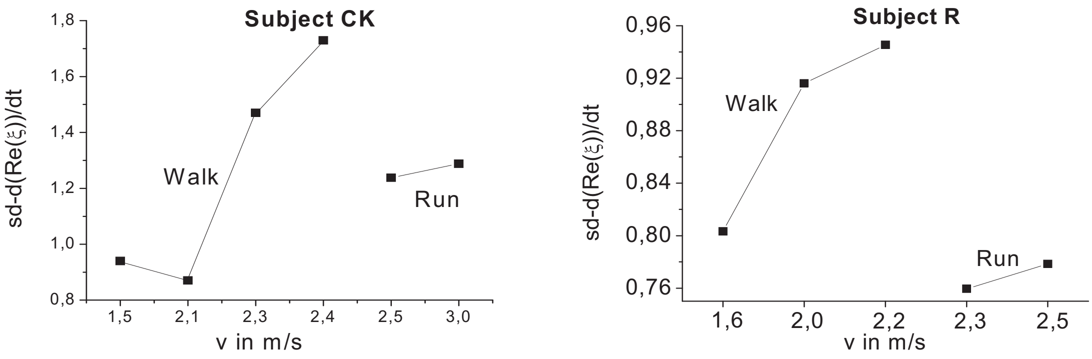


Fig. 3: Mean standard deviation of the temporal derivation of the systems parameter in dependence on the velocity (left: subject CK, right: subject R), (study 1)

Another advantage of describing of the movement coordination by means of a systems parameter is the possibility to quantify the variability of the movement. Of special interest is the dependency on the velocity and on the locomotion. The following analysis relates to the temporal derivation of the order parameter, which was calculated for all single cycles. The mean standard deviation of all single curves was used as a measure of the variability. The diagrams in fig. 3 show this mean standard deviation in dependence on the velocity. We can see that the variability of the dynamics of the order parameter for walking increases with grown velocity. The variability is particularly high immediately before transition. These fluctuations known in synergetics are necessary for self-organization. They drive the system to discover new states and allow a phase change [14]. In this way, the new state of the system, running, is characterized by a particularly high stability.

3.2 Study 2: Characterisation of walking coordination during rehabilitation of patients with a knee joint endoprosthesis

The purpose of this study was the estimation of the number of PCA-components and their eigenvalues relating to the rehabilitation process. The determination of the "pattern" vector was carried out on the basis of the 19 body angles. The angle-time-courses resulted from a three-dimensional video analysis (SIMI Motion, sample rate: 50 Hz). Per researching time point 6 – 8 steps were analysed. Nine (3 female and 6 male) patients in an age from 45 to 70 years participated in this study.

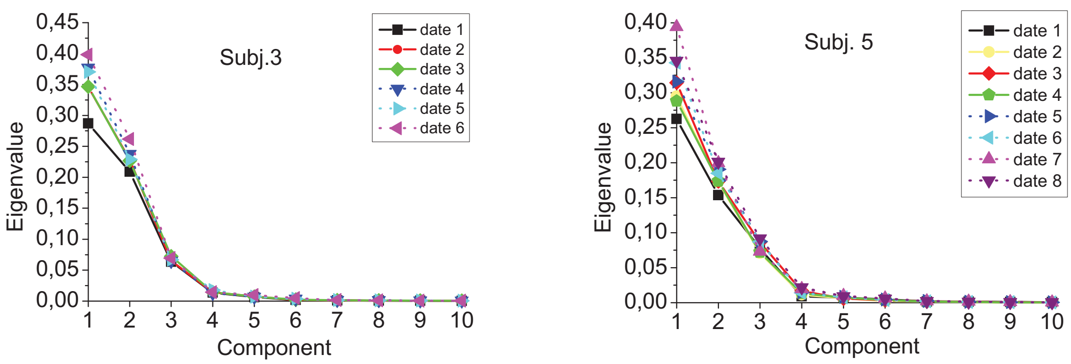


Fig. 4: Mean eigenvalues of the first ten PCA-components for two patients (without crutch). Date1 was the first research and the date 6 (subj. 3) or 8 (sub. 5) were the last researching time point. (Study 2)

The diagrams in fig. 4 show for two patients the mean eigenvalues of the first PCA-components in dependence on the duration of the rehabilitation. The dominance of the first component is obvious. During the rehabilitation the meaning of the first component increases. This can be interpreted that during the rehabilitation the movement coordination gets more stable. In contrast to a normal walk of a healthy person the second PCA-components have relative high eigenvalues. The fig. 5 shows the percentage of the second component to the first component. It can be found that by using of crutches the percentage of the second component in relation to the first component is lower than without crutches. A possible explanation is the influencing on the individual gait coordination by using a crutch.

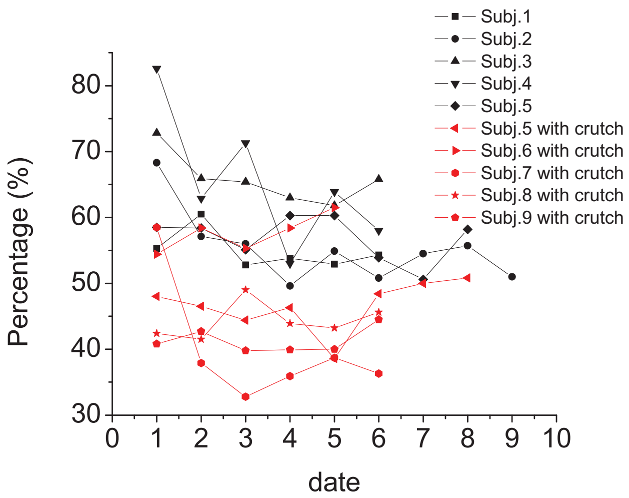


Fig. 5: Percentage of second component in relation to the first component during the rehabilitation for all patients, (study 2)

3.3 Study 3: Characterisation of running coordination after prior cycling exercise in triathlon

It is known that triathletes have the feeling of another coordination of running after prior cycling exercise than running under normal conditions. Therefore, the following question should be answer: There are changes in the coordination of running after prior cycling exercise? Two groups participated in this study. Group 1 consisted of five male experienced triathletes and group 2 was built by five male sports students. After a pre-run-test a cycling-distance (40 km for group 1 and 20 km for group 2) was done. After them a treadmill running of 2 km followed. By means of a kinematical video analysis system 20 cycles were analysed to the following time points: pre-run-test (PreRun), 30 s (Run 1), 2 min (Run 2), 4 min (Run 3) and 7 min (Run 4) after the beginning of the second run-test.

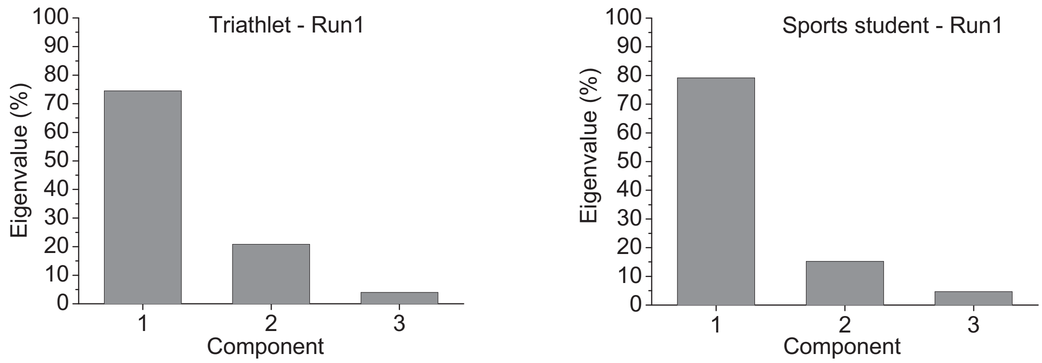


Fig. 6: Mean eigenvalues of the first three PCA-components for Run 1 (study 3)

On consideration of Run 1 (fig. 6) the dominance of the first component is evident. A slight change in movement coordination can be assumed due to the progressive running exercise. In relation to the first PCA-component the eigenvalues of the second KL-component for the triathletes are higher than for the sports students. The fig. 7 shows the percentage of the second and the third PCA-component in relation to the first PCA-component for one triathlete and one sports student. Significant differences between the single runs could not be found. In relation to the first component the eigenvalues of the second component for the triathletes are higher than for the sports students. Also in this study we assume that the second PCA-component characterises the individual style of the movement.

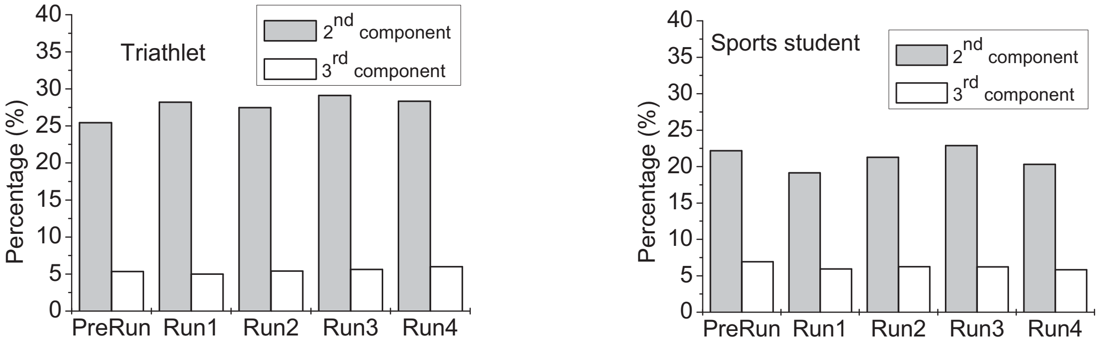


Fig. 7: Percentage of the second and third PCA-component in relation to first component (study 3)

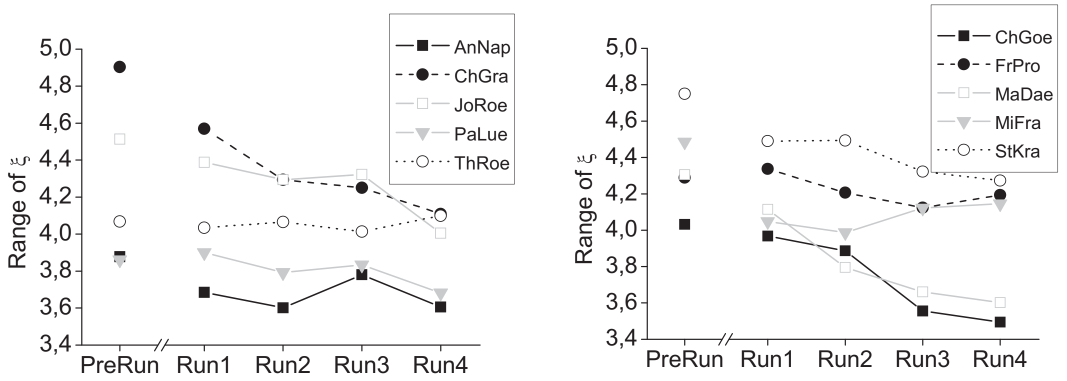


Fig. 8: Mean ranges of the systems parameter to characterise the variability of the movement coordination (study 3)

Another aspect is the calculation of the systems parameter in analogy to study 1. The determination of its range is a measure of the variability of the movement coordination. The diagrams in fig. 8 show that the variability of the movement coordination decreases immediately after the cycling exercise and in the further progress of the running, especially for sports students. Changes in single kinematical parameters of running immediately after a cycling exercise are only individually.

3.4 Motion pattern analysis of gait in horseback riding

Horseback riding concerning the three interactive systems of horse, saddle and rider is a very complex movement which is difficult to characterise by single biomechanical parameters or characteristic curves. All these three components of the combined horse-rider system have their own geometry, inertia, elasticity, degrees of freedom, etc. Two components moving actively driven by their intrinsic muscle system, the third moving passively, coupling the two active parts together. The horse and rider can learn to coordinate successfully their combined motion, which is intuitively felt by equestrians and described as “harmony” [15]. As mathematical procedure the PCA was used. By means of this method it should be prove whether the three gait patterns walk, trot and canter can be characterised by one or a few systems parameters. In the case that it is possible the next step will be the determination of the systems parameter and to represent it’s dynamic in a phase plane plot. On the basis of the qualitative description of the geometrical shape of the phase plot, identification of differences between horseback ridings with two different saddles should be attempted. This could give a hint of the influence of the saddle type to the horse rider system. For application of PCA to horseback riding, the determination of a “pattern vector” composed of the horizontal velocities of a set of 14 body markers was carried out. 13 sound horses were ridden with Side Saddle and English saddle at working walk, trot and canter in a indoor riding hall by the same horseman. The rider provoked the horses by increasing the velocity of different gaits. Motion was recorded with a motion analysis system consisting of six video cameras (sample frequency 120 Hz).

Table 2 shows the mean values of the eigenvalues for each horse according to saddle type. Only the first two components are shown, because the eigenvalues of the 3rd component and those following all fall below 5 %. In trot, the first component dominates in all 13 horses, all other components remain small. Therefore, it can be assumed that the movement coordination of trot can be described as sufficiently accurate with only one systems parameter.

Comparing the eigenvalues of the first component with Wilcoxon tests, no significant differences between Side Saddle (SS) and English Saddle (ES) could be found for walk and canter. For trot, however, a potential influence of the saddle type on the eigenvalue of the first PCA-component became apparent. The English saddle made the first component more pronounced. A more macroscopic speculative explanation might be that at trot horse and rider move in a more coordinated manner - a fact however, which would appear rather counterintuitive to the general equestrian opinion frequently reporting that at trot it is more difficult to maintain a proper and stable sit than at walk or canter.

Tab. 2: Averaged eigenvalues of the first and the second component for walk, trot and canter, separately for Side Saddle (SS) and English Saddle (ES)

Horse	Walk				Trot				Canter			
	1st comp.		2nd comp.		1st comp.		2nd comp.		1st comp.		2nd comp.	
	SS	ES	SS	ES	SS	ES	SS	ES	SS	ES	SS	ES
Ali	62.2	50	31.9	45.6	90	92	5.1	4	57.1	60.7	39.4	36.3
Bla	53	64.4	43	31.2	91.4	84.9	4	7.8	50.3	52.8	47	44.5
Faw	54	53.8	39.8	40.3	82.6	87.6	11.1	6.8	50.9	52.4	45.4	44.4
Fleur	52.8	53	42	42.5	90.6	89.8	5.2	5.1	64.9	53.3	29.1	43.2
Nom	59	58.3	33.6	35.5	81.7	87.7	3.7	8.3	71.6	72.2	25.7	24.8
Ove	52.1	52	41.7	43.1	90.7	92	5.5	4.7	52.2	51.4	45.4	46.3
Roc	52.3	52.5	43.5	43	90.2	91	5	4.7	58.2	60.1	38.9	36.8
Sav	56.2	57.2	37.6	35.8	86.9	87.9	7.4	6.9	55.6	54.3	40.1	41.4
Sca	58.4	52.9	36.5	43	90.7	92.6	4.5	4	57.3	55.5	40.6	42.5
Sch	54.1	50.6	40.4	43.5	89.1	90.2	4.9	5.3	51	53.1	45.8	43.9
Stu	66.7	56.6	23.3	29.5	88.7	89.4	7.7	7.2	52.5	64.4	44.8	45.1
Tri	55.1	53.1	40.5	43.7	92.9	93.4	3.6	3.5	51.6	52.2	46.5	45.8
Xen	56.2	51.4	38.3	42.5	91	90.9	5	5.1	59.3	52.4	37.9	44
$\bar{x}$ $\pm sd$	**58.3 ±4.3**	**54.3 ±3.9**	**37.8 ±5.6**	**39.9 ±5.2**	**89.0 ±3.3**	**89.9 ±2.4**	**5.6 ±2.0**	**6±1, 6**	**56.3 ±6.2**	**56.5 ±6.2**	**40.5 ±6.7**	**41.5 ±5.9**

4 Conclusion

The study shows some examples for practicability of the Principal Component Analysis (PCA) to the movement science. In particular by means of this method the determination of only a few systems parameter helps to characterise the movement coordination. When only one parameter is satisfactory to describe the movement the phase plots of this systems parameter can be used to distinguish between different movement patterns and to calculate the movement variability. Closed curves were obtained for walking and running, because they are cyclic and automated movements. The rehabilitation study shows in which way irregularities influence the gait pattern especially the PCA-components. A qualitative relation between the rehabilitation process and the eigenvalue of the first component could be found. The range or the standard deviation of the systems parameter which was calculated by the Haken-Friedrich-Uhl order parameter analysis is useful to estimate the variability of the movement coordination. This was confirmed by the investigation of the transition from walking to running and the study of changes of running after cycling in triathlon. The PCA presents a useful method to characterise the movement coordination in its entirety. A special application was the analysis of gait in horseback riding. By PCA it was possible to find differences and similarities between walk, trot and canter. In addition, the riding with different saddles shows variations in the movement coordination.

5 References

[1] Haken, H.: *Principles of Brain Functioning*. Berlin, Heidelberg: Springer, 1996.

[2] Kelso, J.A.S.: *Relative timing in brain and behavior: Some observations about the generalized motor program and self-organized coordination dynamics.* Human Movement Science 16 (1997) 453-460.

[3] Sadeghi, H., Allard, P. & Duhaume, M.: *Functional gait asymmetry in able-bodied subjects*. Human Movement Science 16 (1997) 243-258.

[4] Braido, P. & Zhang X.: Quantitative *analysis of finger motion coordination in hand manipulative and gestic acts.* Human Movement Science 22 (2004) 661–678.

[5] Wu, Jianning, Wang, Jue, &Li Liu: Feature extraction via KPCA for classification of gait patterns. Human Movement Science 26 (2007),3, 393-411.

[6] Mah, C. D,. Hulliger, M. & Lee, R. G.: *Quantitative analysis of human movement synergies.* Journal of Motor Behavior, 1994, Vol. 26 (1994) 2, 83-102.

[7] Haas, R.: *Bewegungserkennung und Bewegungsanalyse mit dem Synergetischen Computer*. Aachen: Shaker, 1995.

[8] Chen, H.H. & Saegusa, R.: *The quantification of PCA and NPLCA on change of DF in pedalo locomotion skill learning.* In: 2nd Intern. Congress of Complex Systems in Sport, Madeira, 2008.

[9] Chen, H.H.; Chuang, K.L. & Yang, L.C.: *The structured information for discriminating the table tennis services.* 2nd Intern. Congress of Complex Systems in Sport, Madeira, 2008.

[10] Candler, P.D.: *Selected physiological, kinematical and psychological responses at the walk - to - run and run - to - walk interfaces.* S.A. Journal of Research in Sport Physical Education and Recreation 9 (1986) 2, 15 – 26.

[11] Hreljac A.: *Preferred and energetically optimal walk transition speeds in human locomotion.* Med. Sci. Sports Exerc., 25(10), 1993, 1158-1162.

[12] Diedrich, F.J., Warren W.H. (1995). *Why change gaits? Dynamics of the walk-run transition.* Journal of the Experimental Psychology: Human Perception and Performance, 21(1),1995, 183-202.

[13] Diedrich, F.J., Warren, W.H. (1998). *The dynamics of gait transitions: effects of grade and load*. Journal of Motor Behavior, 30, 1998, 60-78.

[14] Haken, H.: *Synergetics - An Introduction.* Berlin, Heidelberg, New York: Springer, 1983.

[15] Peham C., Licka T., Kapaun M., Scheidl M.: *A new method to quantify harmony of the horse–rider system in dressage.* Sports Engineering, May 2001, vol. 4, no. 2, pp. 95-101.