

Research



Cite this article: Swaminathan K, Tolkova I, Baker L, Arumukhom Revi D, Awad LN, Walsh CJ, Mahadevan L. 2022 A continuous statistical-geometric framework for normative and impaired gaits. *J. R. Soc. Interface* **19**: 20220402.
<https://doi.org/10.1098/rsif.2022.0402>

Received: 26 May 2022
 Accepted: 29 September 2022

Subject Category:
 Life Sciences—Engineering interface

Subject Areas:
 biomechanics, biomathematics

Keywords:
 geometric statistics, gait analysis,
 elastic functional data analysis

Author for correspondence:
 L. Mahadevan
 e-mail: lmahadev@g.harvard.edu

†These authors contributed equally to this study.

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.6251325>.

A continuous statistical-geometric framework for normative and impaired gaits

Krithika Swaminathan^{1,†}, Irina Tolkova^{1,†}, Lauren Baker¹,
 Dheepak Arumukhom Revi^{3,4}, Louis N. Awad^{3,4}, Conor J. Walsh¹ and
 L. Mahadevan^{1,2}

¹School of Engineering and Applied Sciences, and ²Department of Physics, Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA

³College of Health and Rehabilitation Sciences, Sargent College, and ⁴Department of Mechanical Engineering, Boston University, Boston, MA, USA

ID KS, 0000-0001-7570-1408; IT, 0000-0002-2989-7283; DA, 0000-0002-6943-0258;
 LNA, 0000-0002-0159-8011; CJW, 0000-0002-2744-917X; LM, 0000-0002-5114-0519

A quantitative analysis of human gait patterns in space–time provides an opportunity to observe variability within and across individuals of varying motor capabilities. Impaired gait significantly affects independence and quality of life, and thus a large part of clinical research is dedicated to improving gait through rehabilitative therapies. Evaluation of these paradigms relies on understanding the characteristic differences in the kinematics and underlying biomechanics of impaired and unimpaired locomotion, which has motivated quantitative measurement and analysis of the gait cycle. Previous analysis has largely been limited to a statistical comparison of manually selected pointwise metrics identified through expert knowledge. Here, we use a recent statistical-geometric framework, elastic functional data analysis (FDA), to decompose kinematic data into continuous ‘amplitude’ (spatial) and ‘phase’ (temporal) components, which can then be integrated with established dimensionality reduction techniques. We demonstrate the utility of elastic FDA through two unsupervised applications to post-stroke gait datasets. First, we distinguish between unimpaired, paretic and non-paretic gait presentations. Then, we use FDA to reveal robust, interpretable groups of differential response to exosuit assistance. The proposed methods aim to benefit clinical practice for post-stroke gait rehabilitation, and more broadly, to automate the quantitative analysis of motion.

1. Introduction

Human gait biomechanics vary widely based on individual physiological parameters such as age, height, weight and muscle strength [1,2]. Gait can also be affected by neuromotor disorders such as stroke [3] or Parkinson’s disease [4], which can further reduce patient quality of life and community participation [5]. Characterizing gait can provide valuable information for the diagnosis, understanding and treatment of impairment [6], to the extent that gait speed has been termed the ‘sixth vital sign’ [7]. Thus, the development of methods that enable a quantitative analysis of gait, at the population and individual levels, along with associated clinical and biomechanical interpretations, has become an important focus of research in gait diagnostics and rehabilitation.

Common methods for analysing gait presentations from different groups can be broadly categorized into either statistical comparisons of point metrics or dimensionality reduction methods on engineered feature sets. In both cases, the chosen measures encapsulate signal amplitude (the magnitudes of local extrema), signal phase (relative timing within the gait cycle) or a combination of both [8]. These methods, while effective, rely heavily on expert domain knowledge to

both identify metrics of interest and to define significant changes within them [9], which can be labour and time intensive. Moreover, as these methods are confined to a predefined, finite-dimensional space, they are limited in their ability to distinguish between gait characteristics that span across domains, such as multi-joint or spatiotemporal variance. As a result, insight into the biomechanical locomotor strategies used by individuals is limited to the space of chosen metrics. These challenges are further compounded when analysing the effects of an intervention, such as robotic assistance, due to the redundancy in the musculoskeletal system and high heterogeneity of clinical gait and neurophysiology.

In clinical settings, functional measures focus on high-level outcomes such as walking speed, symmetry and capacity, through standardized tests such as the 6-min and 10-m walk tests [10]. These measures require minimal equipment and are thus highly accessible to clinicians. However, there is growing appreciation for studying the geometry of gait through kinematics (e.g. joint angles) and dynamics (e.g. joint torques) and understanding the biomechanical mechanisms underlying a specific gait presentation. For example, people post-stroke often increase the high-level outcome of walking speed through compensatory mechanisms, such as hip circumduction or hip hiking, rather than through increased ankle torque production [11,12]. Such biomechanical measures are typically obtained by using motion capture and force ergometry, in conjunction with inverse dynamics techniques. Both clinical and biomechanical measures have been used to successfully differentiate between gait from clinical and unimpaired populations [8], stratify gait from within a single population [13] and evaluate the efficacy of assistive devices [14].

With access to these kinematic and dynamic measurements, one is naturally led to ask if we can move beyond point metrics to a holistic analysis of gait, using the full time series data. For instance, some studies have directly applied methods such as principal component analysis (PCA) [15] or statistical parametric mapping [16] to gait time series and demonstrated continuous biomechanical differences between groups of participants. However, continuous analysis holds an inherent challenge due to variation in relative timings within the gait cycle. Even when data are time normalized to account for variable stride duration, misalignment in measurements across strides can introduce noise into summary calculations, such as the mean and variance curves. One solution to mitigate this issue is to align events within the gait cycle with algorithms such as dynamic time warping (DTW) before further analysis. Such algorithms stretch and contract the time axis of a signal to optimally match a reference signal [17,18]. Multiple studies have shown that DTW-based analysis can improve stride segmentation [19,20], enable recognition of unimpaired individuals by gait [21] and yield clinically meaningful metrics [22]. However, DTW is designed for pairwise time series alignment, rather than for collections of strides. Moreover, while DTW addresses the challenges of temporal variability within data, it ultimately causes temporal information to be removed from analysis. Finally, this algorithm lacks a rigorous theoretical framework in which desirable properties of alignment, such as symmetry, are met. To mitigate these issues, we look into elastic functional data analysis (FDA), a recent statistical framework under which to perform alignment, termed elastic alignment, and calculate representative templates of

temporally variable data [23–25]. Promising recent work has shown that elastic FDA can enhance statistical analysis of impaired gait, but has yet to be extended for automating extraction of biomechanically interpretable features from gait [26]. In this work, we build on existing literature by integrating ideas from continuous gait analysis, time series alignment and spatiotemporal functional analysis.

Specifically, we demonstrate how elastic alignment can be used alongside common dimensionality reduction techniques to enhance clinical gait analysis at both the population and individual levels, while removing the need for manual feature selection. We conduct this proof of concept using gait kinematics as it both serves as a representation of postural geometry during gait and is an accessible measure that can be estimated through low-cost wearable sensors [27]. First, we show that the decomposition of a signal into its amplitude and phase through elastic FDA, together with PCA, provides interpretable characterizations of unimpaired, post-stroke paretic (more impaired) and post-stroke non-paretic (less impaired) gait presentations. We then show that the integration of elastic FDA with unsupervised clustering methods [28] reveals two types of user responses to gait assistance with a soft exosuit and improves robustness to intra-subject variability relative to conventional methods. We discuss the mathematical background in §2, describe our data collection and pre-processing methods before presenting application-specific details in §3 and present results in §§4 and 5. Finally, we conclude with a discussion on the utility of this approach and future directions in §6.

2. Mathematical background

Elastic FDA presents a theoretical foundation for time series analysis through decomposition into an *amplitude* component, which describes the magnitudes assumed by a function (e.g. joint angles), and a *phase* component, which captures its relative temporal structure [29]. More formally, consider a set of N time series, or functions, $\{f_1, f_2, \dots, f_N\}$, sampled at $T+1$ equally spaced times $t = \{0, 1/T, 2/T, \dots, (T-1)/T, 1\}$. The distortion of a function along the horizontal axis can be formulated by introducing a *warping function* $\gamma(t)$, which represents a transformation between two time axes: $\tilde{t} = \gamma(t)$. Then, time warping can be represented as the composition of an original function f_i with γ , to obtain the warped function $g_i(t) = f_i(\gamma(t))$. Notably, a valid warping function $\gamma(t)$ must satisfy three requirements: (1) it transforms the finite time interval $t \in [0, 1]$, such that $\gamma(0) = 0$ and $\gamma(1) = 1$; (2) warping must be a reversible operation, i.e. γ should be invertible and therefore strictly monotonic; and (3) the function and its inverse γ^{-1} should be smooth to have theoretical guarantees and reduce artifacts.

The problem of pairwise alignment of two functions f_1 and f_2 can be formulated as finding the warping function γ^* which minimizes some cost, $E[f_1, f_2(\gamma)]$, that encodes the desired properties of alignment. Ideally, we would like warping to exhibit inverse symmetry, with optimal alignment of f_2 to f_1 given by $(\gamma^*)^{-1}$. Moreover, γ^* should optimally align $f_1 \circ \tilde{\gamma}$ to $f_2 \circ \tilde{\gamma}$ for any other warping function $\tilde{\gamma}$ to exhibit invariance to simultaneous warping. While a common choice for the cost function E is the Euclidean (\mathbb{L}^2) norm, this cost satisfies neither inverse symmetry nor invariance to simultaneous warping [23]. Instead, we consider the elastic

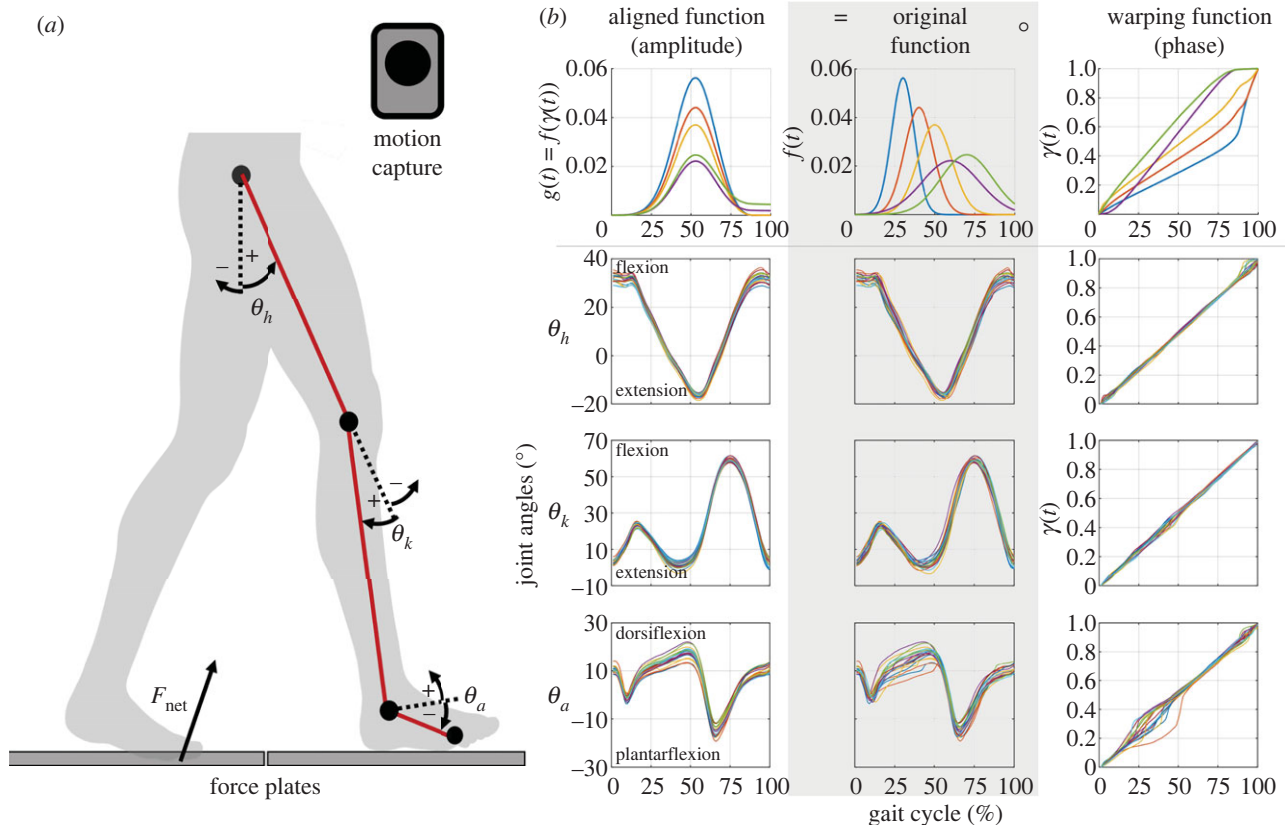


Figure 1. Elastic functional analysis of joint angles. (a) Joint angle conventions for the hip (θ_h), knee (θ_k) and ankle (θ_a). (b) Demonstration of elastic alignment as described in S2. The top row demonstrates the process with Gaussian functions of varying means and standard deviations, and the bottom three rows show example data from the three joints of a representative unimpaired individual walking overground. The middle column (shaded) consists of the original curves, while the left and right columns consist of the aligned functions (amplitudes) and warping functions (phases), respectively. The composition of an original function $f_i(t)$ with the associated warping function $\gamma_i(t)$ yields the aligned function $g_i(t) = f_i(\gamma_i(t))$. We observe improved alignment for this participant's data in certain regions of the gait cycle such as maximum plantarflexion and knee extension.

FDA framework proposed by Srivastava *et al.* [24]. In this framework, we first define the *square-root slope function* operator:

$$q(t) = \text{sign}(f'(t)) \sqrt{|f'(t)|} = \frac{f'(t)}{\sqrt{|f'(t)|}}. \quad (2.1)$$

Conceptually, this function is a normalized representation of the slope $f'(t)$. Then, the corresponding cost function satisfies the desired properties described earlier (see [23] for details):

$$E[f_1, f_2] = \|q_1(t) - q_2(t)\|^2 \quad (2.2)$$

$$= \int_0^1 \left(\frac{f_1'(t)}{\sqrt{|f_1'(t)|}} - \frac{f_2'(t)}{\sqrt{|f_2'(t)|}} \right)^2 dt. \quad (2.3)$$

We can now calculate a mean time series, the *Karcher mean* (\bar{f}^K) [23], that is informed by both amplitude and phase distributions. The iterative calculation of the Karcher mean also yields the corresponding warping functions $\{\gamma_i\}$ for each original function in $\{f_i\}$, such that $f_i(\gamma_i(t))$ are aligned with \bar{f}^K . These warping functions form the phase component of the dataset, while the aligned functions form the amplitudes.

In figure 1, we provide intuition for elastic FDA alongside an overview of the kinematic data used in this work. Specifically, figure 1b (top row) demonstrates amplitude–phase decomposition over a set of Gaussian curves that exhibit variation in the amplitudes along the vertical axis, and in the phases along the horizontal axis. We align a set of functions $\{f_i(t)\}$ (middle column) to obtain the aligned functions $\{f_i(\gamma_i(t))\}$ (left column) and the corresponding warping functions $\{\gamma_i(t)\}$ (right column). A point on the warping function

below the identity line ($\gamma(t) = t$) indicates earlier timing compared to the mean, while a point above the identity line indicates later timing. For example, we see that the blue original function attains the highest and earliest peak; consequently, it has the largest amplitude and the lowest warping function. Conversely, the green and purple functions have similar peak magnitudes but are shifted in time; we see that their amplitudes align closely, while their phases are offset throughout the time interval.

In the context of the current study, the set of time series $\{f_i(t)\}$ represents joint angles across strides, from 0% to 100% of the gait cycle. Figure 1b (rows 2–4) shows the results of the amplitude–phase analysis for the hip, knee and ankle angles during overground walking for a representative unimpaired individual. We observe improved alignment in amplitude, particularly at maximum ankle plantarflexion ($\sim 70\%$ GC) and knee extension ($\sim 50\%$ GC). These regions of large improvements in alignment correspond to regions of high variance in $\{\gamma_i\}$ and indicate sections of the gait cycle that experience high temporal variability across strides, which we can now quantify.

3. Data collection

3.1. Participants

Overground gait data from unimpaired young adults ($N = 10$, 5 female; 26 ± 4 yrs (mean \pm s.d.); 68 ± 17 kg) and from chronic post-stroke patients (>6 months post-stroke; $N = 21$, 9 female; 78 ± 19 yrs; 54 ± 11 kg) were used for this work.

All unimpaired participants reported no previous history of musculoskeletal injury or disease, and all participants provided written informed consent prior to participation. The study procedures were approved by the Harvard Longwood Medical Area and Boston University Institutional Review Boards.

3.2. Experimental protocol

Participants walked around an oval overground track at their self-selected walking speeds for 6 min. Lower limb kinematics and ground reaction forces were measured using optical motion capture (Qualisys Inc., Sweden) and instrumented force plates (Bertec, Columbus, OH), respectively.

Post-stroke participants completed an additional 6-min walk while wearing a mobile, unilateral soft ankle exosuit [30] (figure 4a). The soft exosuit applied assistive stance-phase plantarflexion and swing-phase dorsiflexion torques at the paretic ankle (see [31] for hardware and controller details).

3.3. Software

Inverse dynamics were calculated with biomechanics software, Visual 3D (C-Motion, Germantown, MD). The data were then segmented into strides using heel strikes detected by the force plates and interpolated to 101 points. All ensuing analyses were performed in MATLAB (The MathWorks Inc.). All alignment and Karcher mean calculations were performed using the 'fdasrvf' MATLAB package developed by J. Derek Tucker [32].

3.4. Data pre-processing

We focused on sagittal-plane joint kinematics for this study, using standard joint angle conventions (figure 1a). The range of motion (ROM) of each joint during gait varies widely, with the ankle, knee and hip spanning approximately 30°, 60° and 40°, respectively, in unimpaired individuals [33]. The ROM further varies across people post-stroke depending upon the severity of impairment [34]. Thus, to weight changes at each joint equally, we divided each joint angle curve by a measure of its ROM. For comparison between unimpaired and post-stroke populations (Application I), the average unimpaired ROM was used to normalize the data. For evaluation of individual response to exosuit assistance (Application II), data were normalized by the baseline (no exosuit) ROM for each participant. In the latter application, the normalization aimed to also mitigate any biases from correlations between participant impairment level and their capacity to change gait kinematics. Further detail on application-specific normalization is provided in §§4.1.1 and 5.1.1.

4. Application I: differentiating unimpaired and post-stroke gait

First, we aimed to use this continuous and unsupervised framework to understand the spatial and temporal differences that characterize post-stroke gait kinematics relative to unimpaired gait, using a combination of elastic alignment and PCA.

4.1. Methods

4.1.1. Pre-processing

We analysed data from the right side of the 10 unimpaired participants and the paretic and non-paretic sides of the 21 post-stroke individuals as they walked without any device, for a total of 52 sets of gait measurements, $\{f_{ij}\}$. For each joint j of each set i , we used elastic alignment to calculate the Karcher mean, \bar{f}_{ij}^K , across strides. After normalization by the mean unimpaired joint-level ROM, we concatenated the hip, knee and ankle Karcher means to obtain a single representative gait vector per set. Finally, we aligned all such gait vectors, $\{f_i\}$, yielding an overall \bar{f}_i^K for the dataset, a set of amplitude functions $\{f_i \circ \gamma_i\}$ and a set of warping functions $\{\gamma_i\}$.

4.1.2. Principal component analysis

In this work, we leveraged PCA, an established method for dimensionality reduction [35], to understand structure within the resultant amplitudes and phases. Specifically, we construct a data matrix X in which each column contains a separate observation, subtract the mean \bar{X} and calculate the singular value decomposition of the result using the form $X - \bar{X} = USV^T$. The columns U_p , termed the principal components (PCs), form the basis vectors for the lower-dimensional space, and each data sample i is associated with coefficients V_i , representing coordinates within this space. The diagonal matrix S contains singular values, which represent the relative significance of the components. We separately applied PCA to the aligned amplitudes ($\{f_i \circ \gamma_i\}$) and phases ($\{\gamma_i\}$) to differentiate among unimpaired, paretic and non-paretic gait.

To understand how each component manifests across the gait cycle, we scaled each PC to span the range of coefficients observed in the unimpaired, paretic and non-paretic data, while keeping the other components fixed. Formally, we visualized:

$$g_{C,p,\beta} = \bar{X} + S_p(\bar{V}_{C,p} + \beta \text{SD}[V_{C,p}])U_p, \quad (4.1)$$

where $V_{C,p}$ is a vector containing the coefficients for strides in category C , corresponding to PC p , and $g_{C,p,\beta}$ is the resulting curve for scaling factor β . We considered $C \in \{\text{'unimpaired'}, \text{'paretic'}, \text{'non-paretic'}\}$, $p \in \{1, 2, 3\}$, and chose $\beta \in \{-1, -\frac{1}{2}, 0, \frac{1}{2}, 1\}$ to act over one standard deviation of $V_{C,p}$.

To analyse the phase, we repeated this approach and generated $g_{C,p,\beta}$ for each PC in the space of warping functions $\{\gamma_i\}$. We then composed the unimpaired Karcher mean with $g_{C,p,\beta}$ to understand how variation in the warping functions translated to variation in the gait cycle. Finally, for further visualization of the amplitude and phase PCs, animations of gait defined by $g_{C,p,0}$ for each category and PC are provided in the electronic supplementary material.

4.1.3. Statistical evaluation

The separability of the three categories (unimpaired, paretic and non-paretic) in the PC coefficient space quantifies how much the PCs correspond to the distinguishing characteristics of each category. For each PC, we performed a Kruskal–Wallis test at a significance level of $\alpha = 0.05$ to test for differences in the coefficient distributions across categories. If a significant main effect of category was observed, we conducted *post hoc* pairwise Wilcoxon rank-sum tests with Tukey–Kramer corrections.

4.2. Results

4.2.1. Amplitude

In figure 2a, we applied PCA to the set of amplitudes and found that in the three-dimensional PC coefficient space, the three

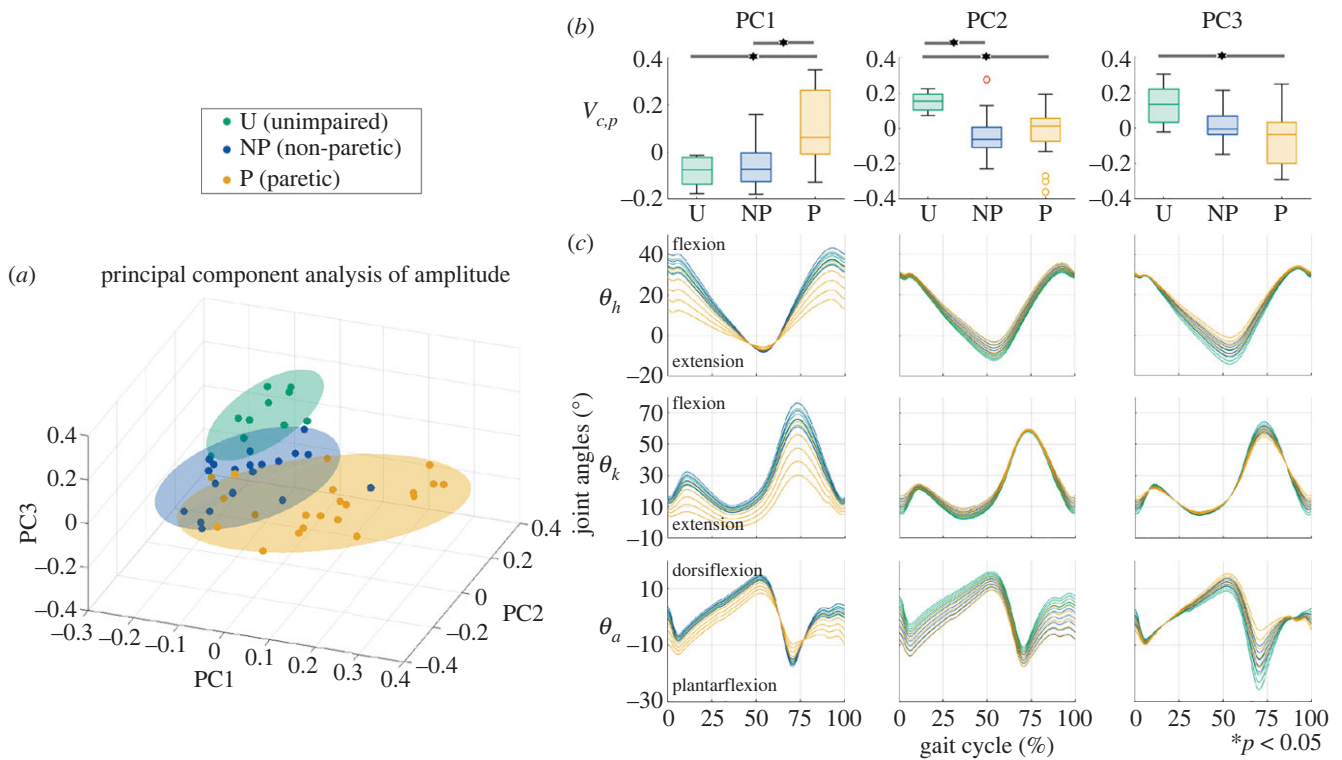


Figure 2. Amplitude-based PCA. (a) Coefficients of the first three PCs obtained by performing PCA on the aligned amplitude curves across the pooled unimpaired (●), non-paretic (●) and paretic (●) data, as described in §4.1.2. The shaded regions represent 2-standard-deviation covariance ellipsoids for each category of gait. (b) Coefficient distributions for the first three PCs across the three categories. Coefficients are significantly different between paretic and unimpaired, between paretic and non-paretic data in PC1 ($p \leq 0.002$), between unimpaired and impaired in PC2 ($p \leq 0.002$) and between unimpaired and paretic in PC3 ($p = 0.001$). (c) Effect of the first three PCs on joint angle time series data for each category, as described in equation (4.1).

categories formed largely separable clusters. The first, second and third PCs accounted for 52%, 20% and 10% of the variation in the data, respectively. Specifically, we found that coefficients of PC1 for paretic strides were significantly different from those for unimpaired ($p = 0.002$) and non-paretic ($p < 0.001$) strides. Similarly, in figure 2b, we see that PC2 separated unimpaired from paretic ($p = 0.002$) and non-paretic ($p < 0.001$) strides, and that PC3 separated unimpaired and paretic strides ($p = 0.001$).

In figure 2c, we isolated the effects of each PC as described in equation (4.1). PC1 underscored the reduced paretic hip and knee flexion across the gait cycle and reduced ankle excursion in both dorsiflexion and plantarflexion, indicating lower ROM in the paretic limb than in the non-paretic or unimpaired limbs. PC2 showed that impaired gait is associated with lower absolute ankle angle across the gait cycle, along with reduced hip extension. Finally, PC3 captured variation in peak ankle plantarflexion and peak hip extension, with the coefficients for non-paretic gait falling between those for unimpaired and paretic gait. For an intuitive picture of how these differences manifest within the gait cycle, we also generated gait animations for these PCA components (see electronic supplementary material).

4.2.2. Phase

In figure 3a, applying PCA to the phase data also led to separability across categories in the PC coefficient space. The first, second and third PCs accounted for 44%, 27% and 8% of the variation in the data, respectively. The coefficients for PC1 were significantly different across all three categories—unimpaired and paretic ($p = 0.025$), unimpaired and non-paretic ($p = 0.029$) and paretic and non-paretic ($p < 0.001$)—with the unimpaired samples occupying a narrow intermediate range, and the paretic and non-paretic samples covering either side. PC2 differentiated unimpaired from impaired data ($p < 0.001$). Finally, PC3 did not provide additional separation between categories.

By examining the gait cycle visualizations, we observed that PC1 extracted differences in the stance-to-swing ratio, with paretic strides spending less time in stance relative to unimpaired data, and the non-paretic strides compensating with increased time in stance. PC2 identified differences in the early-mid stance phase while PC3 represented minor variation in the relative timing of late swing phase. While this third component did not differentiate between categories, we found that paretic strides exhibited higher variation in the corresponding coefficients. Once again, we generated gait animations to visualize these differences within the gait cycle (see electronic supplementary material).

5. Application II: evaluating individual response to exosuit assistance

Functional alignment can also be used to reduce the effects of individual variance and noise during group-level analysis. In this application, we aimed to mitigate the effects of intra-subject variability through alignment and enable unsupervised stratification of user response to exosuit assistance with clustering. We posited that this approach would (i) extract distinct categories of user response to exosuit assistance from geometric representations of their gait and (ii) result in improved robustness compared to analysis without alignment. A summary of the methods is shown in figure 4a.

5.1. Methods

5.1.1. Pre-processing

We analysed data from the paretic limb as post-stroke participants walked with (EXO₊) and without (EXO₋) the soft ankle exosuit, for a total of 42 sets of gait measurements, $\{f_{ij}\}$. As in

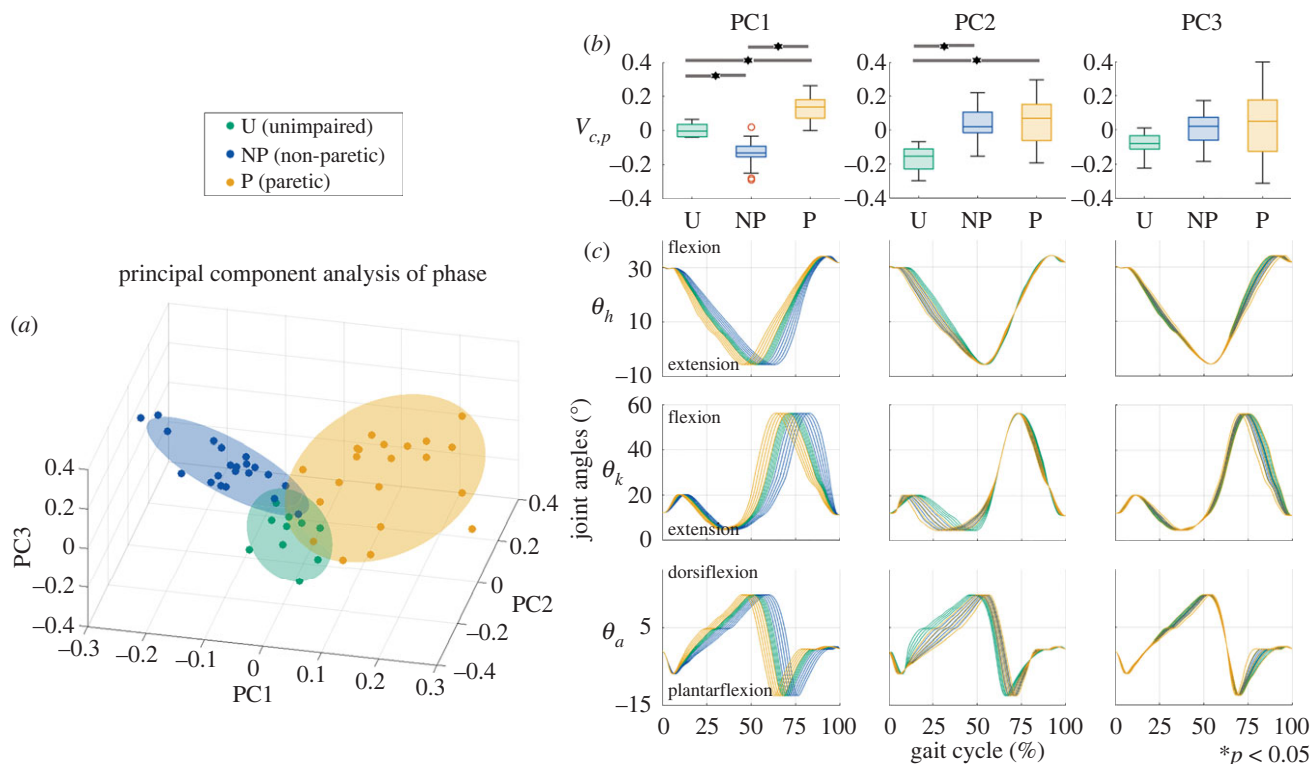


Figure 3. Phase-based PCA. (a) Coefficients of the first three PCs obtained by performing PCA on the warping curves γ_i across the pooled unimpaired (●), non-paretic (●) and paretic (●) data, as described in §4.1.2. The shaded regions represent 2-standard-deviation covariance ellipsoids for each category. (b) Coefficient distributions for the first three PCs across the three categories. Coefficient distributions were different across all three categories in PC1 ($p \leq 0.029$) and between the unimpaired and impaired categories in PC2 ($p \leq 0.001$). (c) Effect of the first three phase PCs on joint kinematics for each category, as described in equation (4.1).

Application I, for each joint j of each set i , we used elastic alignment to calculate the Karcher mean, $\overline{f_{ij}^K}$, across strides. We then normalized each element of $\overline{f_{ij}^K}$ by the ROM observed in joint j during EXO₋ for the corresponding subject. The normalized Karcher means were concatenated across joints to generate a single representative gait vector per set $\overline{f_i^K}$. Finally, we generated a difference curve between EXO₊ and EXO₋ gait, $\delta_s = \overline{f_{s,EXO_+}} - \overline{f_{s,EXO_-}}$, to serve as the input vector for each subject s .

5.1.2. Clustering

Clustering is an unsupervised learning paradigm that separates a set of observations, in our case $\{\delta_s\}$, into distinct subgroups. Here, we applied agglomerative hierarchical clustering as it is deterministic for a given set of data and thus is more robust for small datasets. With this method, each observation begins in its own subgroup, after which the two closest subgroups are merged, and the process continues until all observations are placed in the same subgroup.

We used Euclidean distance for comparing observations, and Ward's linkage method [36] for computing inter-cluster distances d_{ab} :

$$d_{ab} = \frac{n_a n_b}{n_a + n_b} \|\mu_a - \mu_b\|^2,$$

where μ_a and μ_b are the means of clusters a and b , which contain n_a and n_b elements, respectively. d_{ab} quantifies the additional variance resulting from merging clusters a and b . Thus, this distance minimizes the squared total within-cluster variance at each merging step.

To define the optimal number of subgroups, we maximized the silhouette coefficient (SC) [37], a measure of the mean inter-cluster ($\bar{\mu}$) to intra-cluster (\bar{x}) distance ratio:

$$SC = \frac{\bar{\mu} - \bar{x}}{\max(\bar{\mu}, \bar{x})}.$$

We evaluated SC for two to six clusters to determine the final number of subgroups.

We then leveraged the ability of clustering to detect outliers in our data to reduce the sensitivity of the algorithm [38]. We used an automated iterative process to identify outliers as observations that were clustered alone, i.e. a difference curve δ_s was an outlier if the distance between δ_s and every member of $\{\delta_r\}$ for $r \neq s$ was greater than the distance between all pairs of elements in $\{\delta_r\}$. We removed observations until the smallest generated cluster contained at least two samples. Finally, we applied a hierarchical clustering algorithm to the resulting cleaned dataset of normalized difference curves to investigate categories of response to exosuit assistance.

5.1.3. Evaluation

Biomechanics We hypothesized that stroke survivors would use the combination of plantarflexion and dorsiflexion exosuit assistance to increase gait speed with a reliance on hip circumduction, a common compensatory mechanism, maintain walking speed with reduced circumduction, or both increase gait speed and reduce circumduction [39,40]. We further hypothesized that clustering on $\{\delta_s\}$ would elucidate these categories of gait response. Thus, we evaluated changes in gait kinematics, speed, joint torques and hip circumduction. The net change in joint kinematics due to exosuit

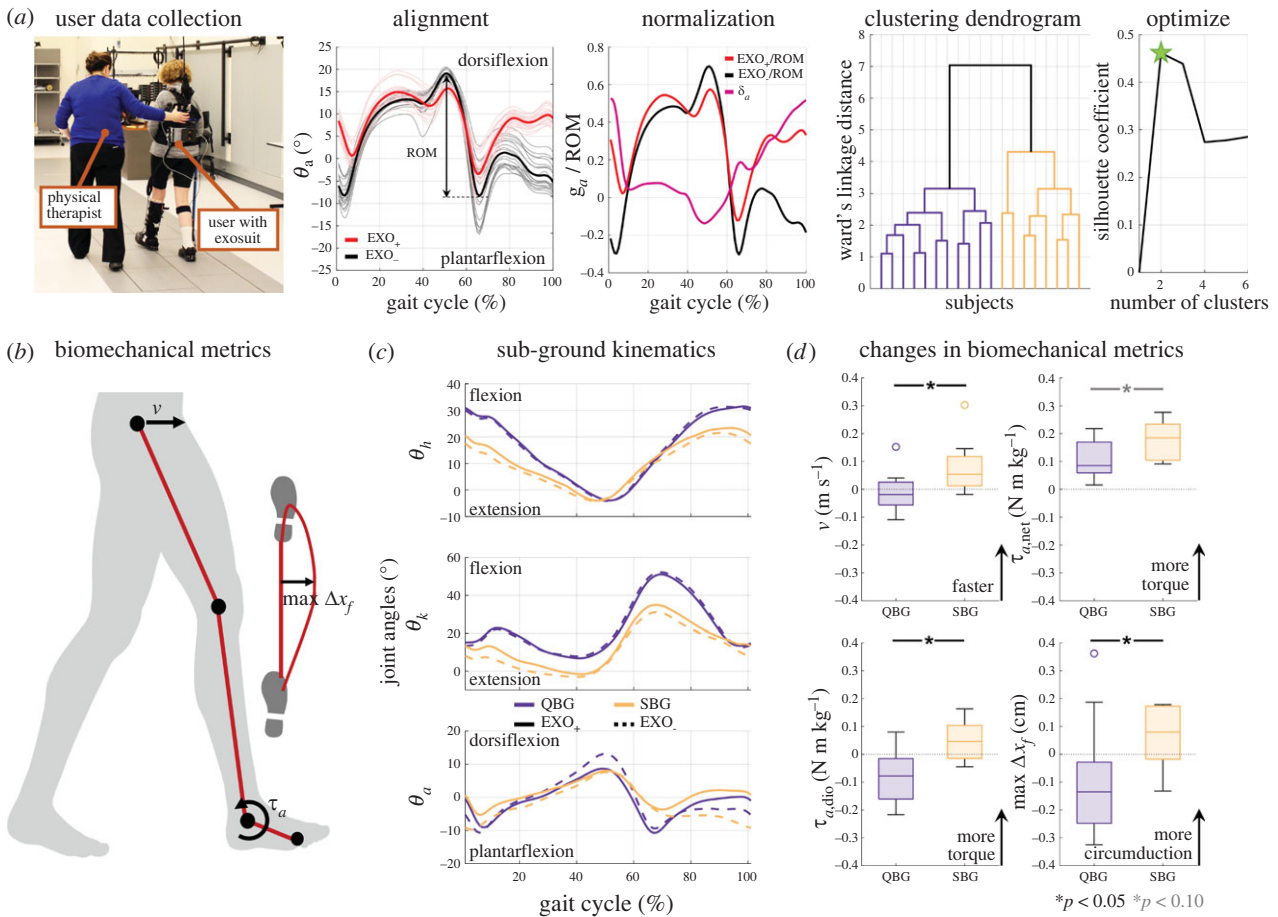


Figure 4. Effects of exosuit assistance on post-stroke gait biomechanics using elastic functional data analysis (FDA). (a) Collection and processing methods for evaluating individual response to exosuit assistance (ankle data for one subject shown). People post-stroke walked along an overground path at their comfortable walking speeds with (EXO₊) and without (EXO₋) a unilateral soft assistive ankle exosuit, while optical motion capture, force plate and exosuit sensor data were collected. For each subject, kinematics for each joint were aligned using FDA to generate a representative curve of EXO₊ and EXO₋ gait, then divided by the EXO₋ range of motion and finally subtracted from each other to obtain a set of difference curves $\{\delta_j\} = \text{EXO}_+ - \text{EXO}_-$ for clustering (see §5.1.2). Two clusters resulted from this process. (b) Metrics used in evaluating dynamic gait quality (see §5.1.3). People post-stroke typically present with low ankle torque (τ_a), low walking speed (v) and high circumduction (Δx_f) [34]. (c) Average joint angle time series for the quality-based group (QBG) and speed-based group (SBG) during EXO₊ and EXO₋. The magnitude of change between EXO₊ and EXO₋ differs between the two clusters, with the largest distinction at the ankle joint. (d) Changes between EXO₊ and EXO₋ in walking speed, peak parietic net and biological ankle torque, and peak hip circumduction in QBG and SBG. The clusters show significant ($p < 0.05$) differences in their response to exosuit assistance. Although both groups increased net ankle torque, the two groups did not significantly differ in the magnitude of increase ($p = 0.083$). SBG also increased in biological ankle torque with an associated increase in walking speed. Conversely, QBG reduced biological ankle torque without any change in walking speed, but reduced hip circumduction (see §5.2.1).

assistance was defined as the mean absolute value of $\delta_{j,s}$ for each joint j and subject s . Net ankle torque ($\tau_{a,\text{net}}$) was computed through inverse dynamics, and biological ankle torque ($\tau_{a,\text{bio}}$) was calculated as the difference between $\tau_{a,\text{net}}$ and exosuit-applied torque [41]. Maximum hip circumduction (Δx_f) was measured from lateral movement of the heel marker. A decrease in Δx_f indicates an improvement in the quality of gait. Given the small sample size, we used the non-parametric Kruskal–Wallis test [42] to check for significant differences in user response to assistance across the resultant clusters, with a significance level of $\alpha = 0.05$.

Robustness To evaluate whether using alignment influenced sensitivity to intra-subject stride-by-stride variability, we used the evaluation score introduced by Gloumakov *et al.* [43]. Briefly, the evaluation score assumes that observations from one individual are likely to be more similar to each other than to observations from any other individual. The score is the percentage of pairs of intra-subject strides that are grouped together when clustering on all strides from all subjects. For example, the evaluation score with

one cluster is 100%, as all strides are placed in the same group. As the number of clusters increases, the evaluation score decreases monotonically. This score can then be used to compare the performance of different clustering algorithms up to k clusters, where k is the number of individuals in the dataset.

We reran the clustering process using individual strides from each subject and condition, with and without alignment. Since the number of strides varied by condition and subject, we first constructed the set of all possible difference curves for each subject using every pairwise combination of EXO₊ and EXO₋ strides. To avoid biasing towards individuals with more strides, we randomly selected n curves for each subject, where $n = 15$ was the size of the smallest set of generated difference curves. We repeated this process 25 times to get a distribution of evaluation scores for 1 to 21 clusters. We further used this metric to compare the performance of hierarchical versus two common non-hierarchical clustering algorithms, k -means and k -medoids, with and without alignment.

5.2. Results

5.2.1. Biomechanics

As hypothesized, clustering the aligned difference curves as described in §5.1.2 resulted in separable clusters based on exosuit-driven changes in sagittal-plane gait kinematics (figure 4a). Two subjects were removed as outliers and two clusters, with 8 and 11 subjects, respectively, were identified from the cleaned dataset. The two groups adopted different functional strategies in response to the exosuit assistance, one that increased walking speed, hereafter referred to as the speed-based group (SBG), and one that reduced hip circumduction, hereafter referred to as the quality-based group (QBG). Individuals in SBG showed significantly higher increases in comfortable walking speed ($p=0.026$) and peak $\tau_{a,bio}$ ($p=0.013$) than in QBG (figure 4d). Specifically, with exosuit assistance, individuals in SBG increased walking speed by $0.08 \pm 0.04 \text{ m s}^{-1}$ (mean \pm standard error; $13.5 \pm 6\%$), which is between the thresholds for small and substantial meaningful changes in the literature [44]. Meanwhile, those in QBG showed negligible changes of $-0.01 \pm 0.02 \text{ m s}^{-1}$ ($-0.8 \pm 2\%$). Although both SBG and QBG increased peak $\tau_{a,net}$ by $0.18 \pm 0.03 \text{ N m kg}^{-1}$ ($30.1 \pm 5.9\%$) and $0.11 \pm 0.02 \text{ N m kg}^{-1}$ ($11.4 \pm 3.1\%$), respectively, the magnitude of change was not significantly different between groups ($p=0.083$). However, peak $\tau_{a,bio}$ increased by $0.04 \pm 0.03 \text{ N m kg}^{-1}$ ($9.5 \pm 4.1\%$) in SBG, but decreased by $0.08 \pm 0.03 \text{ N m kg}^{-1}$ ($-5.2 \pm 2.7\%$) in QBG. Conversely, subjects in QBG demonstrated larger improvements in hip circumduction than those in SBG ($p=0.039$). Specifically, while maximum hip circumduction worsened by $6.4 \pm 4 \text{ mm}$ ($20.9 \pm 12.8\%$) in SBG, it improved by $9.4 \pm 6.3 \text{ mm}$ ($16.6 \pm 12.6\%$) in QBG, both of which are of the same order of magnitude of change observed in past exosuit studies [39]. These results suggest that (i) this approach may differentiate between subcategories of response to external assistance and (ii) individuals may use assistance to improve high-level gait function or low-level gait quality.

Subjects in SBG also showed significantly more kinematic deviation from EXO₋ during EXO₊ than those in QBG at the ankle ($p<0.001$) and knee ($p=0.002$) joints, but not at the hip ($p=0.283$) (figure 4c). The average absolute change of the normalized joint angle (degrees/subject-specific ROM) was 0.19 ± 0.02 , 0.10 ± 0.01 and 0.10 ± 0.02 at the ankle, knee and hip, respectively, for those in SBG, and 0.10 ± 0.01 , 0.05 ± 0.01 and 0.07 ± 0.01 for those in QBG. Rescaling by the original normalization factors, these changes approximately corresponded to $5.0^\circ \pm 1.9^\circ$, $5.0^\circ \pm 2.3^\circ$ and $3.4^\circ \pm 1.7^\circ$ for SBG, and $2.5^\circ \pm 1.1^\circ$, $2.2^\circ \pm 1.1^\circ$ and $2.7^\circ \pm 1.9^\circ$ for QBG. Finally, we applied the FDA framework to align kinematics during EXO₊ and EXO₋ for each subject and investigated the resultant warping functions for the two groups (electronic supplementary material, figure S1). We observed negligible differences between the temporal responses of the two groups. However, both showed the largest changes in phase at the ankle joint, followed by the knee joint and finally the hip joint.

5.2.2. Robustness

The mean evaluation score was higher when using aligned gait data compared to the original data across the range of clusters tested (electronic supplementary material, figure S2). Moreover, the magnitude of improvement increased as the number of clusters grew, and at 21 clusters, we obtained

a mean evaluation score of 78% with FDA versus 65% without. The difference in evaluation score was also present when using non-hierarchical clustering algorithms, with FDA leading to improvements at 21 clusters of 12% in both k -means and k -medoids.

Hierarchical clustering contributed towards increased robustness compared to using k -means and k -medoids, across all numbers of clusters tested (electronic supplementary material, figure S2). At two clusters, both k -means and k -medoids achieved evaluation scores of 89% with FDA, while hierarchical clustering achieved 92%. With 21 clusters and alignment, the scores of k -means and k -medoids were 68% and 70%, respectively, versus 78% with our approach.

6. Discussion

In this work, we have described a framework for gait analysis integrating elastic alignment with conventional dimensionality reduction techniques, and presented results from two complementary applications for post-stroke gait. In our first application, we used elastic alignment followed by PCA to understand the primary modes of gait variation in both the spatial (amplitude) and temporal (phase) dimensions, demonstrating the advantages of continuous, unsupervised analysis. In our second application, we combined alignment with hierarchical clustering to reveal stratification of user response to gait assistance with a soft exosuit and demonstrated robustness to individual variation. Within both applications, this approach yielded biomechanical interpretations that align with and build upon prior work.

Unlike most conventional approaches, this framework automatically leverages information from the full gait cycle while remaining biomechanically interpretable. Specifically, we found that the principal modes of variation in our dataset of unimpaired, paretic and non-paretic gait corresponded to statistically significant differences across the three categories, in both spatial and temporal domains. These results are consistent with recent work that employed a similar approach to identify statistically significant amplitude and phase differences between the centre-of-pressure trajectories of unimpaired individuals and patellofemoral pain syndrome patients [26]. Notably, the gait characteristics corresponding to regions of high spatiotemporal variance in our data were similar to manually selected features and point metrics from prior clinical studies. Analysis of amplitude PC1 was consistent with prior work showing that peak hip flexion, knee flexion at toe-off and peak knee flexion during swing are reduced in the paretic limb due to muscular weakness and low propulsion in late stance [45,46]. Similarly, PC2 of the amplitudes identified knee hyperextension, a common characteristic of post-stroke gait, as a differentiating factor [46]. Phase PCA also extracted a key temporal difference between paretic and non-paretic gait, i.e. the paretic limb spends less time in stance relative to unimpaired gait, while the non-paretic limb compensates with higher stance time. This observation has been attributed clinically to muscle weakness and reduced stability in this population [46]. From phase PC2, we found a slower rise in ankle angle in people post-stroke during early stance (~ 10 – 20% GC), which aligns with observations of reduced tibial progression attributable to reduced walking speed and knee extensor strength [47]. For comparison, we also applied PCA without alignment by using standard

arithmetic means instead of Karcher means for each participant (electronic supplementary material, figure S4). Although we still observed statistical separation, the components were difficult to interpret due to the interactions of spatial and temporal characteristics. While PCA with alignment could also be extended to examine amplitude–phase interactions through analysis of correlation in principal coefficients, we chose to focus on interpretation of individual components and leave more complex analysis of pairwise interactions for future work. These results suggest the potential for this framework to automate identification of key spatiotemporal point metrics, thereby bridging the gap with conventional techniques.

We also demonstrated the utility of FDA with continuous time series gait data at the individual level. We introduced the concept of difference curves to cluster on *changes* in joint kinematic trajectories between EXO₊ and EXO₋ gait, thereby capturing characteristics of user strategy in response to external assistance. While hierarchical clustering has been used to successfully subgroup post-stroke joint kinetics [13] and upper limb motions in unimpaired individuals [48], here we extended these approaches to evaluate user response to exosuit assistance and used a continuous representation of change for this analysis. Similar to our first application, our geometric approach to representing gait enabled us to account for variations in posture of the user throughout the gait cycle without the need for feature engineering. Interestingly, despite only using sagittal-plane geometry as the input to the algorithm, we observed significant differences in both dynamic and out-of-plane measures, which then allowed for functional subgrouping of individuals into quality-based and speed-based groups. For a comparison, we considered whether clustering on changes in walking speed alone would yield informative clusters, as increased speed is both a desired outcome and a factor that may affect gait kinematics [49]. However, we found that while this approach resulted in groups that differed in their walking speed response, no statistically significant differences were found in other biomechanical metrics (electronic supplementary material, figure S3), suggesting that using the full time series data could offer a more holistic understanding of user response. Unlike the group-level analysis of the first application, comparing phase information across subgroups did not provide much additional insight. Given that the magnitude of change in phase can be two orders smaller than the time period of a stride (~1–2%GC), it is possible that this dataset was too small to identify changes at that resolution.

An additional strength of our approach is its capacity to reduce stride-to-stride variability in data during pre-processing. The observed improvements in robustness in our application are consistent with those observed by Gloumakov *et al.*, who used a similar alignment step before clustering upper limb motions [43]. This outcome is particularly beneficial for clinical applications in which the signal-to-noise ratio is low, hindering the ability to find structure across data samples. Thus, we expect that this approach will provide most utility in unsupervised investigations of high-dimensional data, such as in the case of gait, where domain knowledge may be less accessible. Furthermore, as the field largely recognizes that each individual is unique and that a single model cannot capture population-level behaviour, there is a need for methods that identify individual locomotor strategies. The approach presented here may begin to enable

a theoretical method for simultaneously mitigating the effects of variation in the data while also capturing salient features of the data across strides.

This work is a proof of concept that demonstrates the feasibility of the proposed continuous and unsupervised framework for gait analysis. Still, there are a few considerations that future work must account for to enable broader translation. First, the small sample size of the dataset used in this work limits the generalizability of the proposed method, and additional investigation is needed to understand and quantify the tradeoffs between dataset size and validity of the method. With more clinical trials with assistive and rehabilitative devices underway, the field can soon begin to evaluate whether similar techniques can be applied to these larger datasets to obtain separable clusters of response. Moreover, all strides corresponding to a single subject were collected in one experimental visit, and thus we expect future work will evaluate the feasibility of these methods for longitudinal applications, such as categorizing recovery trajectories during rehabilitation. Longitudinal studies would further provide data to address the inability of the current work to separate the effects of age and neuromotor impairment. Finally, we note that while we opted to focus on sagittal-plane gait kinematics given their rising accessibility in real-world environments, these methods are applicable to complex functions and thus can be explored with neurophysiological signals such as electromyography. Hence, we expect that future work will focus on expanding the applications of elastic alignment combined with dimensionality reduction to incorporate larger datasets, longer timescales and more complex physiological input.

In summary, this article presents the use of elastic alignment alongside unsupervised dimensionality reduction methods to characterize post-stroke gait and understand the response of stroke survivors to soft exosuit assistance. Our work aims to enable improved understanding of impaired gait, inform the evaluation of assistive devices and ultimately advance physical rehabilitation for individuals with motor impairment.

Ethics. All participants provided written informed consent prior to participation. The study procedures were approved by the Harvard Longwood Medical Area and Boston University Institutional Review Boards.

Data accessibility. All data needed to support the conclusions of this paper are included in the main text or electronic supplementary material. Source data are available in the electronic supplementary material.

Authors' contributions. K.S.: conceptualization, data curation, formal analysis, investigation, methodology, validation, visualization, writing—original draft, writing—review and editing; I.T.: conceptualization, data curation, formal analysis, investigation, software, validation, visualization, writing—original draft, writing—review and editing; L.B.: investigation, methodology; D.A.R.: investigation, methodology; L.N.A.: funding acquisition, resources, supervision, writing—review and editing; C.J.W.: funding acquisition, resources, supervision, writing—review and editing; L.M.: conceptualization, funding acquisition, project administration, resources, supervision, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. Patents describing the exosuit components documented in this article have been filed with the US Patent Office of which C.J.W. is an inventor. Harvard University has entered into a licensing and collaboration agreement with ReWalk Robotics. C.J.W. is a paid consultant for ReWalk Robotics.

Funding. This work was supported by the National Institute of Health under Award nos. BRG-R01HD088619 and R21 AR076686-01A1, the National Science Foundation under Award no. CMMI-1925085, the Harvard University John A. Paulson School of Engineering and Applied Sciences and the Wyss Institute for Biologically Inspired Engineering. I.T. is supported by the NSF-Simons Center for Mathematical and Statistical Analysis of Biology at Harvard (Award no. 1764269) and the Harvard Quantitative Biology Initiative. K.S. and L.B. are partially supported by the National Science

Foundation Graduate Research Fellowship (grant no. DGE1650114). D.A.R. and L.N.A. were partially supported by the American Heart Association (grant nos. 18TPA34170171 and 18IPA34170487). L.M. was partially supported by the Simons Foundation and the Henri Seydoux Fund.

Acknowledgements. The authors thank Dr Jaehyun Bae, Dr Lizeth Sloat, Dr Franchino Porciuncula, Teresa Baker, Regina Sloutsky, Dabin Choe, Christopher Siviyy, Nicolas Menard, Michael Rouleau and Lexine Schumm for their assistance in data collection and processing.

References

- Samson MM, Crowe A, de Vreede PL, Dessens JAG, Duursma SA, Verhaar HJJ. 2001 Differences in gait parameters at a preferred walking speed in healthy subjects due to age, height and body weight. *Aging Clin. Exp. Res.* **13**, 16–21. (doi:10.1007/BF03351489)
- Buchner DM, Larson EB, Wagner EH, Koepsell TD, de Lateur BJ. 1996 Evidence for a non-linear relationship between leg strength and gait speed. *Age Ageing* **25**, 386–391. (doi:10.1093/ageing/25.5.386)
- Patterson SL, Forrester LW, Rodgers MM, Ryan AS, Ivey FM, Sorkin JD, Macko RF. 2007 Determinants of walking function after stroke: differences by deficit severity. *Arch. Phys. Med. Rehabil.* **88**, 115–119. (doi:10.1016/j.apmr.2006.10.025)
- Tan D, Danoudis M, McGinley J, Morris ME. 2012 Relationships between motor aspects of gait impairments and activity limitations in people with Parkinson's disease: a systematic review. *Park. Relat. Disord.* **18**, 117–124. (doi:10.1016/j.parkreldis.2011.07.014)
- Rand D, Eng JJ, Tang PF, Hung C, Jeng JS. 2010 Daily physical activity and its contribution to the health-related quality of life of ambulatory individuals with chronic stroke. *Health Qual. Life Outcomes* **8**, 80. (doi:10.1186/1477-7525-8-80)
- Buckley C, Alcock L, McArdle R, Rehman R, Del Din S, Mazzà C, Yarnall A, Rochester L. 2019 The role of movement analysis in diagnosing and monitoring neurodegenerative conditions: insights from gait and postural control. *Brain Sci.* **9**, 34. (doi:10.3390/brainsci9020034)
- Fritz S, Lusardi M. 2009 White paper: 'walking speed: the sixth vital sign'. *J. Geriatr. Phys. Ther.* **32**, 46–49. (doi:10.1519/00139143-200932020-00002)
- Filli L *et al.* 2018 Profiling walking dysfunction in multiple sclerosis: characterisation, classification and progression over time. *Sci. Rep.* **8**, 4984. (doi:10.1038/s41598-018-22676-0)
- Kesar TM, Binder-Macleod SA, Hicks GE, Reisman DS. 2011 Minimal detectable change for gait variables collected during treadmill walking in individuals post-stroke. *Gait Posture* **33**, 314–317. (doi:10.1016/j.gaitpost.2010.11.024)
- Cheng DK, Nelson M, Brooks D, Salbach NM. 2020 Validation of stroke-specific protocols for the 10-m walk test and 6-min walk test conducted using 15-m and 30-m walkways. *Top. Stroke Rehabil.* **27**, 251–261. (doi:10.1080/10749357.2019.1691815)
- Wonsetler EC, Bowden MG. 2017 A systematic review of mechanisms of gait speed change post-stroke. Part 2: exercise capacity, muscle activation, kinetics, and kinematics. *Top. Stroke Rehabil.* **24**, 394–403. (doi:10.1080/10749357.2017.1282413)
- Wonsetler EC, Bowden MG. 2017 A systematic review of mechanisms of gait speed change post-stroke. Part 1: spatiotemporal parameters and asymmetry ratios. *Top. Stroke Rehabil.* **24**, 435–446. (doi:10.1080/10749357.2017.1285746)
- Kinsella S, Moran K. 2008 Gait pattern categorization of stroke participants with equinus deformity of the foot. *Gait Posture* **27**, 144–151. (doi:10.1016/j.gaitpost.2007.03.008)
- Shi B, Chen X, Yue Z, Yin S, Weng Q, Zhang X, Wang J, Wen W. 2019 Wearable ankle robots in post-stroke rehabilitation of gait: a systematic review. *Front. Neurobotics* **13**, 63. (doi:10.3389/fnbot.2019.00063)
- Deluzio K, Astephen J. 2007 Biomechanical features of gait waveform data associated with knee osteoarthritis: an application of principal component analysis. *Gait Posture* **25**, 86–93. (doi:10.1016/j.gaitpost.2006.01.007)
- Pataky TC, Robinson MA, Vanrenterghem J. 2013 Vector field statistical analysis of kinematic and force trajectories. *J. Biomech.* **46**, 2394–2401. (doi:10.1016/j.jbiomech.2013.07.031)
- Berndt DJ, Clifford J. 1994 Using dynamic time warping to find patterns in time series. In *AAAI-94: Proc. 3rd Int. Conf. on Knowledge Discovery and Data Mining, Seattle, WA, USA*, pp. 359–370.
- Keogh EJ, Pazzani MJ. 2001 Derivative dynamic time warping. In *Proc. 2001 SIAM Int. Conf. on Data Mining, Chicago, IL, USA, 5–7 April 2001*, pp. 1–11. (doi:10.1137/1.9781611972719.1)
- Helwig NE, Hong S, Hsiao-Weckler ET, Polk JD. 2011 Methods to temporally align gait cycle data. *J. Biomech.* **44**, 561–566. (doi:10.1016/j.jbiomech.2010.09.015)
- Haji Ghassemi N, Hannink J, Martindale C, Gaßner H, Müller M, Klucken J, Eskofier B. 2018 Segmentation of gait sequences in sensor-based movement analysis: a comparison of methods in Parkinson's disease. *Sensors* **18**, 145. (doi:10.3390/s18010145)
- Boulgouris NV, Plataniotis KN, Hatzinakos D. 2004 Gait recognition using dynamic time warping. In *IEEE 6th Workshop on Multimedia Signal Processing, Siena, Italy, 29 September–1 October 2004*, pp. 263–266. (doi:10.1109/MMSP.2004.1436543)
- Bonnyaud C, Roche N, Van Hamme A, Bensmail D, Pradon D. 2016 Locomotor trajectories of stroke patients during oriented gait and turning. *PLoS ONE* **11**, e0149757. (doi:10.1371/journal.pone.0149757)
- Srivastava A, Klassen EP. 2016 Functional data and elastic registration. In *Functional and shape data analysis*, pp. 73–123. New York, NY: Springer. (doi:10.1007/978-1-4939-4020-2_4)
- Srivastava A, Wu W, Kurtek S, Klassen E, Marron JS. 2011 Registration of functional data using Fisher-Rao metric. (<https://arxiv.org/abs/1103.3817>)
- Tucker JD. 2017 Elastic functional data analysis. R Package Version. 1.
- Pataky TC, Robinson MA, Vanrenterghem J, Donnelly CJW. 2022 Simultaneously assessing amplitude and temporal effects in biomechanical trajectories using nonlinear registration and statistical nonparametric mapping. *J. Biomech.* **136**, 111049. (doi:10.1016/j.jbiomech.2022.111049)
- Desmarais Y, Mottet D, Slangen P, Montesinos P. 2021 A review of 3D human pose estimation algorithms for markerless motion capture. *Comput. Vis. Image Underst.* **212**, 103275. (doi:10.1016/j.cviu.2021.103275)
- Omran MGH, Engelbrecht AP, Salman A. 2007 An overview of clustering methods. *Intell. Data Anal.* **11**, 583–605. (doi:10.3233/IDA-2007-11602)
- Marron JS, Ramsay JO, Sangalli LM, Srivastava A. 2015 Functional data analysis of amplitude and phase variation. *Stat. Sci.* **30**, 468–484. (doi:10.1214/15-ST5524)
- Awad LN *et al.* 2017 A soft robotic exosuit improves walking in patients after stroke. *Sci. Transl. Med.* **9**, eaai9084. (doi:10.1126/scitranslmed.aai9084)
- Bae J *et al.* 2018 A lightweight and efficient portable soft exosuit for paretic ankle assistance in walking after stroke. In *2018 IEEE Int. Conf. on Robotics and Automation (ICRA), Brisbane, Australia, 21–25 May 2018*, pp. 2820–2827. (doi:10.1109/ICRA.2018.8461046)
- Tucker JD. 2013 fdasrv documentation. See https://github.com/jdtuck/fdasrvf_MATLAB.
- Perry J. 1992 Normal gait. In *Atlas of limb prosthetics: surgical, prosthetic, and rehabilitation principles*, 2nd edn (eds JH Bowker, JW Michael), ch. 13. Rosemont, IL: American Academy of Orthopedic Surgeons.

34. Olney SJ, Richards C. 1996 Hemiparetic gait following stroke. Part I: characteristics. *Gait Posture* **4**, 136–148. (doi:10.1016/0966-6362(96)01063-6)
35. Abdi H, Williams LJ. 2010 Principal component analysis. *Wiley Interdiscip. Rev. Comput. Stat.* **2**, 433–459. (doi:10.1002/wics.101)
36. Ward Jr JH. 1963 Hierarchical grouping to optimize an objective function. *J. Am. Stat. Assoc.* **58**, 236–244. (doi:10.1080/01621459.1963.10500845)
37. Rousseeuw PJ. 1987 Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53–65. (doi:10.1016/0377-0427(87)90125-7)
38. García-Escudero LA, Gordaliza A, Matrán C, Mayo-Isacar A. 2010 A review of robust clustering methods. *Adv. Data Anal. Classif.* **4**, 89–109. (doi:10.1007/s11634-010-0064-5)
39. Awad LN, Bae J, Kudzia P, Long A, Hendron K, Holt KG, Et A. 2017 Reducing circumduction and hip hiking during hemiparetic walking through targeted assistance of the paretic limb using a soft robotic exosuit. *Am. J. Phys. Med. Rehabil.* **96**, S157–S164. (doi:10.1097/PHM.0000000000000800)
40. Awad LN, Kudzia P, Revi DA, Ellis TD, Walsh CJ. 2020 Walking faster and farther with a soft robotic exosuit: implications for post-stroke gait assistance and rehabilitation. *IEEE Open J. Eng. Med. Biol.* **1**, 108–115. (doi:10.1109/OJEMB.2020.2984429)
41. Yandell MB, Quinlivan BT, Popov D, Walsh C, Zelik KE. 2017 Physical interface dynamics alter how robotic exosuits augment human movement: implications for optimizing wearable assistive devices. *J. Neuroeng. Rehabil.* **14**, 40. (doi:10.1186/s12984-017-0247-9)
42. Kruskal WH, Wallis WA. 1952 Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* **47**, 583–621. (doi:10.1080/01621459.1952.10483441)
43. Gloumakov Y, Spiers AJ, Dollar AM. 2020 Dimensionality reduction and motion clustering during activities of daily living: 3, 4, and 7 degree-of-freedom arm movements. *IEEE Trans. Neural Syst. Rehabil. Eng.* **28**, 2826–2836. (doi:10.1109/TNSRE.2020.3040522)
44. Perera S, Mody SH, Woodman RC, Studenski SA. 2006 Meaningful change and responsiveness in common physical performance measures in older adults. *J. Am. Geriatr. Soc.* **54**, 743–749. (doi:10.1111/j.1532-5415.2006.00701.x)
45. Chen G, Patten C, Kothari DH, Zajac FE. 2005 Gait differences between individuals with post-stroke hemiparesis and non-disabled controls at matched speeds. *Gait Posture* **22**, 51–56. (doi:10.1016/j.gaitpost.2004.06.009)
46. Balaban B, Tok F. 2014 Gait disturbances in patients with stroke. *Phys. Med. Rehabil.* **6**, 635–642. (doi:10.1016/j.pmj.2013.12.017)
47. Kerr A, Rowe P, Clark A, Chandler E, Smith J, Ugbole C, Pomeroy V. 2020 Biomechanical correlates for recovering walking speed following a stroke. The potential of tibia to vertical angle as a therapy target. *Gait Posture* **76**, 162–167. (doi:10.1016/j.gaitpost.2019.12.009)
48. Gloumakov Y, Spiers AJ, Dollar AM. 2020 Dimensionality reduction and motion clustering during activities of daily living: decoupling hand location and orientation. *IEEE Trans. Neural Syst. Rehabil. Eng.* **28**, 2955–2965. (doi:10.1109/TNSRE.2020.3040716)
49. Fukuchi CA, Fukuchi RK, Duarte M. 2019 Effects of walking speed on gait biomechanics in healthy participants: a systematic review and meta-analysis. *Syst. Rev.* **8**, 153. (doi:10.1186/s13643-019-1063-z)