



Faculty of Science
Department of Biology

Methodological Advances in Studying Postnatal Locomotor Development

Thesis submitted for the degree of
doctor of science
at the University of Anwerp, to be defended by
Falk Mielke

supervisors:

Chris Van Ginneken

Peter Aerts

Antwerp, 2024

Methodological Advances in Studying Postnatal Locomotor Development

Methodologische vooruitgang bij het bestuderen van postnatale bewegingsontwikkeling

Institute: University of Antwerp Universiteit Antwerpen
Faculty: Faculty of Science Faculteit Wetenschappen
Degree: Doctor of Science Doctor in de Wetenschappen

DOI: <https://doi.org/10.31237/osf.io/dejkg>

Author: Falk Mielke

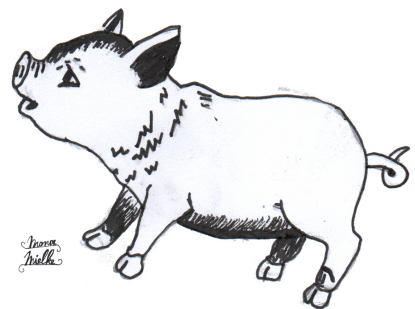
Supervisors: Chris Van Ginneken, University of Antwerp
Peter Aerts, University of Antwerp

Jury: Joris Dirckx (*chair*), University of Antwerp
Ann Hallemans, University of Antwerp
Veerle Segers, Ghent University
Stefan Van Dongen, University of Antwerp
Sam Van Wassenbergh, University of Antwerp

Cover Art, Back: Maja Mielke

Pre-Defence: October 13th, 2023

Defence: March 1st, 2024



Contents

1	General Introduction	6
1.1	Postnatal Locomotor Development	8
1.2	Limited Transferability	16
1.3	Methodological Prior Knowledge	18
1.4	Thesis Outline	22
I	Kinematic Analysis	24
2	Fourier Methods in Locomotor Biomechanics	26
2.1	Abstract	27
2.2	Introduction	28
2.3	Fourier Series in Locomotor Research	33
2.4	Properties of the Fourier Series	35
2.5	Case Study: Predicting Bipedal Walking in Adult Male Baboons	44
2.6	Summary	52
2.7	Supplements	54
3	Fourier Coefficient Affine Superimposition	56
3.1	Abstract	57
3.2	Introduction	58
3.3	Materials and Methods	64
3.4	Results	69
3.5	Discussion	74
3.6	Conclusion	80
3.7	Supplements	81
II	Probabilistic Modeling	90
4	Probabilistic Modeling Workflow	92
4.1	Preface	93
4.2	Introduction: Probabilistic Modeling	94
4.3	Application: Baboon Kinematics	107
4.4	Summary	146
5	Predictive Modeling of Piglet Locomotor Kinematics	148
5.1	Abstract	149
5.2	Introduction	150

5.3	Materials And Methods	154
5.4	Results	162
5.5	Discussion	168
5.6	Supplementary Analysis	172
5.7	Acknowledgements	177
III	Inverse Dynamics	180
6	Inverse Dynamics Workflow	182
6.1	Abstract	183
6.2	Introduction	185
6.3	Dynamics: Theory	187
6.4	Dynamics: Application	198
6.5	Summary	203
7	Inertial Properties	204
7.1	Abstract	205
7.2	Introduction	206
7.3	Materials and Methods	215
7.4	Results	226
7.5	Discussion	233
7.6	Summary	238
7.7	Acknowledgements	239
IV	Epilogue	240
8	General Discussion	242
8.1	First Principles	243
8.2	Predictive Power	244
8.3	Transferability	246
8.4	Methodological Advances	247
8.5	How and Why Piglets Move	249
8.6	Piglets, Baboons, and Humans	252
8.7	Conclusion	255
V	Appendix	258
9.1	Bibliography	259
9.2	Video Digitization	289
9.3	Statistical Parametric Mapping	290
9.4	List of Publications	291
9.5	List of Abbreviations	295
9.6	Acknowledgements	297

Chapter 1

General Introduction

The birth of a precocial creature
Is a wonder of nature, a feature
Of life that's so strong
It can't help but belong
To the world, a miracle, a teacher.

1.1 Postnatal Locomotor Development

1.1.1 Artistic, Computational Motivation

Modern-day computational and scientific tools have led to a deep understanding of life and its processes. We can witness the birth of another living being, and measure and compare parameters to quantify the miracle. We can feel emotionally stirred by that encounter, empathize with the mother or the offspring, feel delight and astonishment, and calculate the physiological improbability of the event. We can be amazed by how little (or how much) time it takes for a little one to stand and walk, to *live* independently, by measuring it with a digital stop watch. We can feel sad for individuals initiated with disadvantageous traits, and we can screen which environmental effects have influenced development, and how.

Thanks to sophisticated analytical tools, we can understand to an ever-growing extent what exactly is going on during the propagation of life. Does analyzing and understanding the wonder make it less beautiful? Does knowing the workings of a “wonder” make it less wonderful?

The limerick above might be considered poetic, you might find it beautiful. But on close read you will find that it is definitely not “art”, and ultimately it is AI-generated nonsense (Fig. 1.2). Does that change your opinion about it?

The duality of research method and research subject, i.e. methodological and biological aspects of research, is also apparent in this thesis. One core aspect of an animals existence is locomotion. Initially, the primary goal of this project was to analyze the locomotion and locomotor development of newborns. More specifically, this thesis project set out to investigate in detail the putative deficits of piglets with lower-than-normal birth weight in locomotor development. Luckily, I could build on the excellent prior studies in our group, in particular the PhD project of Charlotte Vanden Hole (Vanden Hole, 2019). My own project’s purpose then was to extend our understanding of intralimb coordination, i.e. the fine details of how piglets orchestrate the different joints and muscles of their body to get moving, or not. While the heart of the present thesis are newborn piglets learning to walk, its backbone is computational methodology, applied in a novel context. This is a given necessity, because the differences and putative deficits turn out to be most subtle. Because of this, we require highly capable and precise statistical tools to quantify them, just as you would need precision tools to carve the most detailed wooden figures. Yet just as with the novel AI capabilities which currently invade our everyday life, it is important to stay sceptical of the purpose and usefulness of the exapted algorithms. Computational techniques can be abstract, and sometimes their purpose only becomes clear on second application. For this reason, a considerable fraction of the chapters herein branch off to related, but different biological questions (e.g. locomotion of ungulates in general, or locomotor development of bipedal baboons).

It might help if, on the way, we make sure not to lose sight of the actual phenomenon under study. I, the author, personally think that the birth and maturation of a piglet (or any creature) are beautiful and fascinating processes, and I appreciate the opportunity to study them in detail. Nevertheless, herein, I will not ponder further upon the emotional or artistic aspects of the beauty of nature and life. The core purpose of this thesis is to assemble, extend, and scientifically evaluate some relevant facets of current research methods around locomotor development.



Figure 1.1: *"A newborn steampunk piglet galloping on a farm with the sow in the background. (futuristic-biomechanical)", created by the "stable diffusion" AI image generator (<https://stablediffusionweb.com>) on November 3rd, 2023. Impressive as is, this generated image has no deeper meaning or connection to real life, and its only fascinating aspect might be how it was generated.*

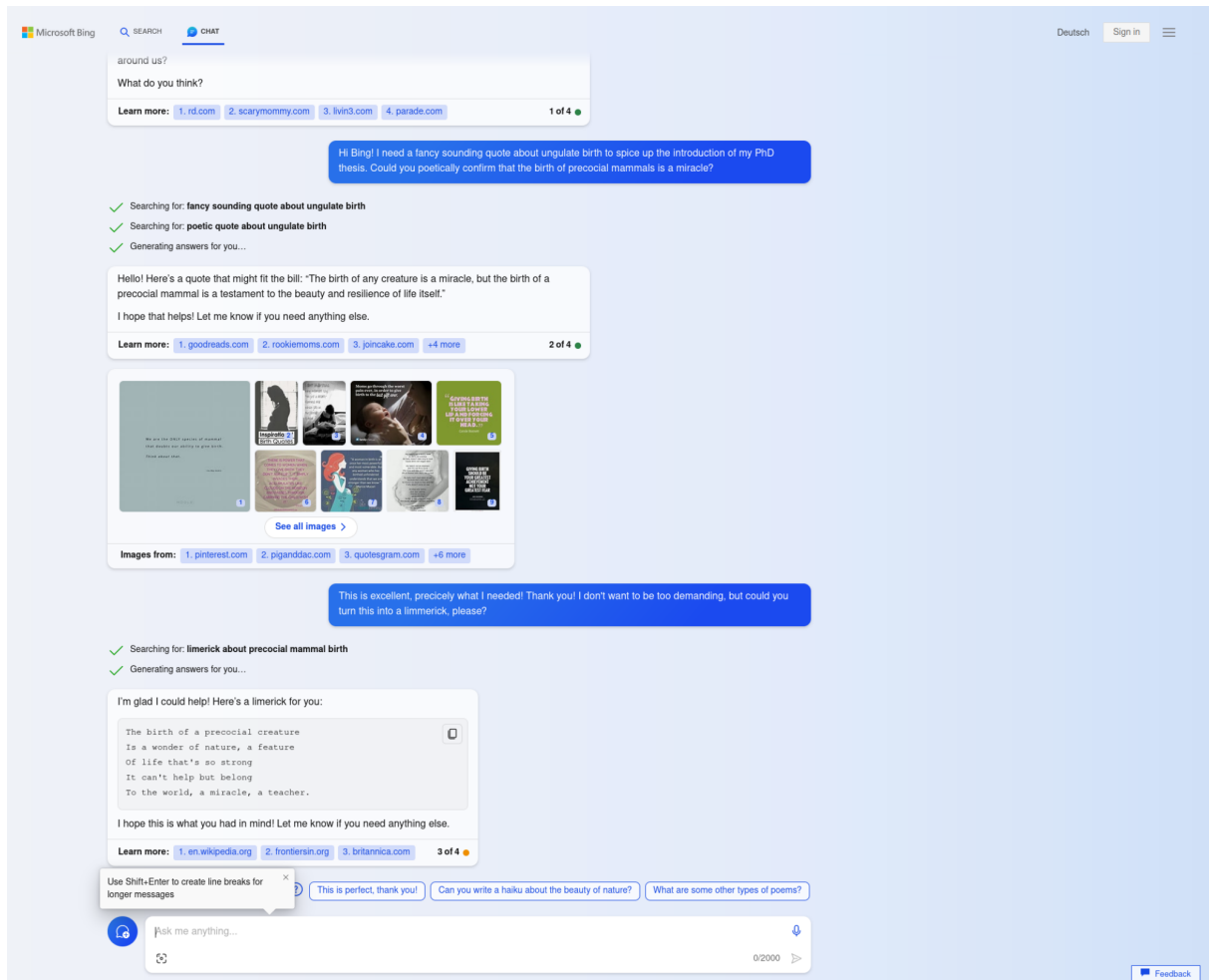


Figure 1.2: *The "bing chat" prompt which lead to the introductory limerick (Nov 3rd, 2023). Besides the use of "DeepLabCut" for landmark tracking, this poem and Fig. 1.1 are the only AI contributions to my thesis.*

1.1.2 Piglets

Any creature’s birth can be fascinating, and studying it can be insightful. Why do we study piglets?

The first is the trivial reason: **because we can**. Our institute at the University of Antwerp has phenomenal equipment. We have incubator-like rearing containers in the “Ark” (a teaching stable in our Veterinary Science department), which can be a cozy warm home to newborn piglets. We have the *3D²YMOX* (3-Dimensional DYnamic MOrphology using X-rays) facility to measure moving animals from hoof to head in light and x-ray vision. The FleXCT scanning facility is just around the block, and the friendly technicians there will scan anything you bring them (if you have the funds, of course). I was lucky to be amidst a great, interdisciplinary collaboration of teams: veterinarians of CoPeD (Comparative Perinatal Development), physicists from the imec-Visionlab and Biomedical Physics (Bimef), and not least biologists at FunMorph (Functional Morphology). The settings could not have been better, and I cannot overstate my praise of all the fantastic people who helped with the projects herein, and all those who previously managed to acquire funding for what enabled my work.

Thank you!

Such equipment and personnel could have been used to study other species of animals. However, there are size limitations. The maximum focal volume of the *3D²YMOX* is about the size of a one week old piglet. The resolution might be a constraint on too small animals, such as rodents. Piglets fit; and as domestic farm animals they are relatively abundant in Belgium.

So we can. And we did.

There is another trivial reason: **curiosity**. We are scientists, and like to study. Piglets are interesting, as I hope this thesis will leave no doubt: they are precocial, but less than most ungulates; they are social, but neonate competition is fierce; they are curious and active; and they share some common features with our own species. Furthermore, the general workings of locomotion are exciting to us. Just observing the outcome of the fine interplay of muscles, tendons, joints, bones - to us scientists, observing this commonplace behavior in high resolution and slow motion is the equivalent of an audiophile listening to a Dvořák symphony. Or to a physicist seeing a rainbow (Lewin and Goldstein, 2012). Have you actually *seen* one?

Yet, back to our cute, pink subject animals. We could have studied any animal; if fact, some of us might have certainly preferred flamingos or boxfish. This is why, at this point and below, I will have to apply a common human strategy: *ex post* rationalization. Walking piglets are graceful in their simplicity. And “graceful”, just like in (human) sports or in playing musical instruments, often describes something which is complex and complicated but performed to look simple. Well, sports and music requires training, and the most adept



Figure 1.3: *"Leihmutter Jolante"*, by Ernst Kahl (date unknown). *Proof that real art can supersede current AI outcomes.*

athletes or musicians seem to us as if they have never done anything else in their lives. But in contrast to hominid trained specialists, the piglets we study indeed have never done anything else in their lives! They are born, and they walk more or less gracefully within eight hours, as my colleague Charlotte concluded from her exemplary PhD work (Vanden Hole, 2019). How much trouble it would save us, and how much different trouble it would cause, if our own toddlers could pull that trick!

Which brings us to the third reason: **piglets are, to a certain degree, analogous to humans**. In other words, we have the working hypothesis that piglets are a good model species for some aspects of human ontogeny (Fig. 1.3). Emphasis on *"to a certain degree"* - are they, really? I will get back to this at the end of the thesis (Ch. 8). Domestic pigs are maybe the only species in the world that is more degenerate than our own (on second thought, maybe not). They are sedentary, comparatively hairless, highly social, omnivorous (which means they can live of the crap humans feed them), training them to work takes considerable efforts. They have probably gone through a couple of founder situations in their evolution, yet are far more prolific than is good for themselves and the planet. That is already quite a lot to have in common. And indeed, many researchers rightfully argue that we can draw conclusions about humans from studying piglets - for some aspects of our humble, shared existence. For example, highly relevant research about infant gut function was achieved by the group of Prof. Per Sangild (Che et al., 2010; Cilieborg et al., 2011; Muk et al., 2023; Sangild et al., 2006). One aspect of newborn piglets is that some individuals are putatively born with a deficit. Those have received different classification labels: small for gestational age, growth retarded, low vitality; for the majority of this thesis, I will stick to the phenomenological classification of "low birth weight", or LBW, yet the fine nuances in the different classifications deserve further discussion (Van Ginneken et al., 2022; Wootton et al., 1983). Kids can be troubling. Prof. Sangild argued in his Francqui honorary lecture here at the University of Antwerp (April 26, 2022) that having kids during a PhD is detrimental for

progress of the project¹; which teases me to contribute contradicting evidence (this thesis). But the question remains: **can piglets serve as a model species for the study of human locomotion?** An obvious “show stopper” might be that, as mentioned above, piglets are precocial, whereas humans are about as altricial as can be. On the other hand, suids are the more altricial of the ungulates: feral *Sus scrofa* females build nests and keep their young in shelter for a couple of weeks (Wischner et al., 2009). Domestication might have involved restraining sows, at least to a certain extent (nests *vs.* group pens *vs.* single pen), and the way we keep them might have favored more mobile young who have reduced mortality from crushing (Grandinson et al., 2002). The picture is complex, and certainly the analogy of piglets and humans should not be taken for granted (Vanden Hole et al., 2017).

¹He might have only referred to a parental leave causing delay, which is good practice, and a valid point.

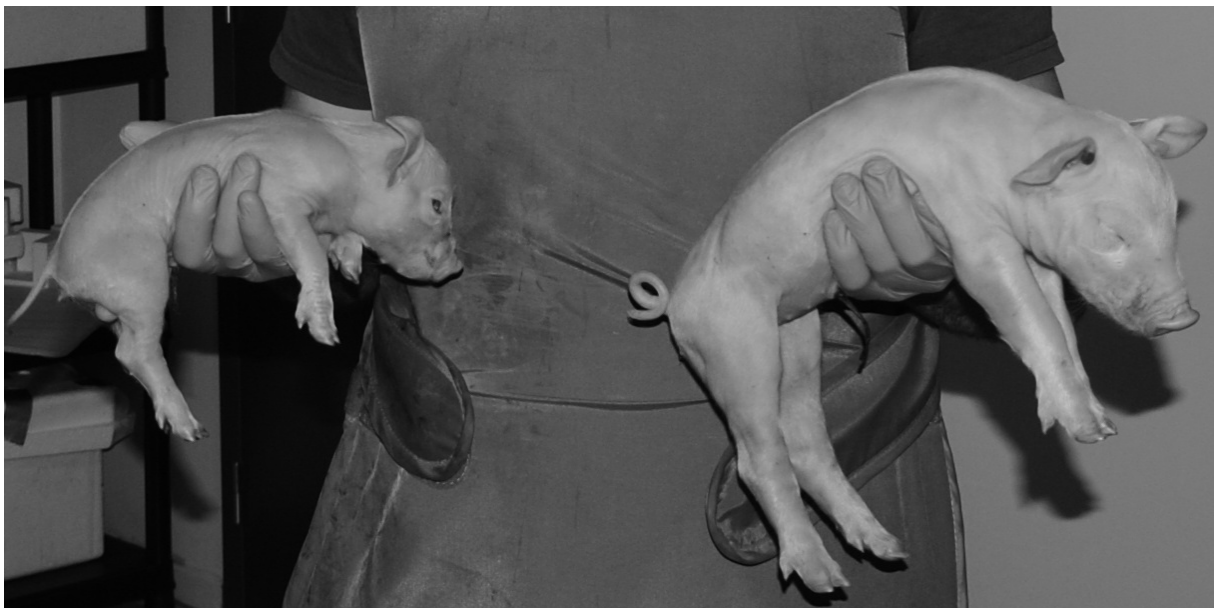


Figure 1.4: *Piglets of low and normal birth weight (left/right, respectively) for size comparison. The animals are female siblings.*

A second “show stopper” for this thesis would be if it turns out that the effect we would like to quantify (a difference between piglets of low- and normal birth weight, see Fig. 1.4) is negligible in magnitude. In other words: if the locomotion of newborn piglets of low and normal birth weight actually does not differ in a relevant or at least measurable way. First off, it is obvious that low and normal birth weights do differ: per definition, in terms of weight. But different weight might go in hand with different physical constraints, and one has to account for the trivial consequences of a different physical appearance (Aerts et al., 2023b). The fact that there are hyperprolific sows, increased litter sizes, and an increased rate of intra-uterine growth restriction and low birth weight (Alvarenga et al., 2013; Quiniou et al., 2002) is in itself an argument for using domestic pigs as a model for human early development: it gives us useful categories for comparison. But how to approach the potential lack of an effect? To set the frontier right from the start, I will insist to apply good scientific methods (Chamberlin, 1890; Platt, 1964; Popper, 2002). If we *searched for* differences in piglet locomotion, just search hard enough, we would surely find them. The null hypothesis in statistical testing is that there are no differences between the study groups, which is why researchers tend to actively search for and find differences. However, with so many variables influencing locomotion (multiple testing), it is inevitable that some variables will differ. It cannot be presumed that these differences generalize to the population from which we sampled. Instead, our attempts must be to **falsify the hypothesis that there are birth-weight dependent differences in piglet locomotion**. If we fail, we learned something.

Common methods of studying locomotion fall short in their falsification capacity. First, some methods miss or loose information in processing. For example, imagine we measured two maximally dissimilar animals, a stork and an elephant, to quantify how they move their limbs when walking. The stork will use high clearance, high flexion, being adapted to cluttered terrain and swampy underground, whereas the savannah-dwelling elephant is large, “graviportal”, and usually walks with minimal joint flexion along the limb. We could observe these animals while they are co-incidentally running at dynamically similar speed; assume that we measure exactly the same (normalized) speed. Although we might not see differences in speed, we cannot exclude that the subjects (likely) differ in the way they move their limbs (coordination). Second, methods might have trouble handling variation. One subject animal was observed to be slower in one trial, but does that mean it *cannot ever* go faster as well? So, on the one hand, failing to incorporate (or even summarize) repeated measurements might be neglecting variation. What does it mean if there is a significant difference which is of the order of magnitude approximately equal to the standard error of the mean of the measurement? On the other hand, methods can amplify variation, by not including all parameters: imagine the “unexplained” variation within a set of recordings of piglet aged 0-10 days, but not documenting the age (which brings us to the next point). Third, the matter is complex: just for kinematics, there are subject characteristics (age, sex,

...), spatiotemporal gait variables (speed, duty factor, clearance, ...), dynamic posture and effective range of motion, and finally coordination. Stay tuned for more on this. Fourth, if we measure, we must keep track of measurement uncertainty. If we do not know how accurate a figure was, it is meaningless. We better have a potent modeling framework to incorporate all that data: if we do not incorporate all measures or fail to explain their variability, differences might be due to sampling biases and “hidden” variables. The falsification approach is a unifying aspect of all the chapters herein.

To summarize, there are some genuine reasons for studying suid ontogeny, and there are some methodological challenges which I will address. Yet whether or not domestic piglets are a good model species, specifically for human locomotor development, is up for debate, and I will return to it in the general discussion (Ch. 8).

1.1.3 Limitations of Piglet Locomotor Research

Part of the natural fate of this (and any) thesis is that others have done research on the topic before. Existing research specifically on piglet neonate locomotion is diverse. It covers activity and welfare (Beattie et al., 2000; Kashiha et al., 2014), muscle workings and general motor function (Alvarenga et al., 2013; Andersen et al., 2016), behavioral and neurological (dys-)function (Roelofs et al., 2019; Sullivan et al., 2013), and gait (Meijer et al., 2014; Vanden Hole, 2019; Vanden Hole et al., 2017; von Wachenfelt et al., 2008).

We already know a lot about many aspects of piglet locomotion, yet there are two major limitations. Firstly, **transferability is limited** (*cf.* Gatesy and Pollard, 2011): it is not immediately possible to mirror findings in morphologically disparate species (such as humans)². Pigs are precocial quadrupeds, and it might seem far-fetched to apply analogous hypotheses to bipedal humans, let alone to our helpless, immobile infants. Secondly, most methods only superficially quantify locomotion: they work on general activity patterns or spatiotemporal gait variables, or remain qualitative in the comparison of kinematic profiles. Few studies dig deep into **intra-limb coordination**, i.e. the temporal succession of limb segment movements.

The compensation strategies I attempt in this thesis are methodological development, and the move to a model species which can be considered “intermediate” (baboons, facultative bipeds).

²“Translational medicine” (i.e. the application of more general scientific findings to human clinical medicine, Denayer et al., 2014) might be considered a specific type of transfer which in the case of locomotor research is impeded by morphological disparity.

1.2 Limited Transferability

1.2.1 Baboons

So there is a limitation to piglets. Or, rather, I consider it a bonus. They have two extra limbs on the ground, thus four limbs in total. This gives us about twice the number of joints to analyze when studying grounded locomotion. Not trivial to handle, but feasible, as I will show.

It came as a great coincidence that, during SEB conference 2021, I watched my colleague François Druelle presenting a data set acquired from bipedally walking baboons (*Papio anubis*). The rationale to study these animals can be directly transferred to this thesis: they are facultatively bipedal, i.e. they occasionally walk on their hind limbs, just like humans often do (Aerts et al., 2023a). Prof. Druelle and colleagues thus analyzed a developmental series of baboons: infant, adolescent, and adult individuals (Druelle et al., 2021). Their intention is to find out if and how the use of a bipedal gait changes with age.

That sounded like a familiar task, and with the euphoria of just having implemented a different, potentially useful method (FCAS, see Ch. 3), I reached out to François and he kindly provided the data. Not only the videos, but also the tracking (i.e. kinematics “to go”). I much appreciate the time and effort François and his colleagues and students put into generating the data, and I am thankful for the exchange. And I regret that this has not been put into another publication, mostly for the lack of time.

Nevertheless, work on the baboon data set has a prominent place in this thesis. The fact that there are fewer joints (bipedal), the fact that the data set has been well studied before, and the fact that this is also a developmental series (though different: age groups are discrete), predestined this data set to be my “playground”. I used it to refine my procedures and code. And I ended up writing two “tutorial”-chapters, one each for Fourier Methods (part I) and Probabilistic Modeling (part II).

Yet in fact, you could argue the other way round. Piglets are a putative model species, but require some transfer work to draw reliable conclusions about humans and human development. *Papio*, on the other hand, are bipedal primates, and their bipedalism develops gradually over time – just like in our own species. On the downside, baboons are less readily available, also for dynamic measurements (e.g. XROMM, part III). Hence it is good to keep both model species in the loop.

Having a biped on board makes transferability more likely, but is no guarantee: it might be hard to transfer findings to human infants, since there is still morphological and behavioral discrepancy. To work towards solving this issue, it might be useful to take one step back, explore what locomotion is *technically*, and see which analytical methods are available.

1.2.2 What is Locomotion?

Fix a single solid metal rod to a support with a frictionless cylindrical joint, deflect it by a small angle, and let gravity do its predictable magic. This is a simple oscillator, a “physical pendulum”, which physics professors enjoy assigning to their students as calculation exercise (LibreTexts, 2022a). Now add a second such element, another cylindrical joint and metal rod. Something astonishing happens: we get chaos (Harrison, 2021). Even more fun to calculate. Due to minuscule variations in the initial conditions of this system, it cannot be solved deterministically any more.

Animal limbs resemble the physical pendulum: they have more restrictive joints, yet usually even more segments. We should expect chaos on this system. But chaos is not what we observe, quite the contrary. Locomotor movements are highly targeted, i.e. pseudo-deterministic, often repetitive. Variations (though ubiquitous) are minimized by training. This observation is due to **coordination**: coordination is herein understood as the temporal pattern of non-random movements of different limbs (inter-limb coordination) or different joints within a limb (intra-limb coordination). Coordination is achieved by (neuro-)motor control. Control is achieved by the central or peripheral nervous structure which establishes coordination. The study of neuromechanics attempts to relate the central and peripheral control mechanisms to the observed motor outcome produced by the musculoskeletal system, i.e. coordination (Full and Koditschek, 1999; Nishikawa et al., 2007). Of course, the control “facilities” of each individual are bespoke for the precise anatomy of that individual: each brain is tuned to the layout of segments and motors of each individual. We do not know the exact “wiring” of each animal, though remarkably we observe “functioning” individuals time and time again, which are the less interesting cases in which the nervous system does a great job in steering its muscles.

Because of this intimate link of neural control and coordination output, the motor system can be considered a looking glass into the workings of the neural system of animals. This can be understood in an evolutionary sense: a change in the behavior, e.g. a transition to bipedality, will reflect in the control system (Zehr et al., 2016). It can also be understood in a medical sense: if an individual does not “walk” as expected, locomotion enables the diagnosis of disease (Mielke et al., 2023). Perturbations, both in the environment (Daley, 2018) and in the two systems (De Groote and Falisse, 2021; Sponberg et al., 2023; Ting et al., 2015), and the organism’s response to it, can be analyzed to understand the robustness, plasticity, and rehabilitation of motor control.

The vertebrate locomotor system is highly complex, therefore sensitive to changes, and therefore a relevant multi-purpose tool for diagnostic and comparative research. Yet to fully understand it, we need in-depth analytical tools which capture the whole ensemble of characteristic parameters.

1.3 Methodological Prior Knowledge

1.3.1 Which Analytical Tools Do We Have?

One would think that our toolbox to study terrestrial locomotion is well-developed and sophisticated (Demuth et al., 2023; McHenry and Hedrick, 2023). After all, locomotion has been intriguing people for hundreds, thousands, maybe tens of thousands of years. Whenever humans develop a new imaging technique, from cave drawings to digital cameras, they use it (first for pornography, and second) for locomotion (Figs. 1.5, 1.6).

In consequence, the *technology* for studying locomotion is available, and sophisticated. Consumer grade cameras have a high temporal and spatial resolution, and can be calibrated to measure 3D space. Compare modern day smartphone cams to, for example, contemporary x-ray techniques for studying the inside of things: both are equally recent, resolution of the latter is usually inferior, price difference is defeating, yet we do not even have color on x-ray (although it is just another form of electromagnetic radiation, see Ch. 7). Visible light videography seems to be far more accessible than other imaging techniques. We can thus expect that far better camera equipment is available than would minimally be required for our research questions. Sounds too easy - which is probably why my supervisors decided to send me out to a pig farm in 2017 with an inferior JVC camcorder (while a synced 4-cam 300fps Norpix system, and a set of high res GoPro's, were available at the time). Although that would not be fair to claim. There is also always an allocation to be set on the spectrum of video quality (especially framerate) and manual processing time (see Ch. 9.2), and I might still be clicking points on videos today if my supervisors had decided differently. And, after all, the JVC video data might be the most relevant data of this PhD (Mielke et al., 2023).

Acquiring videos is just the first step: as always in science, we need numbers and statistics to get reliable results. Where do we stand on the analysis side? You might be underwhelmed. We are in the year ~~2022~~ 2023 2024, and (as if agnostic to the rise of computers) people still essentially rely on footfall patterns (Hildebrand, 1989), spatiotemporal gait parameters (stride frequency, speed, duty factor, etc.), and visually comparing angular profiles (Ekhtator et al., 2023; Jones and Hasiotis, 2023; Plocek and Dunham, 2023; Young et al., 2023b). Those are doubtlessly great methods to get an overview, and they remain the go-to first analysis for any kinematic data set. Yet as I will argue in the chapters to follow, they only capture limited information and are therefore insufficient for a falsification approach, as illustrated with the stork and elephant example above.

People have certainly *tried* to complement the conventional measurements with more in-depth computational methods. “Planar Covariation” theory came out of one attempt to develop a technique which captures the essence of terrestrial locomotion (Borghese et al., 1996; Ivanenko et al., 2008). I think it is flawed, in a sense that it biases the data for high correlation just to find out that it is highly correlated. Throughout the chapters of this

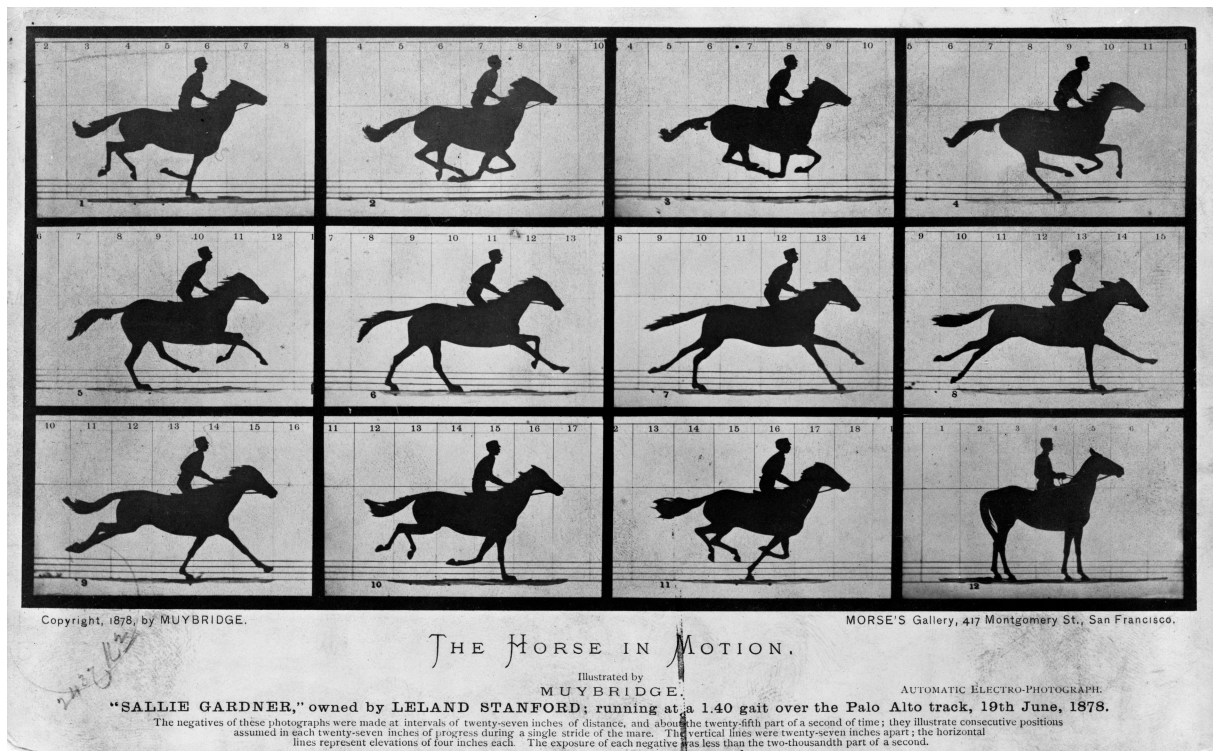


Figure 1.5: *"The Horse in Motion"*, by Eadweard Muybridge (1878). It seems like every PhD thesis on locomotion or cinematography has to start with Muybridge.

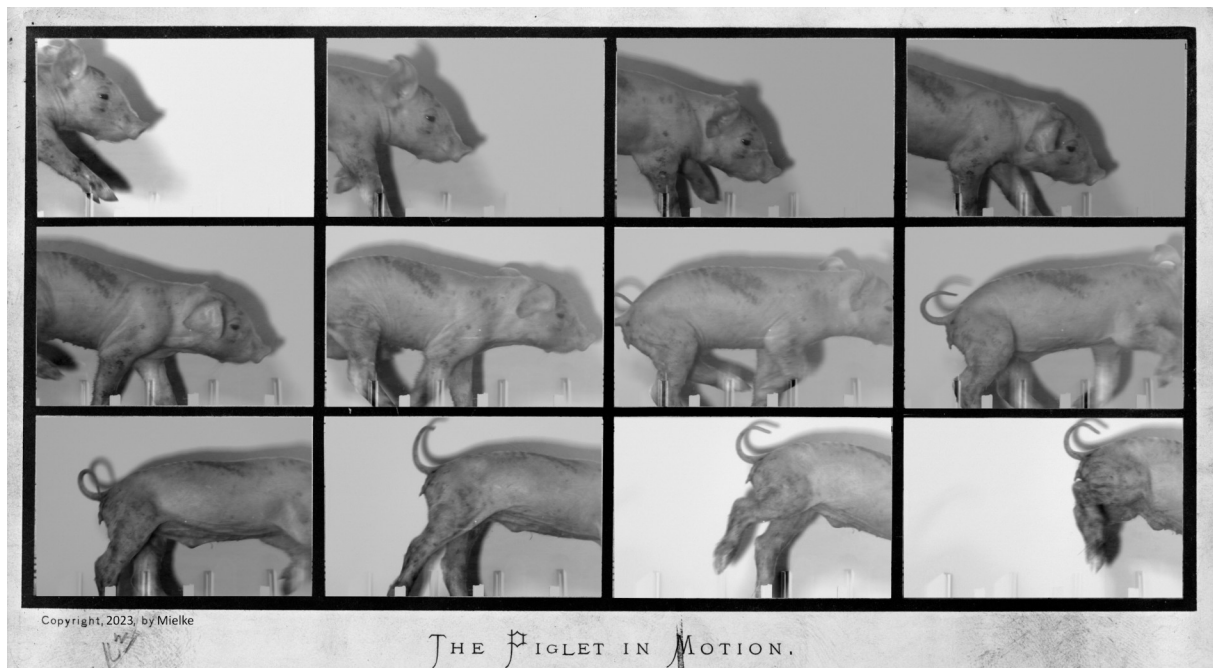


Figure 1.6: *"The Piglet in Motion"*, by Falk Mielke (2020). Sophisticated, modern technology exists for studying animal locomotion, but the idea is the same.

thesis, you will find subtle and polite rants about the planar covariation technique, so I will not go into detail at this point.

Another, potentially useful method is Statistical Parametric Mapping (Flandin and Friston, 2008; Friston et al., 1994; Pataky, 2020; Pataky et al., 2008; Worsley et al., 2004). It is designed to test one-dimensional (e.g. time-varying) data, such as kinematic profiles, for the statistical evaluation of group differences. “Statistical evaluation” refers to classical hypothesis testing, which is normally complicated by the non-independence of continuous temporal measurements. SPM implements various hypothesis tests for continuous data by a clever adaptation of random field theory (Brett et al., 2003; Kemeny et al., 1976). The method first and foremost tackles the temporal interdependence of measurement points (Pataky, 2020). It does not overcome the limitations mentioned above, transferability of kinematics and the quantification of intra-limb coordination. And its current implementations are restricted to classical, frequentist statistics, with all their assumptions and pitfalls. For these reasons, I have not applied SPM in my research, nevertheless appended a brief summary of that method for interested readers (Ch. 9.3; cf. <https://spm1d.org>).

1.3.2 An Overlooked Transformation

And then there is one strain of locomotor science which deviates from the Western mainstream. It was advanced by Nikolai A. Bernstein, and I have dedicated a whole chapter to it (Ch. 2). His group excelled in at least two ways (Bernstein, 1935; Bongaardt and Meijer, 2000). They used clever engineering to get continuous measures of joint positions (Bernstein, 1927b) and applied angular transformations (i.e. Fourier Series) to analyze the data (Bernstein, 1927a). These early works are remarkable, and the day I first read through it and realized their importance was maybe one of the happiest days of my PhD. Essentially, what I did in the work summarized herein, is transferring Bernstein’s ideas to a computer framework, and extending them with modern statistical models. I acknowledge that I am not the first one to take this route (Pike and Alexander, 2002; Webb and Sparrow, 2007). And there is a group of authors who have translated, popularized, and advanced the “Bernstein school” (Laczko and Latash, 2016; Latash, 1998, 2008, 2010). However, given the usefulness of the transformation, I find that ~~there are~~ google scholar returns astonishingly few studies to date who have applied Fourier Analysis to locomotor data.

Venturing along this path, implementing and improving the Fourier analysis method, has partially led this thesis away from its original study subject, neonate piglets. In the end, my work has taken a methodological focus much stronger than anticipated. Yet this is justified in the preliminary observation that differences in low- and normal birth weight are on the subtle side of the effect magnitude spectrum (possibly more visible on intra-limb

coordination), and in consequence only the finest tools would be capable of detecting them. Furthermore, as I will demonstrate in the individual chapters of this thesis, transformation to the frequency domain can help to partly overcome one of the major limitations posed above: the transferability of kinematic data. In the frequency domain, some aspects of locomotion are easily and logically separated (namely: posture, effective range of motion, and intralimb coordination). Apparent differences in either of these separated measures can be related to subject characteristics, such as morphology.

This opens up the path to more refined answers to long-standing comparative and evolutionary research questions.

To summarize, there are a lot of computational tools at hand to study locomotion. Some are old, some more recent. Some capture only a fraction of the information, others try to be inclusive. Some are frequently used, but the ones I find most relevant did not get the recognition in the field that I would think they deserve. This thesis is also an attempt to rectify that.

1.4 Thesis Outline

And this is where my PhD project found its place. We studied piglet locomotion, attempting to falsify a difference that is due to birth weight category. This requires a methodological foundation on topics unfamiliar to most Biologists (Fourier Analysis, probabilistic models), which I introduce in dedicated parts. By adopting and extending these existing techniques for the framework of biomechanical research, I demonstrate a way to overcome previous limitations of transferability and the quantitative analysis of coordination.

The thesis will have the following structure.

Part I: Kinematic Analysis

- I will introduce Fourier Methods with a practical guide, and review its occurrence specifically in the field of locomotor biomechanics (Ch. 2).
- The power of Fourier tools is then unleashed onto a large cross-species study that covers most of the ungulate clade (Ch. 3, Mielke et al., 2019).

Part II: Probabilistic Modeling

- Processing complex kinematic data, I will introduce the basics of probabilistic modeling (Ch. 4) in a “tiny textbook of statistics”, which is based on an example of locomotor development in baboons.
- Moving on to piglets, I will turn the conventional modeling strategy upside down (or leftside right?) and apply a predictive model to piglets (Ch. 5, Mielke et al., 2023).

Part III: Dynamics

- Kinematics are just half of locomotor research: the study of forces and moments is reviewed and summarized in another overview chapter (Ch. 6).
- To calculate joint moments, one needs to measure inertial properties, yet it turns out that these are not trivial to retrieve from x-ray tomographic images (Ch. 7).

The thesis closes with a general discussion to round up and project the presented work to potential future applications (Ch. 8).

Enjoy reading!

Part I

Kinematic Analysis

Chapter 2

Fourier Methods in Locomotor Biomechanics

2.1 Abstract

Analysis workflows which are based on the Fourier theorem are abundant in physics and engineering, yet rarely found in the biological literature. They would be particularly useful in the analysis of locomotor kinematics, where movements can usually be broken down into repetitive cycles. I herein repeat the basics of Fourier analysis and transformation to the frequency domain by reviewing previous studies which applied it in some form in the field of locomotor biomechanics. I then discuss terminological and methodological heterogeneity among these studies, and point out some relevant properties of the method which have received relatively little attention. One of the under-appreciated capabilities is that the transformation enables a diversity of multivariate analysis and modeling methods. I demonstrate this by applying Fourier Series Decomposition, multivariate analysis and probabilistic modeling to a previously analyzed data set of bipedal walking kinematics in baboons. The results are complementary to the prior analysis, and the capability to predict actual kinematic data for unobserved character combinations is of particular interest for future kinematics research.

2.2 Introduction

2.2.1 Recurring Solutions

In 1822, a French mathematician initiated the remarkable discovery that any continuous signal, no matter how complex, can be represented by a sum of relatively simple trigonometric functions (Fourier, 1822). Despite some initial limitations which were mended by scientists to follow, Fourier’s work opened the door to the frequency domain. Fourier analysis has become a widely-used tool in physics and engineering, whereas biologists have only occasionally made the transformation. In particular, although it has been argued that the study of highly repetitive locomotor kinematics would benefit from applying Fourier methods (Mielke et al., 2019; Pike and Alexander, 2002; Schneider and Chao, 1983; Skej  et al., 2021; Webb and Sparrow, 2007), it seems that few researchers have applied this simple method to date. One reason for the hesitant adoption of the method are certainly terminological complexities in Fourier Theory. These become apparent when assembling an overview of past attempts to apply related methods to locomotor kinematics.

In this chapter, I review the history of the application of Fourier Series in terrestrial vertebrate locomotion.

Some properties and practical misunderstandings are highlighted, before applying the method to a previously analyzed data set on bipedal locomotion in baboons (Druelle et al., 2021) to once more demonstrate the enormous potential of Fourier analysis.

This article is not targeted at a mathematically rigorous audience, who will certainly find shortcomings in the simplified explanations I provide. Instead, I aim at facilitating the entry to the topic for applied biologists, especially in biomechanics. Fourier Theory is well studied and widely applied, and numerous textbooks exist for all levels and fields (*cf.* Bracewell, 2000; Osgood, 2019). The rich spectrum of Fourier-based methods comprises versatile techniques which would certainly facilitate the work on many relevant research questions in biomechanics.

2.2.2 Fourier Series

Typical kinematic measurements are joint angles, such as a hip angle, over time (but note that the following will equally apply to any continuous, periodic signal). One recurrent unit of terrestrial locomotor behavior is a “stride”, and it can go as follows (Fig. 2.1 A). If an animal walks, the distal part of its limb would touch the ground at a position ahead of the center of the limb. Our example joint of interest, the hip, would then usually be highly flexed. During stance phase, i.e. while the limb is in contact with the ground, the hip gradually extends, maybe reaching the lowest angle around the initiation of the swing phase. While the limb swings forward to reach another touchdown, the hip joint flexes again. Because the limb is a system of rigid bodies with a given mass (*cf.* Ch. 6), accelerations would never

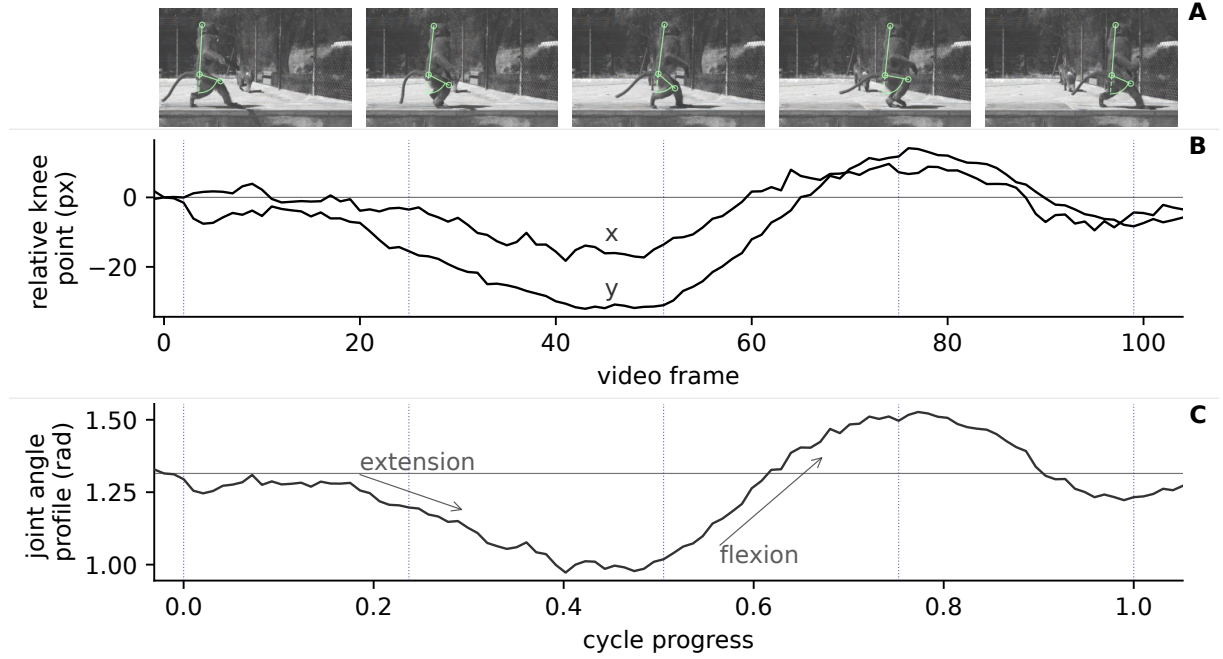


Figure 2.1: Kinematic data generation workflow. (A) Videos of animals are taken, here a baboon walking bipedally from left to right. The frames captured at touch-down, mid-stance, lift-off, mid-swing, and consecutive touch-down are shown. (B) Points of interests, e.g. joints, are tracked, which yields their pixel position over time. In this case the knee x and y positions are shown relative to their first value. Vertical bars indicate timing of the frames from the upper panel. (C) Joint angles are calculated from groups of three of those points of interest. Their change over time is a joint angle profile. I herein define joint angles to be zero at a fully extended (straight) joint, and positive for joint flexion. Per definition of “steady state locomotion”, joint angles in steady state movement will return to approximately their start value over a cycle (indicated by the horizontal line). Mathematically, this is called “periodicity”, the joint angle profiles in steady state locomotion are thus “cyclic”.

be abrupt, and the hip angle plotted over time might be described as a “smooth saw-tooth curve”, or even resemble a sine wave (Fig. 2.1 C). The change of a joint angle over time is called a **joint angle profile**.

One could quantitatively capture joint angle profiles in different ways (Fig. 2.2). The most intuitive, and by far most prevalent way is the time domain representation (Fig. 2.2, left column). It stems directly from the technical method by which kinematic data is usually extracted: landmarks on videos of the walking animals are tracked frame-by-frame (*cf.* Appendix 9.2 and Mielke et al., 2020). The outcome are two columns of numbers: a “time”, capturing the seconds at which a frame was captured, and then the “joint angle” measurement itself, usually assessed from three landmark positions (Fig. 2.2, lower left). The more temporal sampling points, the better the temporal resolution, but the higher the landmark tracking workload and storage requirement.

A key assumption above is that the stride was recurrent, or cyclical, i.e. periodic. This means that an almost (but not perfectly) identical stride will precede and follow the one of interest. This is also called the “steady state assumption”, i.e. the requirement that the observed behavior is regular (and thereby comparable to other observations). If (and only if) that assumption is met, the joint angle profile will start and end at approximately the same value, i.e. the hip angle at each limb touchdown will be almost the same. Plotting this as angle against time will show a line with ends fixed at the same heights of the left and right y-axes. This is metaphorically the same situation as in a guitar string, which is fixed at both ends and performs some excursions from the zero line at positions in between. To further stress this analogy, the hip joint angle profile can be described as the combination of a “base tone” and a number of “harmonics” (or “overtones”). Knowing the base and the relative contribution of the harmonics is sufficient to reproduce the “sound”, or in other words to fully know the exact position of each element of the string or signal, or each instantaneous joint angle of a stride cycle. Base and harmonics are called the “**frequency domain**”, i.e. the representation of a cyclical signal by the relative contribution of its frequency components (Fig. 2.2, right column). Any periodic, continuous signal, and thus also any cyclical joint angle profile, can be described like this, just as any tone the guitar string makes can be described as a sum of harmonics. It might theoretically take infinitely many harmonics to exactly represent the signal (e.g. in case of non-smooth signals). However, in practice, the effects which would cause this are usually considered to be measurement noise. Physiological, “real” signals normally require few harmonics to be appropriately captured.

How can Fourier Series be applied? The method is based on the following formula (Bracewell, 2000; Fourier, 1822; Gray and Goodman, 1995; Mielke et al., 2019), which is presented here for completeness and should not worry the reader.

$$c_n = \frac{1}{T} \sum_{t=0}^T e^{-2\pi i n \frac{t}{T}} \cdot f(t) \quad \forall n > 0 \quad (2.1)$$

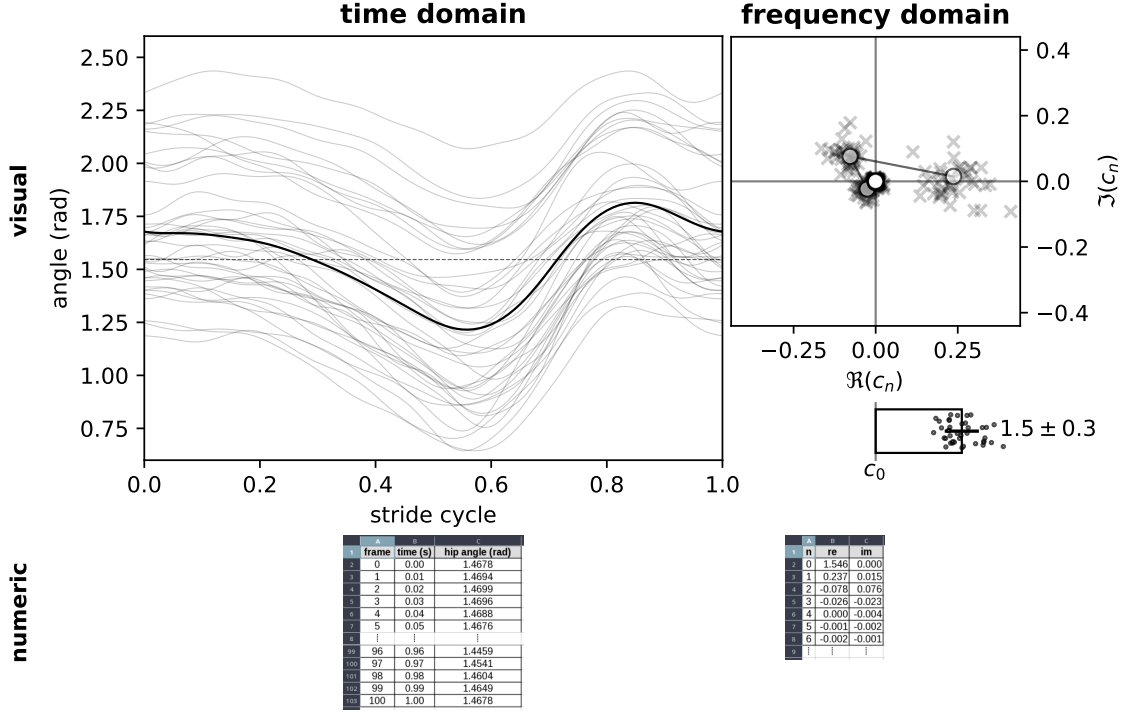


Figure 2.2: The frequency domain. (Upper Left) Hip joint angle profiles of bipedal locomotion in baboons, as conventionally displayed (gray traces: repeated observation; black trace: average profile; dotted line: mean angle). The joint angle profile shows the measured angles over the course of a stride cycle. (Lower Left) The very same information can be stored in an excel list, here shown as a screenshot, consisting of an indefinite number of value pairs of time and angle. (Lower Right) By using Fourier methods, traces can be converted to a list of relatively few coefficients, without loss of information. (Upper Right) These coefficients have a real and imaginary part, which could be plotted in the complex plane. (All) All these displays capture identical data. A visual representation (upper panels) contains exactly the same information as a numeric representation (tables / lower panels). Similarly, the information contained in frequency domain representation (right panels) is identical to that in the time domain (left panels). With the methods described herein, one can transform from the time- to the frequency domain representation or back without loss of information. The frequency domain, which is accessible through Fourier Analysis formulas, is just a different representation of the data, and the transformation procedure is analogous to “plotting” (i.e. deterministic, reversible, favorable in some circumstances).

Where c_n is a complex Fourier coefficient, T is the period of the signal or in our case the stride cycle duration, t is the time (i.e. “x” of the signal), $f(t)$ is the signal (i.e. “y” of the signal), e is the Euler number, i is the square root of minus one. The capital sigma tells mathematicians that this is a series, and thus the equation defines the **Fourier Series**.

At first, the analytical formula might seem quite complex for everyday application, and computer code might be more usable. In supplement 2.7.1 and online (https://git.sr.ht/~falk/fcas_code), I provide code for Python, R and Matlab which can just be copied and used¹:

```
coefficients = FourierSeriesDecomposition(time, signal, order)
```

It can be inverted (just as the analytical formula):

```
signal = FourierSeriesRecomposition(coefficients, time)
```

Knowing the assumptions and limitations, and paying attention to the data types and formats, these functions can be simply applied to transform back and forth between the time- and frequency domain.

There are numerous morphotypes of the Fourier theorem. Equation (2.1) is the exponential form of the Fourier Series. The formula exists in other variants (which are equivalent by Euler’s formula): the rare but useful “amplitude-phase” form, and the more abundant “sine-cosine” representation. Historically, trigonometric functions are more convenient to handle, whereas modern computer implementations easily capture complex numbers, whereas humans have most intuition for amplitude and phase. The Fourier Series is related to, but slightly different from, Fourier Transform, a relation which will be discussed in detail (Ch. 2.4.2). Applying the Fourier Series is a decomposition of a signal into its harmonic components, hence I will use the term **Fourier Series Decomposition (FSD)** herein. “Discrete Fourier Transform” (DFT) can be considered synonymous to “Fourier Transform”, emphasizing the “discrete” nature of sampled data; “Fast Fourier Transform” (FFT) is an efficient algorithm to implement Fourier Transform. DFT/FFT work on any signal and typically involve windowing, whereas FSD demands a periodic input signal. FSD returns a single row of numbers (“**spectrum**”, i.e. harmonic contributions) for that signal, and that spectrum is discrete. FFT natively also returns a spectrum (Welch, 1967), but a continuous one. However, in practice, Fourier Transform is often applied in shifted windows and returns a “spectrogram” (i.e. frequency content over time). Finally, a Fourier Series can be applied to a single joint angle (“1D”), or to multiple signals at the same time (e.g. on phase plots). All of these (and other) variants of the method have been applied in the research of locomotor kinematics.

It is therefore worthwhile to review the historic use of related methods in studies on locomotor kinematics, and look for some common patterns and strategies.

¹The “order” parameter in the decomposition is the number of coefficients to retain.

2.3 Fourier Series in Locomotor Research

The analogy of limbs and a pendulum was established early on (Fischer and Braune, 1904; Weber and Weber, 1836), and maybe that is why Fourier Series was immediately among the relevant considerations for analysis. To my knowledge, the first formal application of it was by Nikolas A. Bernstein (Bernstein, 1927a, 1935), who in the early years of his work was concerned with methodological advances in measuring motor actions (Bongaardt and Meijer, 2000). Bernstein suggested and applied a trigonometrical sum, which consisted only of sines and bares some resemblance in form to a Taylor approximation.

In his later work, Bernstein advocated a hierarchical organization of motor control (*cf.* Bongaardt and Meijer, 2000). This topic was taken up by studies in the field of motor learning, which would find evidence of higher order harmonics being affected by locomotor learning (Gallistel, 1982; Marteniuk and Romanow, 1983). Their methodology is overall similar to that of Bernstein, yet improved by heuristics and standardization (Jackson, 1979; Porges et al., 1980). The use of Fourier methods for asserting gait symmetry and variability were recognized (Soudan, 1982).

Van Weeren *et al.* applied Fourier Series to measure skin marker displacement, not joint angles (van Weeren et al., 1992), yet extracted the coefficients with a regression instead of direct transformation. Schendel and colleagues quantified inter-vertebral angles with a sine-cosine Fourier Series (Schendel et al., 1995), again with a computational regression. It was found that relatively few harmonics can be sufficient to detect symmetry and lameness in horses (Audigié et al., 1999; Peham et al., 1996). Others have applied a Fourier Series to the segment angles of limb segments (Grasso et al., 2000), for example to infer amplitude and phase relations. Fourier Series has recently been applied to the specific analysis of human running (Skejø et al., 2021).

Two studies stand out from the previously mentioned, conventional uses of Fourier Series. Pike and Alexander (2002) have made use of the data transformation to perform multivariate, quantitative comparison of kinematics across a broad range of taxa. Their focus lies not on the method itself and the generated coefficients, but rather on downstream data processing methods. Another hallmark study in terms of advocating Fourier Series in locomotor research is that of Webb and Sparrow (2007). The authors provide a comprehensive introduction to the method, list its advantages and fields of application as well as future perspectives. Thereby, the article has outstanding educational value for those attempting to find a start with the method.

With the rise of computational approaches, inverse dynamic modelling of human gait was facilitated by transforming to the frequency domain (Ren et al., 2007); pattern recognition and classification methods were applied (e.g. Mostayed et al., 2008). Along these lines, Fourier Series was used to get a frequency spectrum and compute development- and pathology-related change in overall locomotor frequency (e.g. Kurz et al., 2011).

Multi-dimensional variants of Fourier analysis for two joint angles in phase plots or “cyclograms” (DiBerardino III et al., 2010) or for 3D joint angles using quaternions (Kenwright, 2015) point at promising future directions.

Finally, recent work by the author (Ch. 3 and 5) has expanded on the aspect of multi-variate analysis and favorable data properties in the frequency domain (Mielke et al., 2019, 2023).

Overall, although some references exist, the studies using Fourier Analysis are sparse among locomotion literature. Authors who apply the technique tend to emphasize its usefulness, but often seem to be unaware of previous implementations. Few studies have explored downstream analysis of the Fourier coefficients. Confusions about mathematical terminology (e.g. “series” or “transform”, “transformation” or “regression”) complicate the situation. To summarize, variability in the use of Fourier Series arises from the following questions:

- Which signals to use? (joint/segment angles)
- How to apply the transformation? (no regression required)
- Which form and implementation to use? (“Series” vs. “transform”, “trigonometric” versus “complex exponential” form)
- If and how to process data in the frequency domain? (multivariate analysis, modeling)
- How many coefficients to retain? (order)
- The possibility of inverting FSD has been neglected, maybe in consequence of the above.

All these questions are the purpose of this review, and will be addressed in the next sections.

2.4 Properties of the Fourier Series

A lot of the text below is written from the applied perspective of a biologist, who used the transformation and its inverse to actual data. Not all of the reported properties of Fourier methods are directly linked to literature references, but most general courses and textbooks cover these concepts (e.g. Bracewell, 2000; Osgood, 2007, 2019).

2.4.1 Not a Regression

As pointed out above, some researchers have applied regression algorithms to extract the weights of the harmonics contributing to the signal of interest. A regression is an iterative procedure which attempts to match a known, parametrized function (the model) to a target signal (the measurement) by adjusting the parameters on each iteration with clever update rules. An intuitive example is linear regression, in which the parameters ‘a’ and ‘b’ of a function $y = a + b \cdot x$ are found so that the formula best describes the observed data relation of ‘y’ and ‘x’. Typical optimization algorithms are the Nelder-Mead simplex (Nelder and Mead, 1965) or variants of the BFGS algorithm (Broyden, 1970; Fletcher, 1970; Goldfrab, 1970; Shanno, 1970; Zhu et al., 1997). Though it is immensely useful, regression has several technical disadvantages: convergence is not guaranteed, start values are required and might affect the result, the algorithm might get stuck in local optima, accuracy is finite, the iterative procedure is computationally expensive, and reversibility requires the inverse function or another regression. Some of these disadvantages are especially hindering when working with trigonometric functions (e.g. inversion).

Fourier Series Decomposition is not a regression. Determining the “parameters”, i.e. Fourier coefficients, works by eqn. (2.1), which can be directly translated to computer code (see above and supplement 2.7.1). The procedure is deterministic, exact, and reversible; no start values are required. The trigonometric formulas provided in many publications resemble regression models, but applying a regression procedure is not necessary. I will demonstrate the overall procedure and provide code for some relevant scientific scripting languages as proof of concept.

2.4.2 Relation to Fourier Transform

As discussed above, Fourier Series Decomposition (FSD) requires a signal of finite length which starts and ends at the same value (periodicity). It computes a discrete spectrum of the whole function, i.e. the amplitudes for each harmonic. A generalization of FSD is the Fourier Transform, which does not assume periodicity, and which can be applied to signals of indefinite length by shifting a computation window along the signal. Fourier Transform yields continuous spectra, measuring frequency content in units related to absolute time. When applied in a sliding window, Fourier Transform provides a spectrogram of the signal,

i.e. the changing contribution of different frequency components to the signal over time. The minimum frequency is limited by the window width, the maximum frequency is determined by the sampling rate of the signal (Shannon, 1949). Fast Fourier Transform (FFT) is the most common algorithm for computation (*cf.* Heideman et al., 1984), and FFT is readily available in most programming environments. Note that there are many other such transforms which circumvent some limitations of FFT or might be otherwise useful in research, such as the S-Transform (Brown et al., 2010; Stockwell et al., 1996) or Empirical Mode Decomposition with the Hilbert-Huang Transform (Huang et al., 1998).

The result of FFT on a single window is a quasi-continuous frequency spectrum, and the frequencies are related to absolute time. Because of this, joint angle profiles which look exactly identical, but differ in stride duration, would have different frequency spectra. In contrast, FSD only considers harmonics, which are the integer multiples of the signal period. Therefore, the outcome of FSD is usually a much shorter array of numbers (Fourier coefficients). The FSD formula involves a normalization for the signal period (i.e. stride cycle duration), and thus strides of different duration can be directly compared. FSD also avoids problems of FFT which are due to windowing. Though both procedures could be applied to stride cycle kinematics, **FSD is favorable if the periodicity/cyclicity assumption holds.**

2.4.3 Time Scale Independence

Another favorable property follows from the cyclicity: **Fourier coefficients are independent of temporal scale or sampling.** Due to the normalization mentioned above, the FSD of the signal does not contain any information about the original duration of the signal (e.g. stride duration), nor does it matter much where the cycle started. On the inverse FSD, a new time frame and temporal sampling must be chosen. It is thus possible to re-sample the time line of a measurement by transforming it to the frequency domain, and then back with a different timeline.

In consequence, FSD is useful for homogenizing and comparing signals of variable duration. This can be advantageous, if groups or paradigms of variable stride durations are to be compared. If discarding that information is unintended, “duration” can simply be retained as a separate variable. The same holds for speed and other spatiotemporal gait variables which are related to the temporal structure of the signal.

It follows that **Fourier coefficients do not replace spatiotemporal gait variables (e.g. duration, speed), but complement them.** In case of joint angle profiles, Fourier coefficients capture the kinematics, while gait variables quantify the system outcome as a whole. Both are related by morphology (i.e. limb segment dimensions), which is another relevant class of variables that can and should be associated with kinematics in modeling approaches (*cf.* Ch. 4).

2.4.4 Approximate Periodicity

How “flexible” is the periodicity assumption? Motor behavior underlies variability (Bernstein, 1935), thus one might never measure a perfect end-start match. Fourier Series Decomposition works on noisy signals, yet it should be assured (as a technical requirement) that the end-start-difference is smaller than the white noise. There are a few common strategies to accomplish this if it is not given, and it has to be discussed how these strategies affect the outcome.

Firstly, one could ignore a start-end-difference, and bluntly apply FSD. The algorithm assumes periodicity, but the formula has no means to check this assumption, and will hence “connect” the last and first sample. This would lead to an abrupt change, which demands high order harmonics, which will show in the spectrum. However, abrupt changes from discontinuities are not physiological, so those values can be considered an artifact.

As a second option, it is common practice to generate an inverted version of the signal (of the form $g(t) = -f(-t)$) and append it to the end of the original signal. This could be visualized as a rotation of a duplicate of the signal about its last measurement point, which results in the values progressing first in forward and then inverted in backward measurement order, all changes accumulating to zero, and the “double” signal ending where it started. This method is common in the Fourier analysis of ground reaction force measurements (Alexander and Jayes, 1980; Schneider and Chao, 1983), where data usually starts and ends at zero.

The procedure shifts harmonics up by one due to period doubling. The rationale of using this method is that it generates symmetry and smoothness, which generally reduces the order (i.e. number of harmonics) required to retain all relevant information of the signal.

A third method to mend end-start differences is to simply spread them over the whole period of the signal (Mielke et al., 2019). This can be achieved in a computationally efficient way by subtracting an array of numbers from the signal which is linearly increasing from zero to the end-start difference, and which has the same number of samples as the signal (pseudo-algorithm: `signal -= linspace(0, end-start, number_of_samples)`). For this to be valid, it must be ensured that the end-start-difference is negligible in magnitude, non-physiological, and non-systematic. If those are givens, spreading the difference over the cycle is least invasive with regard to the outcome coefficients. As with the previous method, this smooths the signal and reduces higher order coefficients, yet without the caveat of altering the relation of relevant harmonics.

If used with caution, the third method is the favorable one for applications on locomotor kinematics. However, with that sort of data, there is another relevant method to minimize end-start-differences. The **choice of where the start and end of a cycle lie** is conventionally made by finding the (left hind-limb) touchdown. This time point is more or less distinctly visible on videos, and might be a requirement for data comparison. Yet, mathematically, that is just an arbitrary point in the cycle, and biomechanically, the

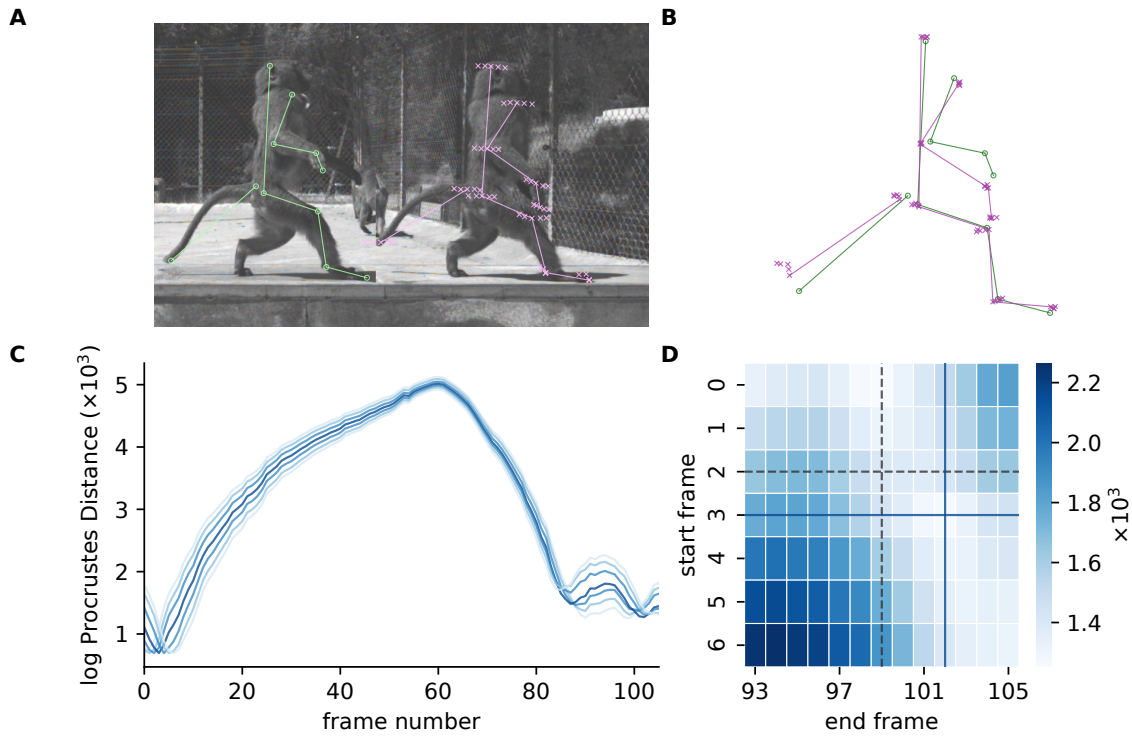


Figure 2.3: Stride cycle end-start matching. A stride cycle can be defined as the time interval between two frames which have the highest similarity in joint configuration and enclose exactly one swing and stance phase. (A) Putative start and end frame configurations, superimposed onto the original video frames. Several frames are candidates for cycle end, indicated by the series of cross markers. (B) Superimposition of end frame candidates onto the reference start frame, using Procrustes Superimposition. Only the axial line points (head to toe) are used for superimposition. (C) Configuration difference (measured as Procrustes distance pd , shifted by the arbitrary formula $\log(pd * 10^3 + 2)$ for visualization) over the whole stride cycle for different candidate start frames. (D) A heat-map of configuration difference can identify mathematically optimal stride cycle intervals. Dashed black crosshair: cycle start and end, conventionally determined by limb touch down. Blue crosshair: start and end frame with maximal configuration similarity, i.e. minimal Procrustes distance of the configurations at candidate frames.

configuration of the limb can be different on two touch downs (two strides might instead have higher cross-resemblance just before touchdown, or at mid-stance; Fig. 2.3). The term configuration herein describes the relative spatial arrangement of a set of points of interest or markers. Thus, if emphasis of the analysis lies more on the cyclic character of the stride, and if researchers find it valid to choose a different point for start and end of the cycle, then optimal points could be found automatically by cross-comparing the configurations of the whole limb in a sufficient frame range around the limb touch downs (as applied in Mielke et al., 2019, 2023). This can be achieved by taking all relevant points tracked in the videos (treating them in each frame as a geometric shape; Fig. 2.3B) and calculating the Procrustes Distance between those frames or shapes (Rohlf and Slice, 1990). That Procrustes Distance is a direct measure for the configuration difference between frames (Fig. 2.3C). In a cross-comparison, the two frames which have least Procrustes Distance are candidates for start and end of the cycle (Fig. 2.3D), and should be verified by visually superimposing the video frames. Conversely, whether or not there are two such frames with a sub-threshold Procrustes Distance, and at what time in the cycle they appear, can be a useful proxy to confirm steady-state locomotion.

All these are technical tricks which usually improve the data set quality and consistency. The bottom line is that, in steady state locomotion, the **periodicity assumption is usually met**, and there exist minimally invasive adjustments to ensure that the data strictly meets the mathematical requirements of FSD.

2.4.5 Reversibility

Some data operations can be undone (math.: invertible, i.e. reversible), others not. For example, consider the Euclidean vector norm of a temporal series of velocity vector measurements: the resulting number (speed) can quantify the magnitude of the original velocity vector, but not its direction. Conversely, when modeling speeds depending on some other parameter and predicting magnitudes, or when averaging speeds, one could not infer a direction. Information is lost in the transformation from a three-dimensional vector to a single number representing its length.

Other operations are invertible: consider again a series of velocity vectors of a physical object moving in space. One can freely change the reference coordinate system. Moving from an earth-bound reference frame to the coordinates defined by the object and back is possible without loss of information. Another common, invertible coordinate transformation is the Principal Component Analysis, which transforms data into the space spanned by the orthogonal eigenvectors of a data covariance matrix. Because these transformations retain all information, and because an inverse mathematical operation exists, one can move the given data forth and back between the different spaces or domains. Even more, one can transform or untransform unrelated, new data between spaces, although that data was not

involved in defining those spaces. This feature is useful and commonly applied to averages and extrapolations.

Fourier Series Decomposition is invertible. As described above, it translates the data from a “time-dependent array” form to a time-independent, complex-valued frequency space. Any set of values, be they observed or synthetic, can also be translated back via an inverse operation. In contrast to PCA, this does not even require the eigenvectors of the original data (there is only one frequency domain). Some operations, such as averaging of multiple signals, are equivalent in the time- and frequency domain.

The formula for the inverse Fourier Series is the following:

$$f(t) = \sum_{n=0}^N (2 \cdot c_n) \cdot e^{2\pi i n \frac{t}{T}} \quad (2.2)$$

Notation herein as in (2.1), with N being the order (number of coefficients). Computer code for application in R, Matlab and Python can again be found in supplement 2.7.1 and online (https://git.sr.ht/~falk/fcas_code).

The existence of this inverse formula means that just as exchanging a coordinate basis for linear coordinates, **one can convert freely between time- and frequency domain**.

However, one limiting factor is the order of the Fourier Series. Some signals (i.e. those with sharp turns or quick changes) require a high number of harmonics to be accurately represented in frequency space. Conversely, when operating with a finite order, some signals might be filtered on the first transformation to the (pruned) frequency domain. Real measurements are usually subject to measurement noise, such noise involves changes as quick as measurement sampling, which are thus lost in FSD. This **filtering property of FSD** can represent a loss of information (if the noise is considered informative, but note that in that case the residual after re-transformation might be even more informative).

However, in cases where this filtering effect is negligible or even favorable, FSD and its inversion can be applied. Repeated back- and forward transformation does not discard further information. When implementing a version of the equations above in computer code, a good check is whether values are unchanged after applying FSD and its inversion in series. And because of the relatively few coefficients needed for accurate representation of physical processes, compared to sampled timelines, it is often **efficient to store** kinematic data in the form of FSD coefficients.

2.4.6 Choice of Order

How to decide how many coefficients should be retained? That number is the “order” of the Fourier Series. It can be easily determined by exploiting the reversibility of a method. Each observation (i.e. each measured joint angle profile) should be converted to the frequency domain with a given order, and then converted back to the time domain. The

(root-mean-square) difference of the original signal and the re-transformed one should be small in magnitude and normally distributed around zero. Additionally, original and re-transformed signals should be plotted on top of each other for visual inspection.

This strategy is of general use, and I will demonstrate it on the test case (Ch. 2.5.1). In the particular case of joint angle kinematics, it has often been concluded that relatively few coefficients are sufficient to capture the essence of the phenomenon. This is physically plausible, because the elements of limbs are rigid bodies with a certain inertia. Such elements cannot perform too abrupt accelerations, and in consequence, the profiles are smooth. Furthermore, in the case of steady-state locomotion, the ensemble of rigid body elements produces the behavior, and it is unlikely that any single element can oscillate an order of magnitude quicker than the whole limb. The stride cycle is the defining time interval, and normally all elements move in relative unison, which limits amplitude in the higher order coefficients.

2.4.7 Affine Components

Some attributes of a signal are emphasized by Fourier methods. Those can be summarized intuitively as those attributes of the signal which can be changed without altering the “perceived shape” of the signal when plotted.

The most obvious one is the average of the signal over time, i.e. its **mean value**. When changing the mean of a signal in the time domain by adding the same scalar value to every sample, the signal shifts “up and down”, but retains its temporal structure. The mean is completely captured by the zero’t h Fourier coefficient, which therefore is always a real number. Apart from that one number, changing the mean of a signal leaves its frequency domain representation unchanged.

Another attribute of the signal is the **amplitude**, or how much values change around the mean. In the time domain, amplitude is altered by centering the signal and multiplying it with a scalar, followed by un-centering. In the frequency domain, amplitude is visible as the distance of coefficients from the origin of the complex plane (i.e. the cumulative magnitude of the complex numbers, or the norm of the complex coefficient vector).

The third special signal aspect is **phase**, and it has to do with the periodicity of the signal. Phase is quantized in the time domain by sampling, and it can be changed by taking a number of samples from the end of the signal and appending it to the start (“rolling” the signal around, or changing the start point of the cycle). In the frequency domain, changing the phase rotates the coefficients in the complex plane (but note that higher order coefficients rotate exactly n times quicker). Because the frequency domain is independent of sampling, phase can be changed by any scalar number here, and is not limited to the sample raster. This allows efficient temporal resampling, as well as an optimal alignment of multiple signals (*cf.* Ch. 3.7.2). Phase is an angle, best used in the range of zero and 2π . It is changed in the

frequency domain, just as any rotation in a complex plane, by multiplication of a complex exponential (“delay/shift theorem”, *cf.* Bracewell, 2000).

The fact that amplitude is the distance and phase the angle of Fourier coefficients in the complex valued frequency space illustrates the relation of the exponential and amplitude-phase form of the Fourier formula: the latter are just the corresponding polar coordinates.

Mean, amplitude and phase are called **affine components** of a signal, which describes the fact that they can be adjusted by scalar operations without altering the “signal shape” (see Ch. 3). They can even be standardized (mean: zero, amplitude: one, phase: zero). They can be associated with biological meaning: the mean quantifies dynamic posture, the amplitude is related to effective range of motion, the phase quantifies relative timing of joint movements. What remains after standardization, i.e. after isolation of the affine components, is in a way the essence of a signal, its “shape”, which is defined by the temporal (or frequential) structure of the behavior. In the case of joint angle profiles, this remainder can be considered coordination *sensu strictu* (Mielke et al., 2019, 2023).

2.4.8 Multivariate Analysis

The numeric representation in the frequency domain is usually shorter, and as mentioned above, storage might be efficient. Nevertheless, an FSD of the order N will yield $N + 1$ coefficients, which are $2N + 1$ numbers when splitting up their real and imaginary parts. Although we usually find an order of $N < 10$ appropriate (Ch. 2.4.6), this would still leave a considerable amount of variables for analysis.

This raises the question of the effect of multivariate analysis methods, such as Principal Component Analysis (PCA). PCA is a coordinate transformation which finds orthogonal coordinate axes in the data set that are oriented towards the largest variability within the data (MacLeod, 2007). A common purpose of PCA is dimensionality reduction, and it is particularly effective if there are strong co-variations within the data. It seems clear that Fourier coefficients of a single joint angle profile are intrinsically linked, for example through the phase rotation in the complex plane (exponential form). However, it is not obvious whether that holds any advantages for PCA, because the phase rotation does not cause linear correlation. On the other hand, higher coefficients are usually of lower amplitude than the main coefficients, and thus their putatively relevant variability might be lost in a PCA-based dimensionality reduction. Thus, PCA does not generally hold benefits for the analysis of single joint FSD in its complex exponential formulation. It might be different in the amplitude-phase formulation, which should be explored in future research.

If more than one joint is of interest, coefficient number is multiplied, but the situation for multivariate analysis changes. Different joints are often interrelated through adjacency along their linking segments. Biarticular muscles and tendons can cause correlated movements, which makes top-down sense, because a typical vertebrate limb during swing phase tends to

be extended and flexed altogether (Fischer and Blickhan, 2006). This situation is favorable for PCA, in a sense that dimensionality reduction can often be achieved with relatively little information loss. Note, however, that PCA must be adjusted in a way that accounts for the FSD properties as follows. It is common practice and often advisable to standardize input variables prior to PCA, so that their value ranges are comparable (usually done by subtraction of the mean and division of a variability measure). Such a standardization would disrupt the temporal structure of the signal, and emphasize nuisance variability in higher order coefficients. A better strategy for multi-joint analyses is the standardization by extraction and isolation of affine FSD components. As mentioned above, mean joint angle and joint amplitude can be standardized, and phase differences of all observations can be minimized by temporal alignment. With such preprocessing, the outcome of a PCA will equally represent each joint of interest, without disrupting the temporal structure of the joint angle profiles, and benefiting from intrinsic correlations of movements of the limb elements. PCA is invertible, just as FSD, and so any downstream modeling outcomes can be related back to the original joint angle profiles.

2.5 Case Study: Predicting Bipedal Walking in Adult Male Baboons

Above, I re-iterated some fundamental properties and best practices related to Fourier Analysis, and Fourier Series Decomposition in particular. I will now outline the typical workflow of the analysis of kinematics with that method, focusing only on the steps which are affected by the Fourier method. There is a lot of flexibility in the procedure (see Fig. 2.4 for a coarse guideline), and researchers should feel encouraged to adapt the steps to their needs. For illustration, I applied the method to a previously analyzed data set of bipedal walking in olive baboons, *Papio anubis* (Druelle et al., 2021).

2.5.1 Data Preparation

After data acquisition is complete and hard drives are filled with videos, the individual stride episodes which were captured on video must be annotated and extracted. As a good approximation, one can use the conventional touch down timepoints (though lift off or mid stance would work equally well, and the choice may vary within the dataset).

A prerequisite for any meaningful analysis is consistent, complete data. Data in this case are points of interest, tracked on videos. There are many tools for landmark tracking on video, which are more or less automatic and accurate (Crall et al., 2015; Dunn et al., 2021; Hedrick, 2008; Karashchuk et al., 2021; Knörlein et al., 2016; Mathis et al., 2018; Mielke et al., 2020). It should be assured that no frames have been involuntarily skipped in the process of tracking, and that there are no discontinuities in the series of point positions. This can be easily achieved by plotting the time series of tracked points, either in the global reference frame, or in a moving reference associated with the subject (Fig. 2.5). After landmark tracking, the aforementioned calculation of Procrustes distances between putative start and end frames (Ch. 2.4.4) can be used to refine the stride interval and minimize end-start-differences.

With strides identified and data quality assured, collective variables such as duration, speed, clearance, or gait asymmetry can be stored for each stride observed in the dataset.

Next, joint angles can be calculated in a variety of ways. Their precise definition and direction is arbitrary, yet should be consistent and clearly documented. Care should be taken that joint angles do not show jumps at any time point of the movement, for example by crossing the 2π angular interval border. Such jumps would lead to a wrong FSD, because that procedure does not natively consider angular value ranges or interval wrapping and thus treats the jump as if it were an actual discontinuity. The joint angle definition chosen herein, with fully extended limb corresponding to a joint angle of zero, and ranging from $-\pi$ to $+\pi$, usually avoids such discontinuities.

In the case of two-dimensional joint coordinates, calculation of joint angles is relatively

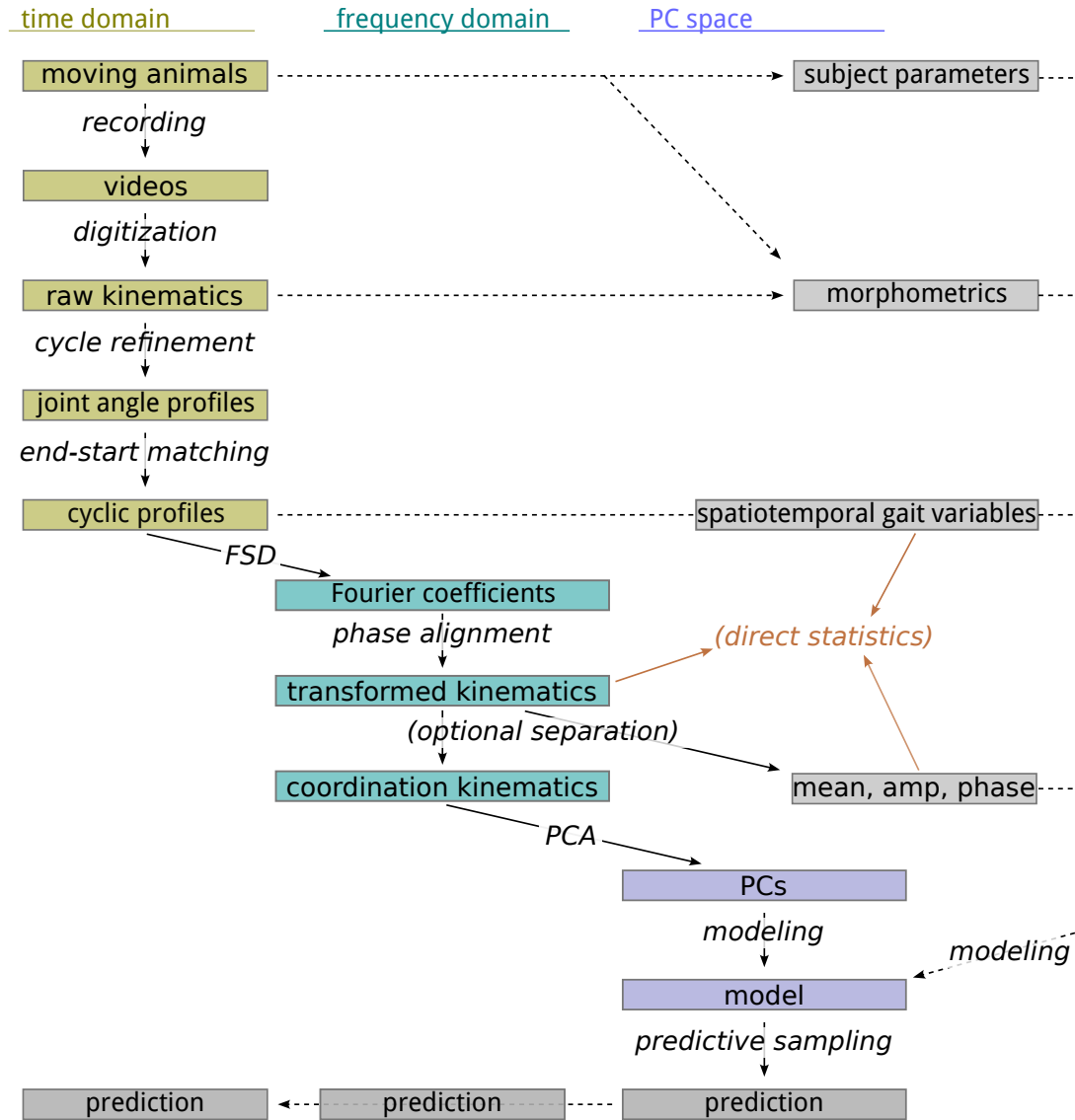


Figure 2.4: Analysis and modeling procedure. Workflow overview of the preparation, transformation and analysis steps involved in the data analysis on baboon bipedal walking as described in the text, from raw observation (top) to statistical inference and multivariate analysis (bottom). At each point of the procedure, transformation between the domains is possible and quality checks should be performed.

straight forward with an `arctan` formula. When working in three dimensions, the question is whether sufficient anatomical reference points can be tracked to calculate actual anatomical joint angles. If not, it must be sufficient to calculate a single joint angle between the segment vectors, though that does not necessarily correspond to anatomical degrees of freedom. If instead there are sufficient reference points, multiple joint rotation axes and angles per joint can be used and integrated just as in a multi-joint analysis. Note that there have been attempts to directly calculate Fourier coefficients from 3D joint angles, using quaternions (Kenwright, 2015), which would be the most elegant solution for this purpose.

After removing end-start-differences as described above (Ch. 2.4.4), joint angle profiles are plugged into the Fourier equation (2.1), e.g. by using the corresponding programming functions supplemented to this manuscript. Thereby, joint angle profiles are transformed to the frequency domain. Though I above emphasized the complex values of the coefficients, they can as well be represented by a one-dimensional array of alternating real- and imaginary parts of the coefficients. The transformation should be inverted, and original and re-transformed joint angle profiles plotted together, to exclude coding errors and confirm that order was chosen high enough (Figs. 2.6, 2.7).

2.5.2 Multivariate analysis

Each stride processed in the analysis counts as exactly one data point, and so far, data points were considered independently. The conventional next step is to analyze their relation to each other.

For example, it might be apparent that two or more data points are phase shifted relative to each other: all the characteristic maxima and minima of the joint angle profile would consistently appear later or earlier in the time domain plot. If one assumed that strides are cyclic, it must be concluded that these phase shifts are due to our arbitrary choice of the start point within the cycle. Phase differences should therefore be minimized. However, different joints, when analyzed together, might theoretically give contradicting information on the relative phase shift of different strides. Obvious solutions are to either choose a reference joint angle for phase alignment (this might be a kinematic angle not otherwise used for analysis: in the baboon test case, I chose the “total limb”, i.e. the head-hip-toe angle), or to use an amplitude-weighted average phase shift as calculated from all joint angles of interest together.

And just as with phase, there are design choices about whether to keep the other affine components (mean, amplitude) implicit in the multivariate data set, or extract and isolate them for subsequent analysis.

Independent of these choices, the result of the previous steps is a data table consisting of multiple observations (strides, in rows) and variables (Fourier coefficients, in columns). The observations can be related to their corresponding master data (i.e. subject characteristics,

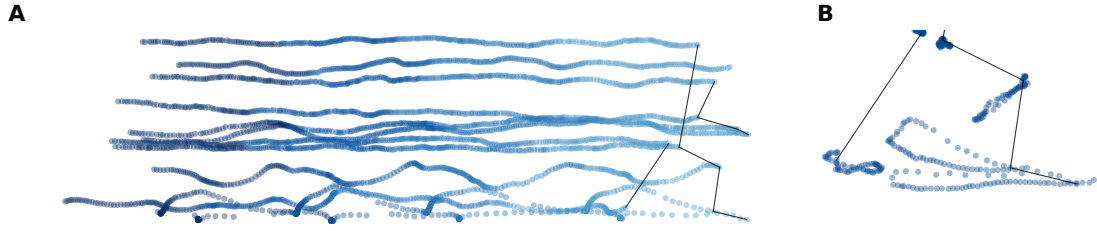


Figure 2.5: Raw data inspection. (A) Plotting a series of stride cycles in the camera reference frame can aid the identification of data discontinuities (visible here in the distal limb markers). The animal is moving from left to right. A stick figure displayed for the last frame facilitates landmark attribution (torso, tail, hind- and forelimb are shown, similar to Fig. 2.3). (B) Plotting one stride in the moving reference frame of the subject (zoomed in on the limbs) can confirm cyclic/steady state movement.

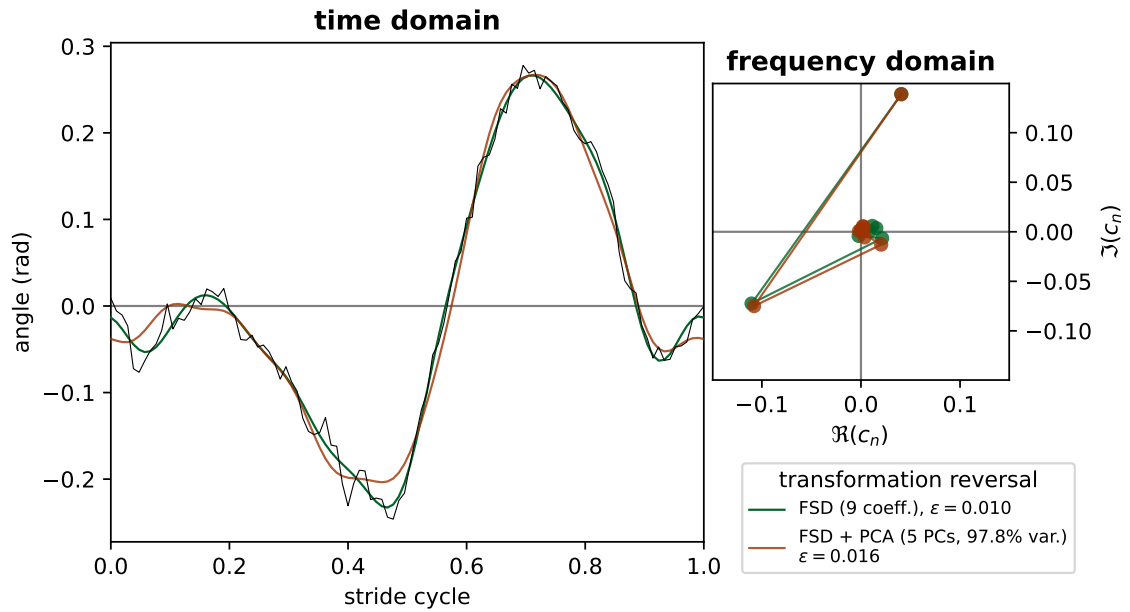


Figure 2.6: Reversibility. A single joint angle profile, processed forth and back with Fourier Series Decomposition (green, using 9 coefficients) and Principal Component Analysis (orange, using first 5 principal components, after FSD). The residual ϵ is the mean of Euclidean distances of all joint angle measurements over time from their corresponding re-transformation in the time domain. See also Fig. 2.7.

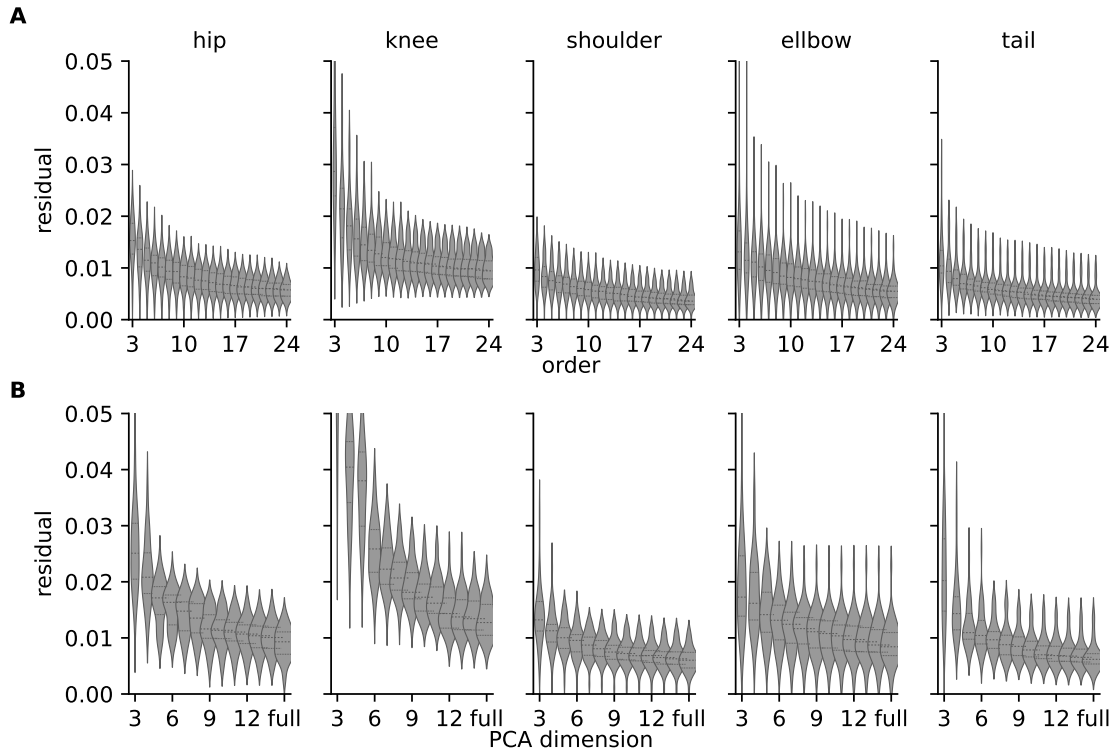


Figure 2.7: Retransformation residuals, (A) after performing FSD only with a given order (i.e. number of coefficients, x-axis), (B) after FSD (9 coefficients) and PCA with a given number of retained components. Residuals ϵ (y-axis) as defined above. Joint angle profiles of all observed baboon strides are included, the distribution of residuals is indicated by grey “violins”. Relatively low numbers of coefficients and components are sufficient to get close to the asymptotic accuracy. The absolute residual is joint-dependent (compare hip and knee, for example), an effect which is primarily determined by landmark tracking accuracy and measurement noise. The data point for “full” PCA dimension is the reference value with just the FSD.

morphology, collective variables, isolated affine components) by a simple stride index lookup. Such a data structure is eligible for multidimensional analysis methods, and one of the simplest such method is Principal Component Analysis (PCA). Often, PCA justifies a significant reduction of data dimensionality (i.e. number of data columns), depending on how much variance is concentrated on the first components. Apart from the residual variance not covered by the retained principal components, PCA is again information preserving and reversible (which should be confirmed, Figs. 2.6, 2.7). Most often, PCA is just used to prepare the data for subsequent multivariate methods (e.g. factor analysis).

To summarize: in the procedure drafted above, I have extracted and quality checked kinematic measurements from raw video digitization. Joint angles were calculated and submitted to two transformation procedures: FSD and PCA. All procedures to this point can be performed without loss of information: any of the resultant data rows could be converted back to an animation of moving points. Thus, the PCA outcome essentially contains the whole of what was captured by the original kinematic data: the spatiotemporal coordination of the moving body appendages of interest.

2.5.3 Statistics and Modeling

Despite the direct link to the raw data, the data table resulting from PCA might seem abstract. Nevertheless, those values are useful, because they are much more compact than the original two-dimensional time series of varying length. And this compactness is crucial for statistical testing and modeling, for which computational complexity can be restrictive. As a proof of concept, I herein briefly present the outcome of one type of analysis approach: probabilistic modeling (to be discussed in all detail in Ch. 4). The two major advantages are that (1) probabilistic models capture the variability of the intrinsically variable process of locomotion, (2) such models can be used for extrapolation (out-of-sample prediction).

The usual modeling steps are:

- data simulation (prior to acquisition; can provide valuable information on required sample size, feasibility, and model structure)
- model construction
- (MCMC) sampling or “model fitting”
- model comparison and refinement
- posterior checks (model “hygiene”)
- predictive sampling

I applied all these to the baboon data set. In total, 40 stride cycles from 17 subject individuals entered the analysis. I applied a stepwise modeling approach, modeling the PCA-transformed

Fourier coefficients (θ) generated from a set of joint angles (hip, knee, and ankle) as a function of sex (**male**), age class (**adol**, **inft**), body mass (**cbm** /centered), limb length (**ll**), clearance (**clr**), duty factor (**df**), trunk angle (**trnk**) and speed-related parameters (**str**, from a PCA of stride duration, length, speed and frequency).

$$\begin{aligned} \theta_i \sim & v_{1,i} \cdot \alpha_i + \\ & + v_{male} \cdot \beta_{male,i} + v_{adol} \cdot \beta_{adol,i} + v_{inft} \cdot \beta_{inft,i} + v_{cbm} \cdot \beta_{cbm,i} + v_{ll} \cdot \beta_{ll,i} + \\ & + v_{clr} \cdot \beta_{clr,i} + v_{df} \cdot \beta_{df,i} + v_{trnk} \cdot \beta_{trnk,i} + v_{str1} \cdot \beta_{str1,i} + v_{str2} \cdot \beta_{str2,i} + \\ & + \epsilon_i \end{aligned} \quad (2.3)$$

In the case of the baboon data set, I was able to successfully train this complex model despite limited sample size. I then confirmed model convergence and ensured that the model is favorable over alternative models with more or less parameters. The implementation in PyMC (a Python library, <https://www.pymc.io>) has the capability of posterior predictive sampling: the trained model can be used to generate an arbitrarily high number of virtual data points, which underlie the same variability as the original data. Most notably, this includes predicting “out-of-sample”, i.e. parameter combinations which were not directly observed (in this case, male adult baboons were not included in the data, but could be predicted; Fig. 2.8). Though the model infers abstract PCA values, the much emphasized reversibility of the method enables the computation of joint angle profiles from the predicted values. All data and documented code for all the steps described above are available online (https://git.sr.ht/~falk/papio_fcas). I will return to this data set and extend the explanation of the modeling procedure later in this thesis (Ch. 4).

This modeling and prediction is complementary to and consistent with the analysis of Druelle et al. (2021). A targeted model design could for example serve to infer effects of age class, speed, or their interaction, as was done in the original treatment of this data set. Such research questions can be addressed without transformation to the frequency domain. However, the point highlighted here is that the frequency domain data retains almost the full kinematic information, and thereby enables assessing a broader range of quantitative analysis questions, and predictive modeling of joint angle profiles and coordination.

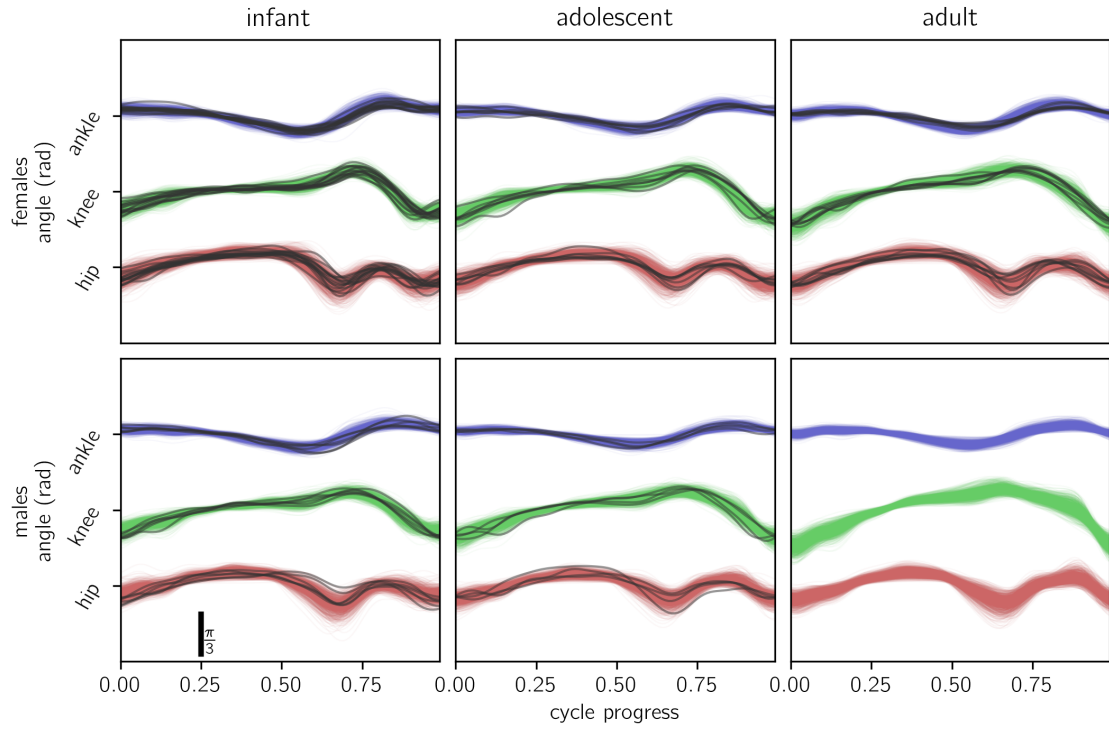


Figure 2.8: *Posterior predictive sampling.* A probabilistic model which is trained on the kinematic data (dark grey lines) is capable of predicting joint angle profiles (colored, thin lines; 1000 predictions per category). This can be extrapolated, for example to unobserved category combinations (here: adult males, which were not part of the dataset). Model design and training are enabled by transformation of the data to a PCA-space of the frequency domain. Joint angle profiles are centered around their mean for visualization; black bar in the lower left plot indicates angular units.

2.6 Summary

When reviewing prior attempts to use Fourier-based methods for the analysis of kinematics, a lack of consistency becomes apparent. Prior studies using the method exist, but hardly reference each other, which indicates independent events of lateral Fourier gene transfer into the population of biologists. Authors seem to be unaware of the different available methods and how to apply them efficiently. Some of the choices for application are of little relevance for the outcome of the analysis (e.g. whether to use the sine-cosine, exponential, or phase-amplitude form of Fourier Series), whereas others have severe impact (e.g. choice of order). I pointed out that Fourier methods do not require a regression algorithm for data conversion, and that often the most appropriate tool in the Fourier toolbox to analyze joint angle profiles is Fourier Series (and not FFT). Usually, few coefficients are sufficient to capture all relevant kinematic information, and multivariate methods such as PCA are readily available for further dimensionality reduction. Some properties of joint angle profiles become directly accessible by FSD (affine components), and a mathematically precise phase alignment is possible.

Yet, as demonstrated above, I would argue that the biggest advantage of transforming the data to the frequency domain is that it enables the quantitative analysis of coordination. Though joint angle profiles could theoretically be resampled and submitted to multivariate analysis directly, this approach would face major technical challenges if sample size is limited, or data points are structurally different, misaligned, or undersampled (all of which is practically always the case in comparative kinematics). FSD circumvents these problems by providing an elegant, concise representation of the data: a set of relatively few, meaningful complex coefficients which can be converted back to the original data.

This availability for quantitative analysis has far-reaching potential. I demonstrated it on a relatively small, but well studied data set of bipedal walking in baboons, and was able to predict stride cycles for a co-incidentally unobserved combination of subject attributes. Using phylogenetic and morphometric bracketing, the same procedure could be applied to infer the locomotion of extinct species, e.g. hominids. Another potential purpose of this quantitative method is its combination with dynamics, i.e. force and moment measurements: consider a perturbation analysis, where normal locomotion is interrupted, and a subject can succeed or not in maintaining dynamic balance. Certainly, success will depend on the specific movement of some limb elements, and quantitative kinematic modeling can highlight which.

I will get back to and extend the use of FSD and the probabilistic modeling approach in the following chapters (Ch. 4, Ch. 5).

None of these analysis components are novel. Fourier Analysis is a standard tool in physics and engineering. I see equally great potential for the biological sciences, where cyclic phenomena are abundant (genetics, physiology, behavior, ...). I suspect that the reason for the limited use of these tools to date is limited accessibility, caused by terminological

confusion and the lack of attention for example studies. I referenced some excellent work that made use of the Fourier theorem (Bernstein, 1935; Pike and Alexander, 2002; Webb and Sparrow, 2007), as well as my own attempts to extend the capability of the technique in the context of kinematics (Mielke et al., 2019, 2023). Thereby, I hope that this review can increase availability of the Fourier method to biological sciences, potentially even beyond the study of locomotor kinematics.

2.7 Supplements

2.7.1 Computer Code to Apply Fourier Series

This is an example how the Fourier Series can be implemented in the Python programming language (<https://python.org>).

Similar implementations for R and Matlab are provided “ready-to-use” on the following git repository:

https://git.sr.ht/~falk/fcas_code

```
def FourierSeriesDecomposition(time, signal, order):
    # calculate the Fourier Series decomposition of a signal,
    # given the sample time array ("time")
    # and a chosen "order" (i.e. highest coefficient returned)
    # returns complex coefficients

    # the period of the signal
    period = numpy.max(time)-numpy.min(time)

    # number of samples taken
    n_samples = len(time)

    # the exponential formula for each coefficient
    SingleCoefficient = lambda t, T, n: numpy.exp(-1j*2*numpy.pi*n*t/T) / (2 if n == 0 else 1)

    # calculate the Fourier Series as a list of coefficients
    fsd = (2/period) * numpy.array([ \
        numpy.sum(signal*SingleCoefficient(time, period, n)) / n_samples \
        for n in range(order+1) \
    ])

    return fsd
```

```
def FourierSeriesRecomposition(coefficients , output_time):
    # reconstruct a signal from its frequency space representation
    # i.e. take coefficients list and return the signal at given output_time points

    # the exponential used in this formula
    FourierSummand = lambda t, T, cn, n: cn*numpy.exp(1j*2*numpy.pi*n*t/T)

    # the Fourier function at a single time point
    FourierFunctionT = lambda t, T, coefficients: numpy.sum(numpy.array([ FourierSummand(t, T, cn, n) \
                                                                           for n,cn in enumerate(coefficients) \
                                                                           ]))

    # the period "T" of the signal
    period = numpy.max(output_time)-numpy.min(output_time)

    # for every point in time, sum up the coefficients
    signal = period * numpy.array( [FourierFunctionT(t, period, coefficients) \
                                     for t in output_time \
                                     ])

    return numpy.real(signal)
```

Chapter 3

Fourier Coefficient Affine Superimposition

3.1 Abstract

Many phenomena related to motor behaviour in animals are spatially and temporally periodic, making them accessible for transformation to the frequency domain via Fourier Series. Although this has been applied previously, it had not been noticed that the characteristic arrangement of Fourier coefficients in their complex-valued representation resembles landmarks in geometric morphometrics. We define a superimposition procedure in the frequency domain which extracts and isolates affine differences (mean, amplitude, phase) to reveal and compare the shape of periodic kinematic measures. This procedure is conceptually linked to dynamic similarity, which can thereby be assessed on the level of individual limb elements. We demonstrate how to make intralimb coordination accessible for large scale, quantitative analyses. By applying this to a data set from terrestrial ungulates, dominant patterns in forelimb coordination during walking are identified. This analysis shows that typical strides of these animals differ mostly in how much the limbs are lifted in the presence or absence of obstructive substrate features. This is shown to be independent of morphological features. Besides revealing fundamental characteristics of ungulate locomotion, we argue that the suggested method is generally suitable for the large scale quantitative assessment of coordination and dynamics in periodic locomotor phenomena.

This chapter is a slightly modified version of a previously published article (Mielke et al., 2019). The publication preceded the review in Ch. 2, which unfortunately causes some redundancy. The concept of affine components and the analogy to geometric morphometrics are extended and applied to a comprehensive data set. I must apologize that the relevant publication by Pike and Alexander (2002) had completely slipped my attention prior to this publication, I was unaware of its existence, but added reference ex post.

Furthermore, it must be noted that this chapter, contrary to the title of the thesis, does not analyze developmental data. the variable “age” was not available for the animals shown on ‘YouTube’, and could not be included. No neonate or early juvenile data was incorporated. It is not clear whether age plays a role on this precocial group of animals. Due to the large sample size, I would expect effects to average out. In consequence, the intention of this chapter is explorative: I applied a novel method to a broad data set, inspecting and interpreting the outcome.

3.2 Introduction

A prevalent feature of motor behaviour of animals is temporal and spatial periodicity. Rhythmic or cyclical recurrent patterns can occur on all scales, from large (e.g. seasonal migration) to small (e.g. neural pattern generators, actin-myosin binding). Locomotion is one behavioural class which is highly relevant for the organism, and which often shows a considerable degree of recurrence, most prominently during steady-state locomotion. In consequence, locomotion was one of the first subjects of motor behaviour to be formally studied (Bernstein, 1927a,b; Braune and Fischer, 1904; Marey, 1878; Muybridge, 1893). In the study of terrestrial locomotion, limb kinematics (e.g. joint angles) and dynamics (e.g. joint moments, ground reaction forces) represent relevant periodic profiles accessible to comparative analyses. Groups for comparison can be defined by differences in various attributes, such as for example phylogeny, age and development, physiological background, or pathological status. Locomotion holds diagnostic potential in all of these: the periodic profiles are ultimately linked to the ecology, ontogeny, and morphology through evolutionary processes, and proximately modified by immediate external and internal constraints (e.g. Barrett et al., 2008; McGibbon, 2003; Mohling et al., 2014; Nyakatura et al., 2012, 2019; Vanden Hole et al., 2018a; Vanhooydonck et al., 2014).

The analysis of locomotor patterns can be insightful and has been applied in numerous cases, targeting different levels of detail. A common tool to quantitatively compare kinematics are spatiotemporal gait parameters, which are summary statistics over limbs, stride cycles, or both (see e.g. Biancardi and Minetti, 2012; Christiansen, 2002). The limitation of gait parameters is that they do not resolve differences in intralimb coordination. Intralimb coordination is a complex phenomenon, which has been difficult to capture in quantitative methods. In some hallmark studies, the comparative analysis of intralimb coordination has been performed by relatively sparse sampling from large groups, with nevertheless respectable workload (Fischer et al., 2002; Isler, 2005; Stoessel and Fischer, 2012). Gatesy and Pollard (2011) have attributed the limited quantitative accessibility of coordination to limitations in transferability of kinematic profiles. According to their theoretical considerations, the temporal change of a joint angle within a limb can hardly be disentangled from morphology, posture, and the interaction with adjacent joints, which makes it hard to transfer observations between groups. These authors call for novel methods to resolve these dependencies. We herein present such a method, and exemplify it on a broad sample of terrestrial ungulates. The backbone of our approach is Fourier Series decomposition (FSD). Its novelty lies in an attempt to superimpose temporal profiles in a manner analogous to geometric morphometric techniques, i.e. using FSD to extract the “shape” of cyclic curves.

To extract shapes from such temporal profiles (or any periodic signal in general), one requires data that resembles landmarks, as well as a procedure to superimpose

observations. Profile landmarks can be found by applying Fourier Series (Ch. 2). Fourier Series is a classical method in the study of kinematics (Bernstein, 1927a, 1935), although there are few recent studies exploring it further (Grasso et al., 2000; Pike and Alexander, 2002; van Weeren et al., 1992; Webb and Sparrow, 2007). Note that Fourier Series is related to, but different from other techniques previously applied in the context of locomotion (namely Fourier Transform and Elliptical Fourier Descriptors, see discussion). By applying FSD, the signal is decomposed into its harmonics, which describe it in what is called the frequency domain. The outcome of FSD is usually reported as sine- and cosine coefficients, which through Euler’s formula are equivalent to complex exponentials. As will be illustrated further, this latter representation has some advantages for our purpose. When plotted in the complex plane, Fourier coefficients have characteristic arrangements, which we find are to an extent analogous to landmarks in geometric morphometrics (Bookstein, 1991; Kendall, 1989). Furthermore, the *affine* components of a signal become mathematically accessible in the complex plane representation. Affine components are those that can be scaled linearly without altering the shape of the signal. In the case of complex Fourier coefficients, those are (i) the mean profile value (i.e. “y-value”) over time, captured as the zero’tth Fourier coefficient, (ii) the amplitude of the harmonics, visible as the distance from the origin of the complex plane, and (iii) the phase, which is related to a (non-trivial) rotation in the complex plane (*cf.* methods, Fig. 3.1). These terms are analogous to the parameters of a harmonic oscillation of frequency f over time t of the form “ $y = c_0 + A \cdot \sin(2\pi f \cdot t + \phi)$ ”. Such an oscillation has the mean c_0 , amplitude A , and the phase ϕ . Assume there is a mathematical operation which extracts the affine components of two such oscillations by shifting (i.e. changing c_0), scaling (changing A), and rotating (changing ϕ) the complex coefficients in the frequency domain to a standardized arrangement. This would be a Procrustes procedure which results in a mathematically optimal superimposition (two harmonic oscillations of the same frequency would become identical). The differences would not be removed, as in “discarded”, but isolated, and can still be part of subsequent analyses. We explore such a superimposition in this study, terming it Fourier Coefficient Affine Superimposition (FCAS).

The motivation behind this is the search for a quantitative, comparative measure of coordination of joints along a limb. For such measurements to be comparable, one should normalize the measures according to the dynamic similarity principle. Dynamic similarity, in general and in its biological application (*cf.* Alexander and Jayes, 1983; Vaughan and O’Malley, 2005), is also concerned with the isolation of affine components. In general, dynamic similarity applies when all spatial dimensions of two mobile systems scale with the same factor (i.e. geometric similarity), the temporal aspects of all involved movements scale with another, again uniform factor (i.e. similar coordination), and therefore all forces are similar except for a uniform third factor. Scaling the system by such factors represents an affine transformation. In the specific case of terrestrial locomotion, dynamic similarity

allows predictions about kinematics and kinetics of different animals (i.e. about leg phasing, relative stride length, duty factors, forces and power output; Alexander and Jayes, 1983). In practical application, a common purpose of the dynamic similarity principle is to compare animal locomotion at equal Froude numbers (e.g. Holmes et al., 2006; Steudel-Numbers and Weaver, 2006). Perfect dynamic similarity is rarely observed and not expected, but finding the reason for deviations from it can be instructive (e.g. Kramer and Sylvester, 2013; Raichlen et al., 2013; Steudel-Numbers and Weaver, 2006). Such studies are usually concerned with spatiotemporal gait variables. We attempt to apply the dynamic similarity principle on the level of joint angles, i.e. to measure intralimb coordination, using FCAS as a method to separate affine and non-affine differences in time profiles of joint angles.

A classical illustration of the concept of dynamic similarity is the case of terrestrial ungulates (Alexander and Jayes, 1983). These are a group of mammals that successfully populated an impressive variety of habitats, such as cluttered rain forests, swamps, rugged mountainsides, plain grasslands, and even deserts. Ungulates have achieved this without fundamental adjustments in the topology of the locomotor apparatus (Alexander, 1984; McMahon, 1975), albeit variation in their limb geometry. The musculoskeletal layout and posture of ungulates favour movement that is characterized by parasagittal limb excursions (Jenkins, 1971; Stein and Casinos, 1997). Therefore, their locomotor characteristics can be adequately approximated by time profiles of two-dimensional joint angles. Despite this simplification, their enormous diversity makes ungulates a challenging subject for conventional analyses of intralimb coordination. During steady state locomotion, joint angle profiles are approximately periodic, thus FSD and FCAS can be applied. We herein use ungulate kinematics as a case study for FCAS.

FCAS does not reveal intralimb coordination *per se*. However, the fact that superimposition operators (and with them affine differences between groups) are transferable to adjacent joints can be exploited (analogous to Mielke et al., 2018a). More specifically, in this study, we will perform superimposition of elbow joint angle profiles with respect to the carpus joint. This example application of FCAS yields *relative* joint angle profiles, which are quantitative measures of intralimb coordination because they contain the combined variance from two joints. Using the frequency domain representation, the relative profiles become accessible for multivariate analysis. These ingredients allow us to investigate intralimb coordination in a quantitative comparison of unprecedented phylogenetic scope, overcoming previous limitations (Gatesy and Pollard, 2011).

In this study, we focus on the practical application of FCAS and the insights it extracts about ungulate locomotion and evolution. We supplement a detailed mathematical description (supplements 3.7.1 and 3.7.2), an extensively commented, open computer code tutorial (supplementary data 3.7.5), and all data that was generated (supplementary data 3.7.5). We present data of walking gait, densely sampled across terrestrial ungulates (87 of

117 genera). For this purpose, we tracked landmarks on freely available video material from online video platforms, as others did before (Biancardi and Minetti, 2012; Lees et al., 2016). Rather than fixing an *a priori* hypothesis, we are interested in extracting major patterns of variation in intralimb coordination in the proximal forelimb of ungulates. We discuss evidence for deviations from dynamic similarity, and possible reasons. This provides hypotheses for future, more controlled experimental tests that can likewise benefit from the herein proposed analytical methodology.

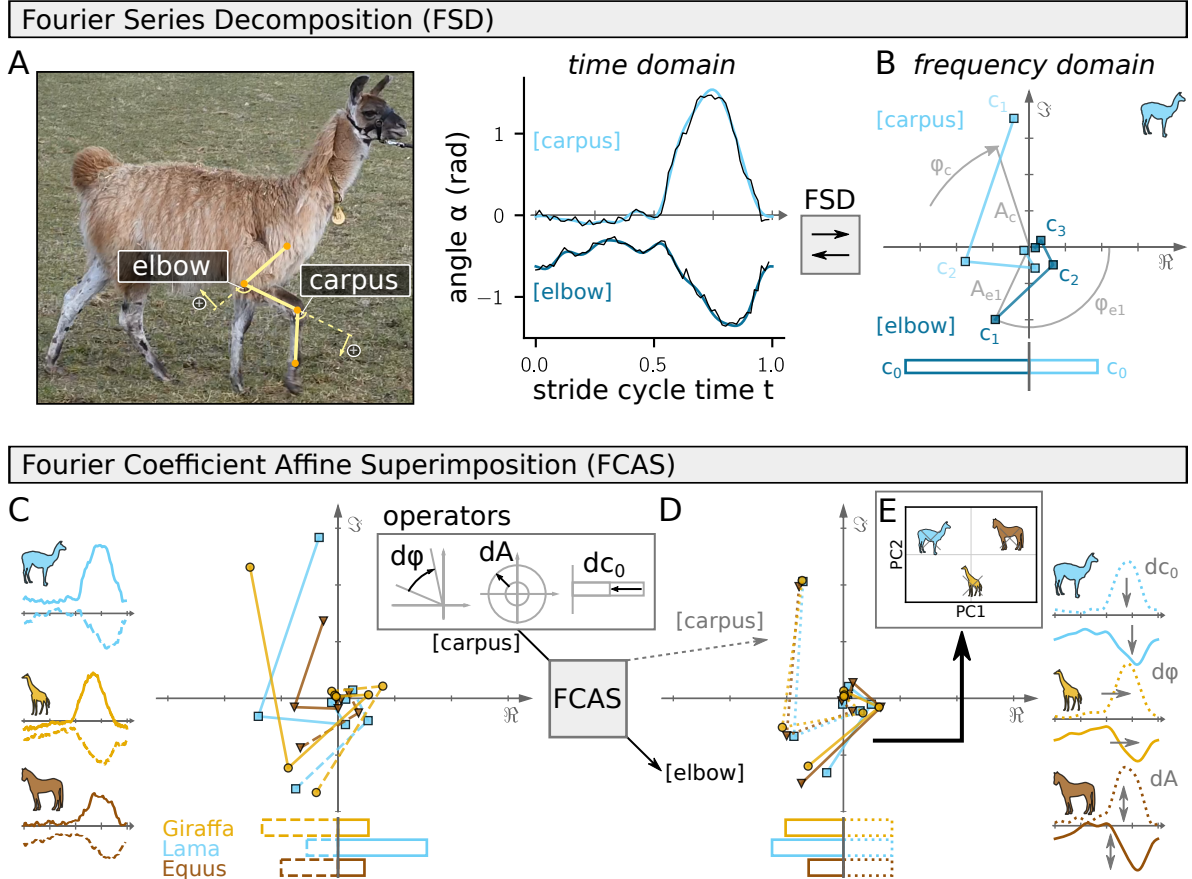


Figure 3.1: Analysis work flow, schematic. **A** Video material of walking ungulates is digitized by tracking trace marks (shoulder, elbow, carpus and fetlock joints; see supplementary movie 3.7.5) over a stride cycle, and joint angle profiles are computed (joint angle defined zero at fully extended limb, positive direction indicated by arrows). Thin black lines are the original measurement, whereas blue lines show the same profile, filtered by FSD with 7 components. (*continues next page*)

Figure 3.1: (*contd.*) **B** Fourier Series decomposition and synthesis allow transformation of temporally periodic joint angle profiles back and forth between time domain and frequency domain. The zero'th Fourier coefficient (c_0 , lower bar plot) corresponds to the temporal average of the joint angle profile (i.e. mean joint angle). The higher coefficients ($c_{n>0}$) can be plotted in the complex plane ($\Re, \Im$); line connections are only for visualization. Coefficients i of joints j have phases (ϕ_{ji}) and amplitudes (A_{ji}), as indicated in gray for c_1 of the elbow. Average phase and amplitude of the joint angle profile can be derived (for example ϕ_c , A_c for carpus). Mean joint angle, amplitude, and phase are the affine components of the signal.

C Multiple groups are compared, in this case ungulate genera (*Lama*, *Giraffa*, and *Equus*; coloured in blue, yellow, and brown respectively). By quantifying how they differ in affine components, operators can be extracted ($d\phi$, dA , dc_0). This is facilitated by the transformation to the frequency domain, but a synthesis of time domain traces (line plots on the left and right margin) is always possible.

D Affine operators are applied to superimpose the joint angle profiles. Superimposition operations from the carpus (reference joint) are applied to the carpus itself, but also transferred to the elbow (focal joint).

E (inset) Fourier coefficients of carpus-superimposed elbow joint profiles are available for multivariate analyses, such as PCA.

3.3 Materials and Methods

Data Acquisition

The data for this study was acquired from an online video sharing platform ([youtube.com](https://www.youtube.com)). Genera were searched in Latin, English, or various local languages and search results were scanned for episodes of walking in approximately lateral perspective that did not show signs of *post hoc* manipulation (cutting, etc.). Slow motion recordings were included if the frame rate was available in the video description. To exclude confounding variability due to the non-standardized recording circumstances, i.e. to get controlled reference measurements that ensure the validity of the online data, we supplemented high quality recordings of some genera available at our institute or local zoological gardens (34 stride cycles from 11 videos; *Bos*, *Equus*, *Lama*, *Tapirus*). In total, 388 videos contained candidate episodes and were annotated and cut (i.e. full stride cycles) in ELAN software (Max Planck Institute for Psycholinguistics, Nijmegen; Brugman and Russel, 2004). Episodes were selected to have sufficient temporal and spatial quality and correct perspective (movement perpendicular to the view axis). Videos of juvenile individuals were rarely included, i.e. only in cases where few recordings were available for a genus. Video cutting was performed in the free video manipulation software “ffmpeg” (<http://ffmpeg.org>).

Tracking of the animal movement (Fig. 3.1A) was done in a custom-written tool (Mielke et al., 2020) using the Open Source Computer Vision Library (“OpenCV”, <https://opencv.org>) interface for the Python programming language (Python Software Foundation, <https://www.python.org>). We manually tracked joint center pixel positions over time, tracking four joints along the forelimb which was ipsilateral to the camera. From this data, joint angles were calculated, various quality checks applied, and it was ensured that the joint angle profiles (i.e. joint angles over time) were cyclical (for a detailed procedure, see supplement 3.7.1). The videos with episodes that passed quality criteria were reviewed to register species, sex, speed, duty factor, and degree of obstruction. The final material covers 87 of 117 genera (109 of 251 species) of terrestrial ungulates. 556 stride cycles passed strict quality checks, stemming from 217 of the videos (supplementary data table 3.7.5).

Fourier Series Decomposition (FSD) Exposes Affine Signal Components

We herein study periodic joint angle profiles, but with methods that generally apply to periodic signals. There are two fundamental ways in which two or multiple signal observations can differ. One type of differences are those which can be changed by a linear operation (i.e. addition or multiplication of scalars). Those are the mean, amplitude, and phase, which can be standardized by “y-shifting”, scaling around the mean, and phase-shifting, respectively. These differences are termed affine differences. Other affine operations exist (reflection, shear, planar rotation), but they do not apply to time profiles

because they disrupt the spatiotemporal integrity of the signal. Differences that cannot be standardized by affine transformations are subsumed as non-affine.

In the time domain (plot of the signal value over time, Fig. 3.1A), the affine differences are intuitively visible, but hard to capture mathematically due to temporal periodicity. Hence, Fourier Series decomposition (FSD, as defined in supplement 3.7.2) is applied to the signals we study herein. This transformation decomposes a signal into a finite sum of its harmonics. These harmonics are wave functions for which the product of their signal period T (i.e. full cycle recording duration) and frequency f is an integer (e.g. $T = 2.1$ s, $f = 3/2.1$ Hz, which is the $n = 3^{\text{rd}}$ harmonic). The harmonics are the Fourier coefficients, but instead of the commonly used sine- and cosine representation (a_n, b_n) , we extract complex exponentials (c_n) ; both representations are related through Euler’s formula). Hence, the coefficients herein are complex numbers which, with their real and imaginary part, constitute the frequency domain representation of the signal (Fig. 3.1B).

Complex Fourier coefficients facilitate the extraction of information about affine components of the signal (see supplement 3.7.2) because rotational operations are conveniently applied with complex exponential factors. Affine differences translate to the coordinate positions of the coefficients in the complex plane: the zero’t h coefficient c_0 reflects the temporal mean, whereas polar coordinates of the other coefficients (distance from origin, rotation relative to positive real coordinate axis) map the amplitude and phase of the signal (Fig. 3.1B). The transformation has an interesting effect on the temporal sampling. Take, for example, a signal composed of 20 measurement samples over the cycle. To phase shift, one could “roll” the vector of samples, e.g. appending the first sample at the end. Intermediate phase-shifts are inaccurate because they require interpolation. In the frequency domain, the signal becomes independent of temporal samples (*cf.* supplement 3.7.2), except that sampling gives an upper limit to the number of harmonics which can be extracted. In consequence, phase information becomes continuous, phase shifting is not restricted to the temporal raster defined by sampling, but no interpolation is required. Conversely, series decomposition with a finite number of harmonics leads to low pass filtering, hence FSD removes high frequency components from the original signal.

An overview of how affine changes affect the signal in time and frequency domain can be found in the online supplements, together with an implementation of FSD (supplementary data 3.7.5). It is important to note that all of FSD and the procedures described herein are deterministic. Hence, none of these require a fitting- or optimization algorithm. In previous studies, it has been quite common to determine Fourier coefficients by approximation (e.g. Alexander and Jayes, 1980) or an iterative algorithm (regression, e.g. Hubel and Usherwood, 2015). A deterministic implementation is more efficient and less error prone and should thus be strictly preferred (the space that an optimizer has to traverse is complex due to angular ambiguity; there are inherent phase relationships of coefficients that common samplers ignore). We also recommend to reconstruct test signals

from coefficient values for visual comparison with the original signal to confirm that the decomposition was accurate.

Fourier Coefficient Affine Superimposition (FCAS) Reveals Signal Shape

Taking all coefficients derived from the FSD of a signal together is data that is structurally analogous to two-dimensional landmarks. Hence, methods of geometric morphometrics can be applied (Bookstein, 1991; Dryden and Mardia, 2016; Gower, 1975; Kendall, 1989), in particular Procrustes superimposition, a method which standardizes affine components of the form of an object to reveal its shape. For the complex Fourier coefficients, it is possible to define standard values for the affine components (i.e. mean of zero, amplitude of one, phase zero) and operators that shift or scale the signal to achieve standardization. Similarly, it is possible to find, extract, and apply operators that superimpose two signals by standardizing affine components (Fig. 3.1C), which we term FCAS. In contrast to geometric morphometrics, the rotation is not uniform, but depends on the harmonic number. Also, the translational component is restricted to c_0 , and all higher coefficients are only affected by rotation and scaling. These modifications do not hinder the general procedure. Analogous to the process of Generalized Procrustes Analysis in geometric morphometrics, the superimposition can be applied to find improved averages of signals. The (non-affine) residual after isolation of affine differences is, in analogy to geometric morphometrics, the “shape” of the signal, and it is possible to define distance metrics to quantify shape differences (though not relevant for the present study).

In case kinematic measures, such as joint angle profiles, are the signals of interest, both affine and non-affine differences can have consequences for the kinetics. However, only non-affine differences imply different coordination, whereas affine components are merely a change in temporal or spatial scale of the same patterns. Coordination should be evaluated in context of the limb. Hence, when regarding multiple joints along the limb, it is crucial to retain the coupling of the limb joints. For example, when phase aligning carpus joint profiles (reference joint, for reasons to be discussed), the same time shift should be applied to other joints of the limb (focal joints, herein the elbow, Fig. 3.1D). Similarly, a normalization of the mean and amplitude of the reference joint can be applied to focal joints, which then express relative mean and amplitude. When applying operators from a carpus joint superimposition to the elbow joint profiles, we term this “carpal superimposition of the elbow”, and retrieve “relative elbow profiles”. This transfer shifts the affine variance from the carpus to the elbow joint. For example, if in species A the carpus phase is lagging +0.05 stride cycles relative to species B, but the elbow phase difference of the two species is -0.1 (A before B), then the phase of the relative elbow profile of species A will be -0.05 with respect to that of B. Hence, relative profiles capture intralimb coordination. They are also represented in the frequency domain, but temporal profiles can be reconstructed at any point for visualization (margins of Fig. 3.1C and D).

The Fourier coefficients of relative elbow profiles are subjected to further analysis. Because we are interested in dynamic temporal changes relative to the mean joint angles, we isolate differences in c_0 by subtracting it from the profiles, despite acknowledging that posture or average joint angle might affect the dynamics.

Multivariate Analysis

A variety of multivariate methods exist that could be explored with relative elbow profile Fourier coefficients. Because our study is introducing a previously unexplored method (FCAS), we restricted post processing to plain Principal Component Analysis (PCA, Fig. 3.1E). PCA was applied to the real and imaginary parts of relative elbow FSD coefficients. By an eigenanalysis of the covariance matrix of the coefficient table, PCA finds orthonormal axes which orient the data in a way that shows the major variation (*cf.* MacLeod, 2007). This was manually implemented in Python, using singular value decomposition (`numpy.cov`, `scipy.linalg.eig`). We apply it to genus averages (Mitteroecker and Bookstein, 2011) of FSD coefficients of relative elbow profile (i.e. after FCAS with regard to carpus).

The component axes are not *per se* biologically meaningful, but can be associated with covariates. Covariate choice is arbitrary, but depends on the question of interest. In our case, four classes of covariates which potentially have a relevant connection to intralimb coordination were collected. In addition to that, we quantified the affine composition of the relative profiles. The covariate classes are morphology, spatiotemporal gait parameters, ecology, and phylogeny.

First, characteristics of general and musculoskeletal morphology determine locomotor function. Species- and sex-specific average measures of body length, shoulder height, and body mass were taken from online data repositories (Animal Diversity Web, 2019; Huffman, 2019). For measures on the video, segment lengths were taken from pixel distances of the tracked points, but normalized as indicated below to remove differences in video resolution. To evaluate lever relations at the critical joints, zeugopodial length (elbow-wrist, normalized to withers-croup distance) and the meta- to zeugopodial length ratio were included. Under ideal circumstances, segments should measure equal length in every frame. However, there is uncertainty in the landmark tracking procedure. In addition, in cases where the fetlock joint was covered by vegetation during ground contact (but not during the swing phase), tracking points were shifted proximally along the limb segment to still be able to capture the (carpus) joint angle over time. The framewise 90% quantile value turned out to be robust to episodes when joints were hidden during stance, while being sufficiently close to the true mean segment length (i.e. robust to inadvertent tracking outliers).

Second, intralimb coordination potentially covaries with spatiotemporal gait parameters (kinematics). From the video data, duty factor (ratio of stance- and stride time) and

approximate speed (body lengths per second, from back line to reference landmark movement) could be measured. The exact Froude number was unavailable, since the spatial scale of the videos was unknown. Furthermore, a proxy for clearance (*cf.* Austin et al., 1999) was calculated, i.e. how much the hoof is lifted above the ground during a limb swing. This required the effective limb extension (ξ), defined as follows, from segment (i) lengths (l_i) and their proximal joint angles (α_i , Fig. 3.1):

$$\xi = \sum_i l_i \cdot \cos \alpha_i$$

We define clearance as the quotient of the 5% and 95% quantiles of ξ over time (which quantifies limb flexion ratio), subtracted from one.

$$\text{clearance} := 1 - \frac{\xi_{5\%}}{\xi_{95\%}}$$

Again, quantile values were used because the extremal values would be at the far tail of the Gaussian noise associated with point tracking.

The third group of covariates are ecological parameters. In particular, we were interested in the species habitat and the momentary environment at video recording. We followed the method of Stankowich and Caro (2009) to quantify habitat openness (Caro et al., 2004; Stankowich and Caro, 2009). Web resources indicated multiple habitat preferences for most species, in which case an uninformed average of the openness values was taken. The momentary openness, or its inverse (“degree of obstruction”), was determined by noting how far substrate clutter reached up the supporting limbs of the animal in each video. Values ranged from 0 (hoof visible) to 0.9 (limb visible to just below the carpus joint, see supplementary data table 3.7.5).

The fourth covariate was phylogeny (see supplement 3.7.3 and Fig. 3.4; Adams, 2014; Zurano et al., 2019).

Except for phylogeny (which is directly estimated), associations of covariate genus averages and principal component scores of genus means (Mitteroecker and Bookstein, 2011) were assessed by calculating Pearson correlation (Python: `scipy.stats.pearsonr`). Within-group variability was high and observation count heterogeneous across groups, and we applied many correlation tests. Hence, we only deemed associations with $p < 10^{-2}$ relevant.

It can be instructive to quantify how much each of the affine differences contributes to the overall relative elbow trace difference. Therefore, the affine operations that would align the genera averages with their global average were computed. We report the correlation of these affine differences to the PC axes, as well as the residual shape difference that would remain after their isolation.

3.4 Results

3.4.1 Superimposition of Joint Angle Profiles Improves Profile Averaging

To demonstrate the effect of FCAS, we applied the method to our most extensively sampled genera (*Equus*, *Lama*, and *Giraffa*; Fig. 3.2). Raw data (“video aligned”) was manually aligned for the video frame of limb touch down for comparison with the computational alignment. Although this was done at maximum possible accuracy (± 1 frame), averages revealed no clear similarities or differences between the three genera. Standard deviation ranges were comparatively high and averages could potentially suffer from incorrect phase relations (analogous to destructive interference) or artificial amplitude weight (because at a fixed time point, particularly high absolute values might attract the average). FCAS was subsequently performed with regard to carpus joint profiles (Fig. 3.2, “superimposed”, “[c]” for carpus joint). Afterwards, carpus profiles were practically identical for the three example genera. Elbow joint profiles (“[e]” in Fig. 3.2) were kept in synchronization and amplitude relation with the carpus profiles and were therefore processed with the same operators, resulting in relative profiles (see methods). Note that no additional superimposition or alignment was performed, except for centering the profiles around their mean. Relative elbow profiles revealed a precise similarity of lamas and giraffes with regard to coordination of the two joints, which was not obvious from the “video aligned” traces. In contrast, equid intralimb coordination indicated a higher amplitude and a slight phase delay at the elbow (triangle marker, Fig. 3.2), relative to carpus joint. Thus, equids turn out to be more distinct from the other groups than was visible prior to FCAS.

This example indicates the increased accuracy of joint profile averaging after FCAS, even in the presence of noise.

3.4.2 Patterns of Variation of Ungulate Intralimb Coordination

We aim to extract the most prevalent patterns of variation in ungulate intralimb coordination. Therefore, we used the same superimposition procedure as above to generate genus averages for the full data set. Average joint angle profiles of all joints and genera (*cf.* Fig. 3.4 for overview of phylogeny) were then superimposed with reference to the global mean carpus (carpal superimposition). Thereafter, a PCA was performed on the Fourier coefficients of genus average relative elbow profiles (Fig. 3.3). The first component axis captured 37% of the variance, and the next three principal components further 49% (PC2: 27%, PC3: 12%, PC4: 10%). Together, they span the ungulate elbow “coordination space”. In that space, bovids and cervids, which contributed the majority of genera, dominate the central region. On PC1, we observed a spread of bovids from Bovinae to Caprinae. For cervids, who spanned almost the same range on that axis, there was no clear grouping.

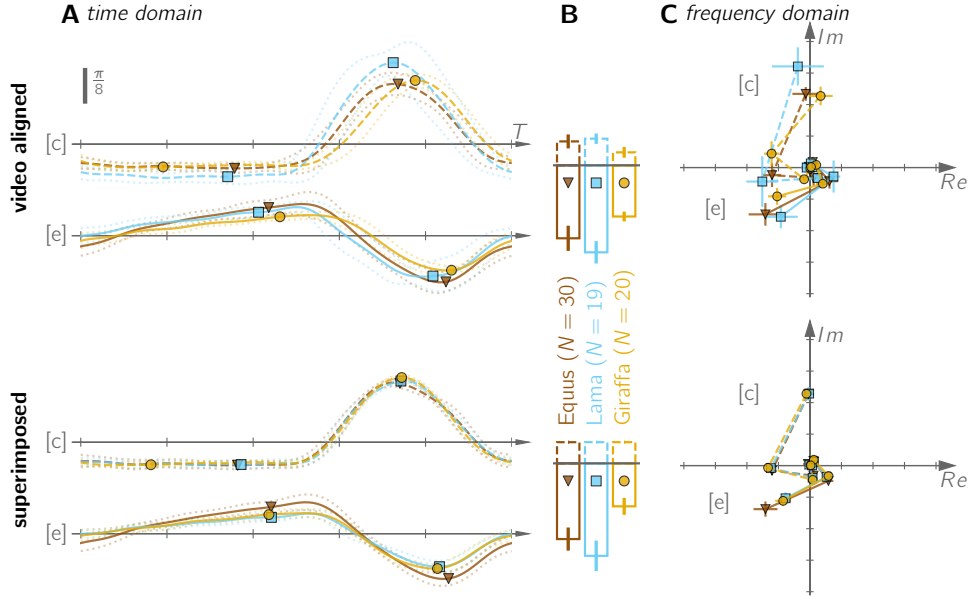


Figure 3.2: Superimposition can improve joint angle profile averaging. Mean joint angle profiles of different forelimb joints ([e]: elbow, [c]: carpus joint) for the genera *Equus* (brown, ∇), *Lama* (blue, $\square$), and *Giraffa* (yellow, $\circ$). These examples were chosen for their high sample size (more examples in Fig. 3.5). See methods and supplementary tutorial 3.7.5 for nomenclature and definition of FCAS. **Upper row:** input data (“video aligned”). **Lower row:** after carpal superimposition using FCAS. **A** Forelimb joint angle profiles (α) over time (t , stride cycle period T), centered around their temporal averages (gray horizontal axes) to emphasize temporal differences. Thick lines are genus averages for elbow (solid line) and carpus (dashed line) joint angles. Thin dotted lines indicate $\pm$ standard deviation range of individual observations. Scale bar 22.5° , time axis T . **B** Mean (bars) and standard deviation (whiskers) of the joint angle temporal averages (c_0) per genus. This indicates the actual y -offset that was removed for visualization in (A) and (C). Joint angle values are to scale with (A). **C** The same joint angle profiles as in (A), but converted to the frequency domain via FSD. Whiskers indicate standard deviation.

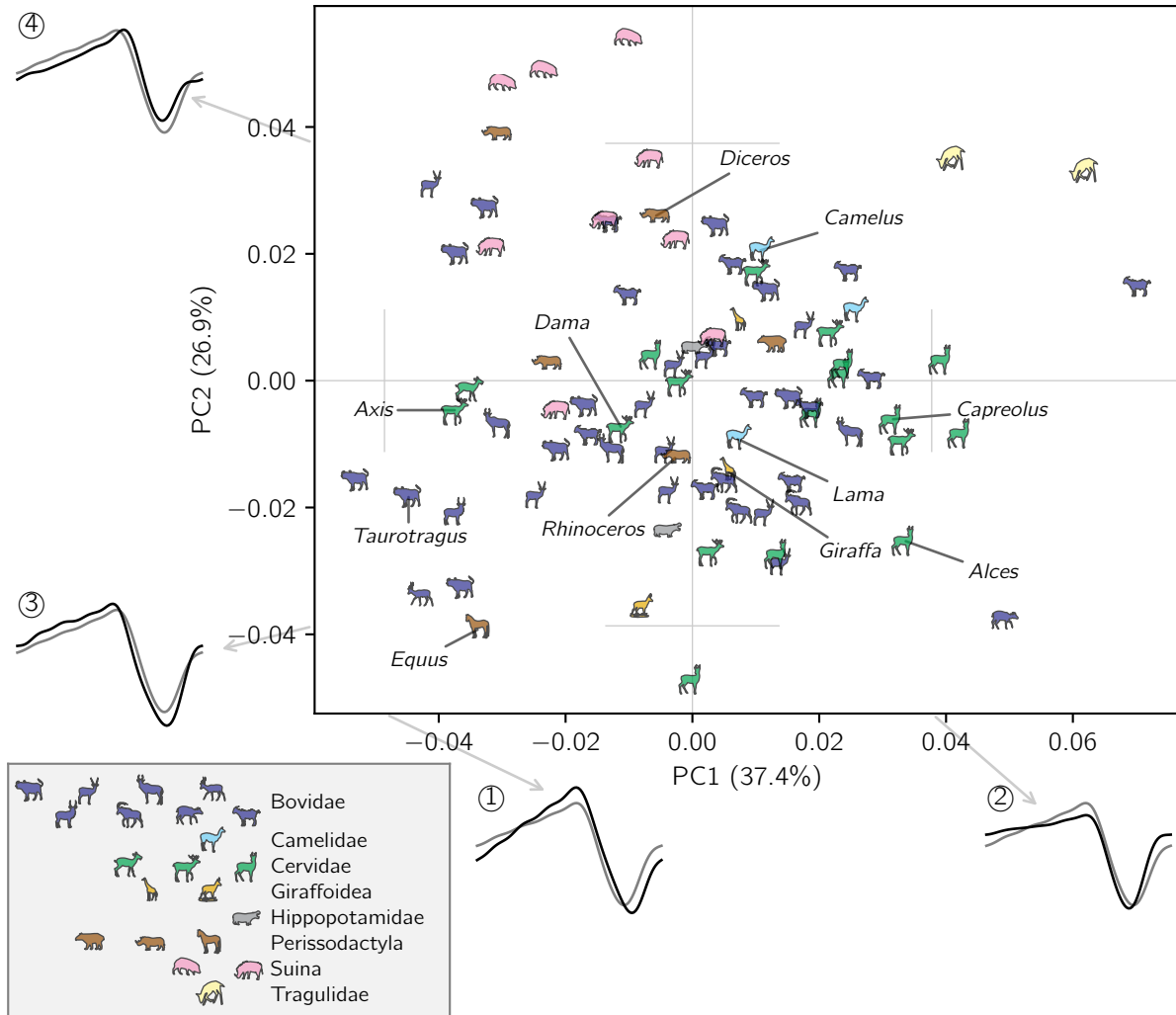


Figure 3.3: Bivariate plot of the Principal Component Analysis of elbow-carpus coordination visualizes a slice of the ungulate “coordination space”. The PCA is based on the genus average Fourier coefficients of relative elbow profiles after carpal superimposition. Gray grid lines on the biplot indicate the (0.1, 0.9) quantiles of all observations; reconstructions were done on their intersections. Outset line plots on the margins represent these reconstructed angular profiles (black) at the quantile values for each component axis, relative to the reconstructed mean (gray). Their labels (circled numbers no. 1-4) are for reference in the text. Annotations on the plot mark the examples used in Figures 3.2 and 3.5. See Fig. 3.6 for further components and complete labeling of taxa.

Tragulids were found on higher values of PC1. A useful aspect of PCA is the possibility to reconstruct hypothetical joint angle profiles at arbitrary points of coordination space (black line plots at the lower figure margin, Fig. 3.3, no. 1-4). From these, it can be seen that the major difference is how much the relative joint angle changes during ground contact. Ground contact (stance phase) begins with limb touch down (left end of the line plot), continues for approximately the first half of the line, and is followed by a shorter swing phase. At low values of PC1, the relative elbow profiles (no. 1) are “saw-tooth-like” curves with steeply sloped, increasing joint angles during stance phase. These profiles are characterized by two periods of relatively constant slope, one rising (elbow extension, during stance) and one falling (flexion, during swing). The slope of these curves changes in between these periods, which approximately coincides with hoof touch down and lift off. Extension of the elbow joint happens during ground contact, hence we call this type of profile “grounded elbow extension”. In contrast to that, elbow joint angles that cluster on high PC1 values (no. 2) only change during the swing phase (brief period of consecutive down- and upward slope), whereas they are close to zero during stance, indicating a straightened joint. Elbow extension is almost absent during stance (approximately horizontal curve), and after swing-related flexion, an extension follows immediately while still in the swing phase (“swinging extension” type profile).

PC2 separated a cluster of Rhinocerotidae and Suina (upper left) from the rest. Their relative elbow swings (represented by no. 4) were of lower relative amplitude and slightly phase delayed than those of groups on the opposite end of PC2 (reconstruction no. 3).

Standard deviations within groups were high (not shown; approx. ± 0.02 for *Giraffa* on PC1) and sample size was low for many groups. However, overall group and observation counts were sufficient to describe population effects. In addition to the shape reconstruction, the direct comparison of geometrically or phylogenetically related genera along the major axes could confirm the mentioned trends (i.e. stance phase slope, relative peak timing and amplitude; Fig. 3.5). This further supported the observations from PCA reconstruction.

In summary, the PCA, profile reconstruction, and taxa comparison allow the identification and visualization of major differences in intralimb coordination among the studied groups.

3.4.3 Covariates of Intralimb Coordination

To evaluate possible explanations for the variation along principal component axes, we computed Pearson correlation of the components with various morphological, kinematic, and ecological factors, considering effects with $p < 0.01$ to be relevant (Tab. 3.1). We also split the joint profile shape after superimposition into (non-)affine components, and quantified which of these affine changes were associated with the PCs or whether the shape (“residual”, i.e. coordination *sensu stricto*) was affected (Tab. 3.1, “[A]” category). These

values quantitatively characterized the profile reconstructions from the PCA (Fig. 3.3) and were consistent with the findings reported above.

Of the first four component axes, PC1 showed the strongest correlation with clearance ($r = 0.49$, $p < 10^{-5}$) and degree of obstruction ($r = 0.33$, $p < 10^{-2}$). Comparatively smaller taxa (i.e. lower body mass) tended to appear on higher PC1 values ($r = -0.34$, $p < 10^{-2}$). PC2 was associated with limb morphology, particularly lever relations (ratio of meta- to zeugopodial length, $r = -0.51$, $p < 10^{-6}$; normalized elbow-wrist segment length, $r = -0.46$, $p < 10^{-4}$), and with duty factor ($r = 0.4$, $p < 10^{-3}$). PC3 scales with habitat and multiple aspects of size, whereas PC4 captured non-affine shape differences (visible in that it correlated to “residual”).

Hence, by the use of simple correlation, it is possible to associate the variational components identified by PCA with meaningful extrinsic (e.g. ecology) and intrinsic (e.g. morphology) covariates.

Table 3.1: Correlations of [M]orphological, [K]inematic, [E]cological and [A]ffine covariates (see methods) with the major axes of the PCA (Figs. 3.3, 3.6). Note that affine covariates are components of the joint angle profiles.

covariate	<u>PC1</u>		<u>PC2</u>		<u>PC3</u>		<u>PC4</u>	
	r_{xy}	$p < \cdot$	r_{xy}	$p < \cdot$	r_{xy}	$p < \cdot$	r_{xy}	$p < \cdot$
[M] body length	-0.24	10^{-1}	-0.17	10^0	-0.31	10^{-2}	-0.11	10^0
[M] shoulder height	-0.19	10^{-1}	-0.22	10^{-1}	-0.33	10^{-2}	-0.11	10^0
[M] log(body mass)	-0.34	10^{-2}	-0.19	10^{-1}	-0.28	10^{-2}	-0.11	10^0
[M] elbow-wrist	0.05	10^0	-0.46	10^{-4}	-0.24	10^{-1}	0.05	10^0
[M] meta-/zeugopodial	0.08	10^0	-0.51	10^{-6}	0.01	10^0	0.16	10^0
[K] approx. speed	0.06	10^0	0.11	10^0	0.12	10^0	0.15	10^0
[K] duty factor	0.16	10^0	0.40	10^{-3}	0.19	10^{-1}	-0.16	10^0
[K] clearance	0.49	10^{-5}	-0.02	10^0	0.14	10^0	-0.27	10^{-1}
[E] habitat	-0.09	10^0	-0.15	10^0	-0.29	10^{-2}	-0.06	10^0
[E] d.o.obstruction	0.33	10^{-2}	-0.20	10^{-1}	0.10	10^0	-0.06	10^0
[A] amplitude	0.12	10^0	0.47	10^{-5}	-0.70	10^{-12}	-0.08	10^0
[A] phase	0.47	10^{-5}	-0.51	10^{-6}	-0.36	10^{-3}	0.13	10^0
[A] residual	0.67	10^{-11}	0.17	10^0	0.15	10^0	0.36	10^{-3}

r_{xy} : Pearson correlation coefficient, “ $p < \cdot$ ”: p-value, order of magnitude. Bold numbers are below significance threshold ($p < 0.01$).

3.5 Discussion

In this study, we demonstrated how an application of Fourier theory can be used to extract non-affine differences in joint angle profiles, thereby enabling their mathematically accurate superimposition (FCAS). The new method has to be classified with relation to existing tools. As an example of FCAS, we analyze intralimb coordination on a large scale, involving the majority of terrestrial ungulate genera, which has not been possible previously. The kinematic data from the proximal forelimb of ungulates reveals systematic patterns of variation in intralimb coordination. These can be linked to extrinsic and intrinsic covariates, thereby proposing hypotheses for specific tests of dynamic similarity at the joint level.

3.5.1 Quantifying Intralimb Coordination with FCAS

Conventional kinematic analysis of intralimb coordination involves visual inspection of joint angle plots in space and time (Day and Jayne, 2007; Fischer et al., 2002; Irschick and Jayne, 1999; Nyakatura et al., 2010; Polk, 2002; Schmidt, 2008; Stoessel and Fischer, 2012). Joint angle values for statistical evaluation are extracted at fixed “timemarks” in the stride cycle. Taken together, these sampled joint angles allow for coarse assessment of intralimb coordination, but they potentially sacrifice information. Timemarks are reference points of putative biological meaning, and it is possible to use them for temporal alignment (e.g. Hsiao-Wecksler et al., 2010). However, the number and choice of timemarks involves some critical design decisions. When for example superimposing only based on limb touch down, the alignment is insensitive to differences in hoof placement trajectories before and after that time point. The hoof can reach the ground at a flat impact angle with smooth horizontal deceleration. But the impact could instead be vertical, preceded by a more or less brief period where the hoof hovers above the substrate before being placed. The end of hoof forward movement might provide a more homologous timemark, but it might still be that the two cases differ fundamentally in the underlying dynamics at that particular episode of the stride cycle. FCAS represents a mathematically optimal superimposition that equally incorporates all the information contained in the profiles. The advantage of this strategy is its independence of selected points, and instead the sensitivity to the joint angle trajectory over the whole stride cycle. Working in the frequency domain also has the advantage of providing a manageable feature vector for multivariate analysis or classification via machine learning. Fourier Series requires periodic signals to work upon. As we argued initially, this requirement is often met in the study of locomotion, in which case a transformation to the frequency domain opens up new ways of analyzing the data. The direct application of Fourier Series (Bracewell, 2000; Fourier, 1822; Gray and Goodman, 1995) is not novel in the context of kinematics (e.g. Bernstein, 1935; Pike and Alexander, 2002; Webb and Sparrow, 2007; Wheat and Glazier, 2006). Note that Fourier

Series is related to Fourier Transform (cf. Bracewell, 2000; Robertson et al., 2018): it is also a transformation, but considers only the reduced frequency domain constituted by harmonic frequencies. Because the interrelation of segments is of interest, Fourier methods have previously been applied to cyclogram curves. Cyclograms, or phase portraits, are a direct, two- or three-dimensional visualization of segmental coupling (Bernstein, 1934; D’Août et al., 2002; Goswami, 1998). Several quantitative descriptors of cyclograms exists. Elliptic Fourier Descriptors (Kuhl and Giardina, 1982; Wheat and Glazier, 2006) are an application of Fourier Series to two-dimensional phase portraits. This technique is abundant in the study of intralimb coordination (e.g. Hsiao-Wecksler et al., 2010; Polk et al., 2008; Rosengren et al., 2009), but influenced by variance in reference point determination. Although affine-invariant outline descriptors exist (e.g. Arbter et al., 1990), we are not aware of an attempt to apply those to kinematic data. Planar covariation (Borghese et al., 1996; Hallemans and Aerts, 2009; Ogihara et al., 2014) originated in the pursuit of unifying, general features of gaits, and finds high covariation values by design (choice of segment angles, common temporal swing/stance structure for these, use of PCA; Hicheur et al., 2006; Ivanenko et al., 2008). In contrast, we seek a method to highlight differences across taxa. Another method that has found application in the analysis of human locomotion is continuous relative phase (CRP, cf. Lamb and Stöckl, 2014). This method calculates the difference in instantaneous phases of two signals (using phase portrait plots or the Hilbert transform). CRP requires removal of differences in mean joint angle and amplitude (i.e. normalization), whereas our method retains the other affine components. By intentionally phase-shifting random test signals (not shown), we found that both the relative phase from CRP (averaged over a stride cycle) and the FCAS phase difference as introduced herein (supplement 3.7.2, Eqn. 3.6) accurately recover the shift. However, the methods do not yield identical results when other aspects of the signal are manipulated (which in general complicates the definition of a signal phase). It is thus possible that Hilbert transform and CRP offer alternatives for phase determination. Because ungulate joint angle profiles are sufficiently homogeneous, we expect general agreement between the methods for the test case presented herein.

All the mentioned related methods have originally been applied to phase plots and cyclograms. Our approach initially goes one step back, to the transformation of individual time profiles of joint angles. FSD preserves their shape information, and no *a priori* definition of temporal reference points is required (we herein performed the conventional limb touch down alignment, prior to FCAS, only for comparison). To generate measures of coordination, one has to combine variance from multiple mobile elements (joints or segments) by aligning with regard to one of them while keeping the intralimb coupling intact. We exploit the empirically observed uniformity of carpus joint profiles, and use this joint as reference, which is motivated by mathematical convenience. Although superimposition based on the elbow joint might be biologically more meaningful (for

example with regard to a proximo-distal sequence of neural control), it would be a less suitable reference because profiles might be topologically heterogeneous. Transferring the affine components from references to adjacent, focal joints results in relative joint angle profiles, which are optimally synchronized (phase) and depict the relative amplitude (scale) and differential mean joint angle. If mean joint angles are not of interest, profiles can optionally be centered. We argue that this way of superimposition results in measures that overcome the theoretical limitations in transferability of angular parameters (Gatesy and Pollard, 2011): (i) posture can be removed by centering relative profiles; (ii) adjacent joint angles are combined; (iii) morphology, i.e. segment lengths, can be correlated *ex post* as shown, or could even be multiplied in (effective segment length $l \cdot \cos(\alpha)$; not explored). FCAS crucially depends on reference and focal joints, but is not limited in the choice (or number) of these. Also, it is not limited to a single limb or to phylogenetic comparisons (for example, pseudo-cyclic episodes during gait transition could be compared) and it could be generalized to three-dimensional measurements. FCAS results in a mathematically optimal alignment, and is robust against low video frame rate due to the filtering properties of the Fourier Series with a finite number of coefficients. The capacity of a one-dimensional FSD to standardize affine components in the frequency domain has not been emphasized in previous work. The delay theorem, from which we derive a formula that defines a zero phase (see supplement 3.7.2, Eqn. 3.6), is of particular interest. FCAS is not limited to kinematics, and could potentially be applied to a broad variety of other periodic measurements, such as ground reaction force profiles.

3.5.2 Dynamic Dissimilarities in Terrestrial Ungulate Intralimb Coordination

The purpose of the PCA step in the present analysis is to orient the multivariate data so that maximum variation is revealed. Variation in this case means dissimilarity. It is herein applied to relative elbow profiles. These combine information from carpus and elbow, but the carpus only contributes affine changes (Tab. 3.1, category [A], “amplitude” and “phase”). Non-affine variation (Tab. 3.1, category [A], “residual”) will stem exclusively from the elbow. Hence, relative elbow profiles in our case show elbow profiles that are centered, normalized and time-aligned relative to the carpus profiles. The change of one joint angle relative to another is the most basic measure of intralimb coordination. From these measures, PCA extracts the “coordination space” of the relative elbow joint profiles (Fig. 3.3). The method also allows to reconstruct hypothetical profiles at arbitrary points in that space. Overall, the group differences might seem subtle at visual inspection, attributing to the fact that “Bauplan” and movement mode are homogeneous in the data set. However, even subtle changes in the coordination of the proximal limb have a considerable lever towards the ground contact point. Also, the reconstructed profiles are

derived quantitatively, incorporating a lot of information. Hence, they are a sharp image of the most pronounced statistical differences that remain after the standardization of affine parts of background variance. Likewise, dynamic similarity (as usually applied to spatiotemporal gait parameters) refers to similarity in non-affine dynamic measures. This conceptual link indicates that FCAS enables the direct comparison of dynamic differences in the coordination of large data sets of animals. For the present measurements of walking ungulates, PCA of FSD coefficients of relative elbow joint angle profiles revealed variation associated with several factors.

The most prominent effect revealed by the PCA (PC1, 37% of variance) was related to clearance. Clearance quantifies how much the animal lifts the foot from the ground during swing phase (Austin et al., 1999; MacLellan and McFadyen, 2010; Perrot et al., 2011). The component axis was also correlated with animal body mass and momentary degree of obstruction (i.e. how free the path of the animal is), but independent of morphology. Classical gait characteristics (duty factor, symmetry, i.e. inter-limb coordination) doubtlessly classify all included strides as “walk” (Hildebrand, 1989), but do not consider intralimb coordination or clearance. Because all observations are “walk”, and because no change in duty factor was detected (Tab. 3.1), the timing of stride- and stance phase is uniform along PC1. However, the temporal pattern of the net joint moments at a limb joint (i.e. the net result of the moments generated by muscles crossing the joint) will likely differ among the profile types as the movements patterns throughout the cycle differ. For example, the ungulate elbow in the swinging extension type of relative joint profiles (Fig. 3.3, no. 2) is commonly held stiff during the period of ground contact (constant joint angle value), but quickly flexed and extended during forwards swing. It thereby resembles the common pattern of the ungulate carpus joint (see Fig. 3.2). In contrast, the elbow joint oscillation of grounded extension type profiles is distributed across the whole stride cycle. This suggests either that the elbow can be a pivot point of limb movement, in which case more proximal elements could be fixed passively, or that the elbow contributes to limb compliance during stance to minimize vertical oscillation of the center of mass (*cf.* Geyer et al., 2006). Animals commonly use the grounded extension profiles on obstruction-free substrate, whereas the swinging extension could be interpreted as an alternative mode of walking that animals use on cluttered terrain.

The elbow profile modes have different implications for the energetic demands of the locomotor behaviour. Energetics are affected by the substrate that an animal frequently encounters. Grounded elbow extension implies non-zero angular velocity at the elbow during all of the stride cycle. If, as discussed above, more proximal elements are mostly fixed, this would lead to an effective shortening of the part of the limb which is mobile. This is consistent with evidence of stabilizing modification of the shoulder of equids (Hermanson and MacFadden, 1992). Conversely, muscle arrangement in tapirs, which usually exert higher clearance than horses, shows adjustments that favour shoulder motion

(MacLaren and Nauwelaerts, 2016). The effective limb shortening, together with the passive pendulum dynamics of the distal elements on unobstructed ground, would reduce energy expenditure. In contrast, the brief flexion-extension period in swinging flexion appears to require more energy per stride (quicker angular acceleration in opposing directions, longer effective limb, higher clearance). At the same time, because collision with substrate features is avoided, this behaviour reduces collisional energy loss (i.e. loss from collision with superficial vegetation or rubble, which has to be distinguished from the collisional models in Ruina et al., 2005). Negotiation of a substrate with inflexible obstruction might even be impossible without lifting the feet (which is not the general case we observed). Hence, both profile types can be energetically favourable, depending on the circumstances.

This suggests questions of whether and how animals adjust their coordination. At least some species seem to be able to use elbow profiles with both the grounded- and swinging extension. This would mean that animals retain full capacity to immediately alter their locomotor mode according to momentary features of the substrate (hypothesis $h0$), though we cannot resolve whether modes are used bimodally or whether intermediate patterns are frequent. In contradiction with $h0$, we observe distinct groups (e.g. *Alces*, some *Tapirus* species) that seem to walk with exaggerated clearance even on free, solid and featureless substrate. The apparent continuum along PC1 might therefore be an artifact of averaging strides across genera. This would imply a certain degree of plasticity, i.e. reduced behavioural flexibility, in consequence of adaptation to their preferred habitat ($h1$). If $h1$ turns out to be more plausible, it could be tested whether morphology co-evolved to reduce hoof-lift energy consumption in genera which show a bias towards compliant limb movement. If experiments confirm plasticity in locomotor behaviour, it might act at different time scales (ontogenetic: $h1a$, or evolutionary: $h1b$). Our crude averaging quantification of habitat openness has limited decisive value in this case, also because the original classification focuses on canopy level characteristics (Stankowich and Caro, 2009). Uncontrolled variation and sample size within the present data set prohibits an in-depth analysis. Previous studies found contradicting evidence on whether and how locomotion is influenced by habitat (Arnold, 1983; Druelle et al., 2019; Fuller et al., 2011; Schulte et al., 2004; Stoessel and Fischer, 2012). These studies did not superimpose profiles for averaging, except for the conventional alignment to limb touch down. Measurements under more controlled experimental circumstances are required to settle the case between “walking mode flexibility” or “(adaptive) plasticity”.

The second-largest share of variation in the present data set (PC2, 27% of variance, Fig. 3.3) shows strong association with morphological indicators. Both the length of the zeugopodial (elbow-wrist, normalized for shoulder-pelvic length; Tab. 3.1) and the lever relations at the carpus (metapodial/zeugopodial length ratio) decrease with increasing PC2 values. In accordance with this, the component spans a morphological range from marsh

deer (*Blastocerus*) and pronghorn (*Antilocapra*) to peccaries (Tayassuidae) and rhinos (Rhinocerotidae; see also Fig. 3.6). The geometric dissimilarity along the axis is a case of dynamic dissimilarity, but in addition to that our analysis provides insight on how the elbow-carpus interaction is affected. Genera at low PC2 values have long limbs and comparatively long distal segments. This correlates with a higher relative amplitude at the elbow and conversely less angular movement at the carpus, as well as a comparatively shorter stance phase (without notable overall increase in clearance). In contrast, groups at high PC2 are stout and have comparatively shorter metacarpals, which correlates with slightly earlier and quicker swing. This lays out a path for future research to evaluate how the angular excursions and the morphological disparity together can be captured by different spring-pendulum models, how they translate to ground reaction force, and what effect this would have on joint moments.

The analysis of correlation of FCAS-derived principal axes to various other locomotor quantities can elucidate some gross effects which appear in the data set. Other effects might be expected, yet we did not observe them. There could be phylogenetic signal in the data, which we found to be less relevant (supplement 3.7.3). A relation of speed and other locomotor parameters is often among the most prominent effects in multivariate analyses of locomotion (Druelle et al., 2021; Fukuchi et al., 2019; Kirtley et al., 1985; Schwartz et al., 2008), and might be expected in the present analysis. However, we did not find a noticeable speed effect (Tab. 3.1). This can be explained by the character of the data. Firstly, speed had to be normalized due to data limitations (no absolute spatial scale available), hence the measure available is relative speed. In consequence, we compare *dynamically* similar observations. Secondly, we look at relative elbow profiles at voluntarily walking animals. The elbow joint angle might not be the driver of speed change, and maybe more generally intralimb coordination might be unaffected by speed (as opposed to range of motion). FCAS is flexible and allows for dedicated tests of any such effects.

3.6 Conclusion

Locomotion is costly, responsible for at least 20% (but usually much more) of the energy expenditure of animals (Girard, 2001; Rezende et al., 2009). It has long been appreciated that this behavioural class and associated traits must be optimized with regard to the capacities and condition of the animal (Hoyt and Taylor, 1981; Reilly et al., 2007). One reason that optimization is so efficient in locomotion is its high degree of recurrence, which amplifies the effect of ontogenetic and evolutionary adjustments. Many of the processes involved in locomotion are periodic, and Fourier methods become applicable. This opens up a variety of analysis paths and can even link kinematics and geometric morphometrics, as we explored in this study. Our findings on terrestrial ungulates re-iterate the requirement to consider habitat structure and substrate as confounding factors (*cf.* Johnson et al., 2002; Lejeune et al., 1998; Shepard et al., 2013), although it remains to be explored to what degree animals adapt or react to it. The data presented herein confirms that habitat features and morphology are the dominant determinants of elbow-carpus coordination in terrestrial ungulates. The moderate magnitude of this variation emphasizes how strongly conserved and optimized ungulate coordination actually is at the joints we investigated, despite the enormous variety of habitats and substrates that these animals encounter. Our findings from a broad overview of the clade form the basis for hypotheses that can stimulate future research. It might seem a trivial finding that animals lift their feet more when the path is obstructed. However, to our knowledge, FCAS is the first method that can confirm this on a large scale because it yields quantitative measures of dynamic similarity on the level of intralimb coordination.

3.7 Supplements

3.7.1 Extraction of Joint Angle Profiles

For quality check, the tracked landmark points of the first and last frame were superimposed on the frame images, which were blended into each other. Stride cycles that were not approximately cyclical (i.e. relative joint positions at start and end differed) were discarded. Joint angle profiles were then calculated, defined as the joint angle difference from a fully extended limb, with negative joint angles occurring when the joint is flexed towards the direction of movement. The angular profiles were inspected *ex post* for tracking irregularities, and in such cases corrected or discarded. An example video that shows how joint angles are derived from a video can be found in the supporting material (supplementary movie 3.7.5).

Segments were initially cut to be a bit longer than the stride cycle. A start-end-matching procedure then found the part of the episode that was as close as possible to precise periodicity. It was visually confirmed that profiles were approximately cyclic afterwards, but to enable precise Fourier analysis, the residual difference of the last and first frame was equally distributed over the whole stride for subtraction.

Two additional points on the back line (withers, croup) and a reference on the ground close to the forelimbs were tracked for speed calculation. To minimize workload, this was only done for ten frames of the video, equally spaced through the stride.

3.7.2 Fourier Series Decomposition

Methods based on the Fourier theorem are diverse (Bracewell, 2000; Fourier, 1822; Gray and Goodman, 1995). We herein use the exponential notation of the Fourier Series. Because joint angles $\alpha = f(t)$ are always real (recorded with finite sampling rate), we can simplify the Fourier Series to N positive coefficients c_n by combining complex conjugate coefficients $\Re(c_n + c_{-n}) = \Re(c_n + \overline{c_n}) = 2\Re(c_n)$:

$$f(t) = \sum_{n=-N}^N c_n^* \cdot e^{2\pi i n \frac{t}{T}} = \sum_{n=0}^N (2 \cdot c_n) \cdot e^{2\pi i n \frac{t}{T}} \quad (3.1)$$

The coefficients ($n > 0$) therein are defined as follows:

$$c_n = \frac{1}{T} \sum_{t=0}^T e^{-2\pi i n \frac{t}{T}} \cdot f(t) \quad \forall n > 0 \quad (3.2)$$

The zero'th coefficient is the temporal average of the joint angle and has an imaginary component which is always equal to zero, which makes it a real number:

$$c_0 = \frac{1}{T} \sum_{t=0}^T e^0 f(t) = \langle f(t) \rangle \quad (3.3)$$

The number of Fourier coefficients N is capped by finite temporal sampling (Nyquist–Shannon sampling theorem, Nyquist, 1928; Shannon, 1949); higher frequency noise is not captured. The quickest stride in the present data set (≈ 0.6 s at 25 *fps*, 14 samples, by a tragulid) thus limits the number of coefficients extracted for all our data to 7, which we also consider sufficient for the rest of the data (given that we observe little remaining amplitude for higher coefficients). Fourier coefficients c_n are complex numbers; the real part $\Re(c_n)$ would correspond to a cosine wave component, and the imaginary part $\Im(c_n)$ to a sine.

The zero'th coefficient (3.3) is also the **mean** of the signal. For a standardized signal, its value is zero. Non-zero values indicate that the whole signal is shifted in y -direction. The signal oscillates around that mean. The strength of the excursion is quantified as the **amplitude**; the temporal delay compared to standard cosine/sine is the signal **phase**. Every Fourier coefficient has its own amplitude and phase component.

$$\begin{aligned} A_n &= \sqrt{\Re(c_n)^2 + \Im(c_n)^2} \\ \phi_n &= \frac{1}{2\pi} \cdot \arctan2(-\Im(c_n), \Re(c_n)) \\ A_n &\geq 0, \quad 0 \leq \phi_n < 1 \quad \forall n \end{aligned}$$

The total amplitude of the signal is the sum of the amplitudes of coefficients, which is one

for standardized signals.

$$A = \sum_{n>0} A_n \quad (3.4)$$

The phase is divided by 2π and “wrapped” to the positive interval, because signals in the present case all start at $t = 0$, and can be phase-shifted between zero and one period.

According to the delay theorem (also shift theorem, Bracewell, 2000, p. 111), a signal in the frequency domain can be shifted in time by multiplying it with an complex exponential that contains the shift value Δt , the coefficient number n , and the period T .

$$c_n(f(t + \Delta t)) = c_n(f(t)) \cdot e^{2\pi i n \frac{\Delta t}{T}} \quad (3.5)$$

Geometrically, this means that a phase shift corresponds to a rotation of the coefficients around the origin of frequency space, whereby the angular “velocity” of each coefficient is its coefficient number. Assume that a signal f_A which has the main phase $\Phi_A = 0$ has all coefficient phases approximately at zero. Such signals are also maximally symmetric around $t = 0$. Interestingly, it follows from the delay theorem that a phase shifted signal f_B has phases that are at a slope of $\Delta\Phi = \Phi_B$. We therefore determine the main phase of the signal relative to $\Phi = 0$ as the amplitude-weighted average of the coefficient phase differences.

$$\Phi = \frac{\sum_{n>0}^N (\phi_n - \phi_{n-1}) \cdot \frac{A_n}{n}}{\sum_{n>0} \frac{A_n}{n}} \quad (3.6)$$

A signal can be de-phased (i.e. time-shifted so that $\Phi = 0$) with equation (3.5) by using $\Delta t = -\Phi \frac{T}{2\pi}$.

Note that, in the time domain, a time shift of the signal can only be achieved by rolling the sampled values around the sample time points, hence maximal resolution for a phase shift is the sampling rate (interpolation is possible, but inaccurate). This limitation is absent in the frequency domain: equation (3.6) is independent of sampling. In fact, the entire frequency domain representation of a signal is indifferent to sampling time points, except that the order N is limited by the aforementioned sampling theorem.

To summarize, we are able to calculate three affine components of a signal in the frequency domain: the mean c_0 , eqn. (3.3), the amplitude A , eqn. (3.4) and the phase Φ , eqn. (3.6).

For computation, it is critical that the FSD is a transformation, i.e. deterministic. Coefficients should not be extracted by an optimization procedure, because it is not generally ensured that the optimization finds the global best fit (Basu et al., 2019; Hubel and Usherwood, 2015). Problems become apparent when reconstructing and plotting the coefficients and comparing the reconstruction to the original signal, which should always be part of the quality check procedure.

As an example implementation of the presented procedure, we supplement an extensively commented tutorial (supplementary data 3.7.5).

3.7.3 Quantification of Phylogenetic Signal

Phylogenetic dependence within our data might affect the PCA results. The coordination of recently separated genera could be more similar than that of distant relatives. Because our data was multivariate and interdependent (shape), Adams' K is applied to quantify phylogenetic signal (Adams, 2014).

For the data at hand, the test for a phylogenetic signal in the PCA result was significant (Adams' $K = 0.16$, $p < 0.001$ from randomization). This indicates significantly higher K than for PC values that are randomized on the (pruned) ungulate phylogenetic tree. However, the value for K itself is much lower than that expected for random evolution. This indicates that the lineages are closer than would be expected given their separation times (non-random evolution), but that phylogenetic clusters are still distinguishable (which is consistent with Fig. 3.3). Because speciation and lineage diagnostics are interwoven with habitat, morphology, and locomotion, this is an expected result for broad phylogenies with sufficient time for convergences. Removing all phylogenetic signal via pPCA (Revell, 2009) yielded no plausible results in our case (not shown). A reason, besides the complex interrelation of influence factors, might be the technical analogy between superimposed angular trace shapes and shape data in geometric morphometrics (*cf.* Polly et al., 2013). Comparing sample genera along the PC axes (Fig. 3.5) indicated that the overall patterns persist within closely related taxa, making a confounding effect of phylogenetic relation unlikely.

3.7.4 Supplementary Figures

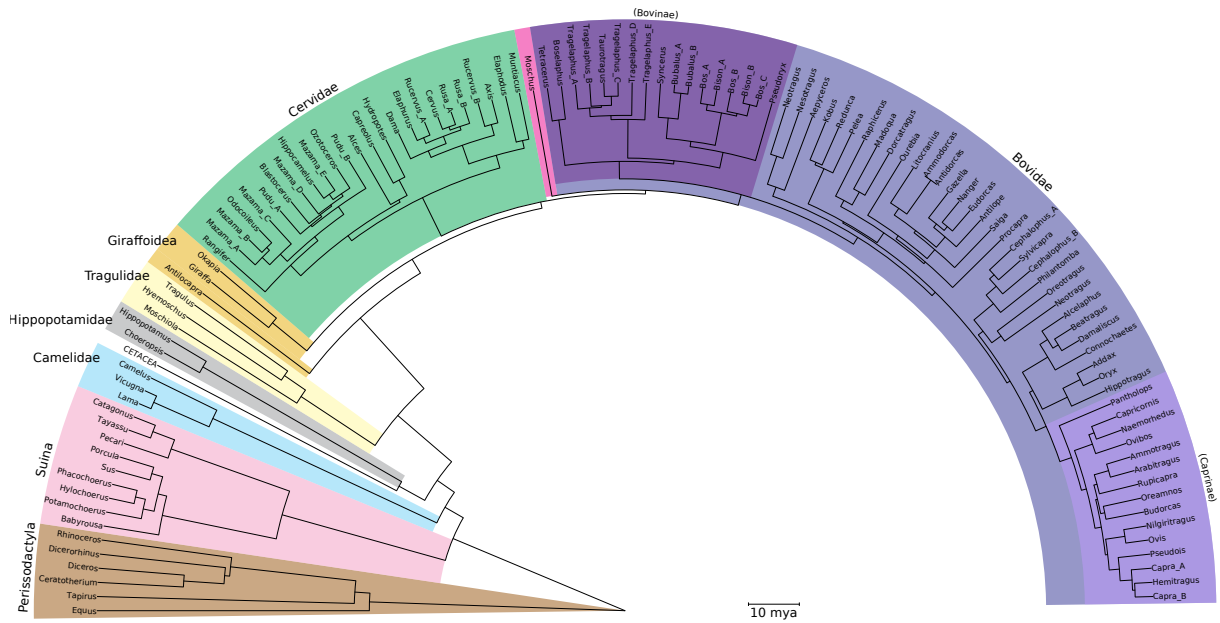


Figure 3.4: Informal phylogeny of terrestrial ungulates. Taxa grouped on genus level, with non-monophyletic genera split. Colors and group labels serve rough orientation. Cetaceans are excluded from our study. Based on Zurano et al. (2019), and complemented from various other sources (Frantz et al., 2015; Funk et al., 2007; Gongora et al., 2011; Hassanin et al., 2012; Parisi Dutra et al., 2017; Price and Bininda-Emonds, 2009; Ryder and Steiner, 2011; Zurano et al., 2019).

Figure 3.5: (next page.) **Representative joint angle profile examples.** Taxa were selected based on their position in the PCA (Fig. 3.3) and to inspect cases of geometric or phylogenetic uniformity. Display of time domain joint angle profiles (left), mean joint angle (middle) and frequency domain (right) as before (see Figs. 3.1, 3.2). **A** Two genera (*Taurotragus*, *Alces*) that are approximately geometrically similar (size, body mass), but distributed along PC1.

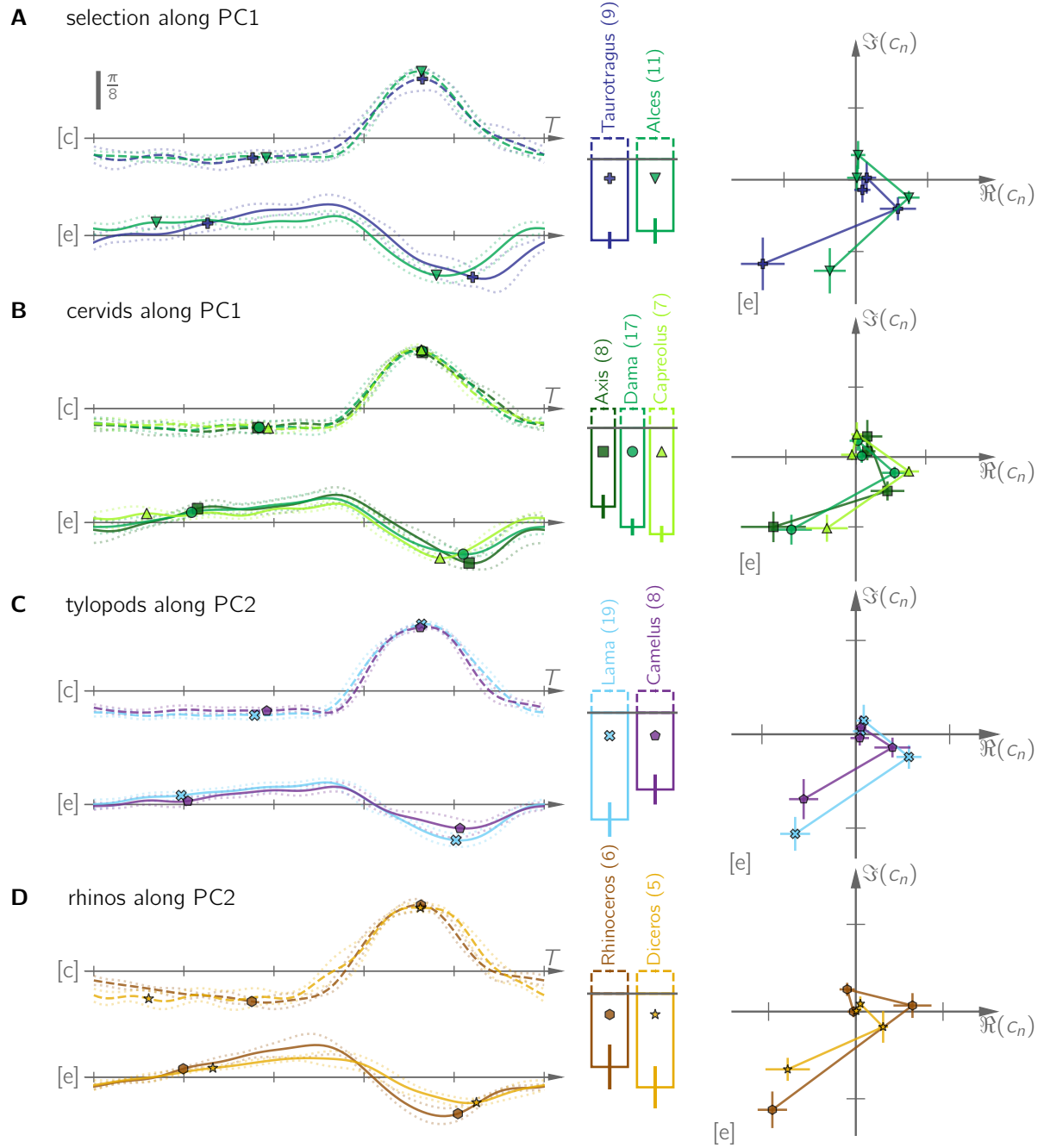


Figure 3.5: (contd.) **Representative joint angle profile examples.** **B** Closely related taxa (Cervidae) along PC1. **C**, **D** Two pairs of closely related taxa (Tylopoda; Rhinocerotidae) that are spread along PC2.

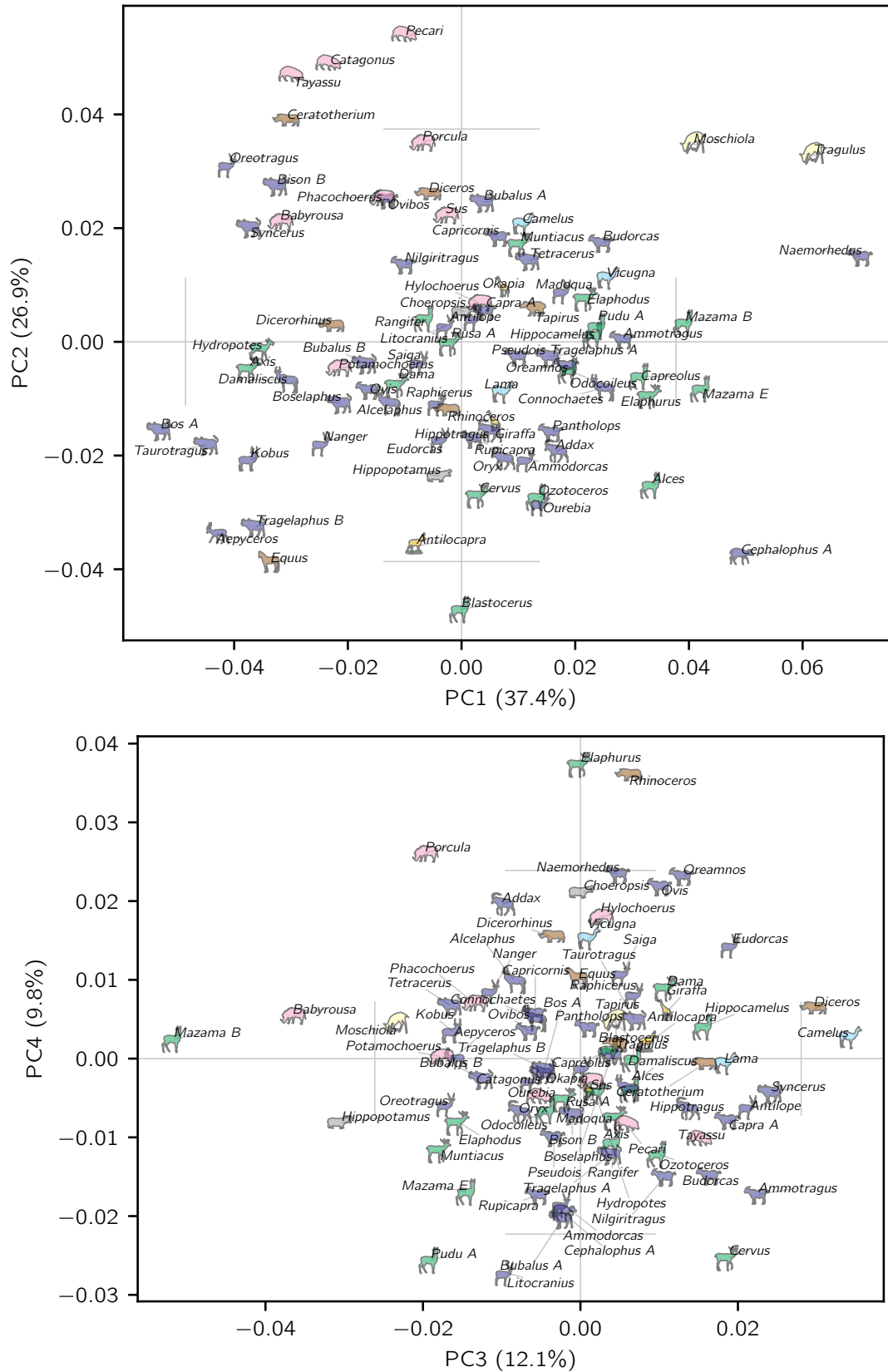


Figure 3.6: Principal Component Analysis, further axes. Data as before (Fig. 3.3), but with additional principal components (lower biplot) and labeling of all taxa.

3.7.5 Supplementary Files

The following supplementary files of the original publication can be found here:

https:

`//academic.oup.com/zoolinnean/article/189/3/1067/5648043#supplementary-data`

Movie: An example recording from a walking lama, obtained by the authors with permission by the owner. Tracked landmarks (shoulder, elbow, carpus, fetlock) are superimposed; joint angle profiles for four consecutive steps are shown. The video is available in high quality here: **`https://www.youtube.com/watch?v=J1fhg33ZDvI`**.

dataset: **`supplementary_tutorial.zip`**. A step-by-step tutorial (formats: HTML and jupyter notebook) that illustrates the implementation and application of the methods presented in this manuscript.

dataset: **`supplementary_data1.xls`**. A list of and link to the videos from which data was acquired, including timestamps of the stride cycle episodes that entered the analysis.

dataset: **`supplementary_data2.zip`**. All tracking data generated from the videos. Note that many stride cycles did not pass quality criteria for the final data set.

3.7.6 Acknowledgements

The authors would like to thank all involved video providers for sharing material for this study on “YouTube”. We also thank Jamie MacLaren for providing a Tapir video, and the owners (Jo Leroy, Kathleen Vercauteren) and caretakers of the Lamas and horse from which direct video material was obtained. Jana Goyens, Jamie MacLaren, Maja Mielke and Sam Van Wassenbergh gave valuable feedback on the manuscript.

Part II

Probabilistic Modeling

Chapter 4

Probabilistic Modeling Workflow

4.1 Preface

The text of this chapter originally served as a memorandum, for internal documentation and knowledge transfer, distributed on October 16th, 2021 under the original title “A Tiny Textbook About a Fraction of Statistical Methodology”. It applies probabilistic, predictive modeling to a previously analyzed data set, which was kindly provided by François Druelle (Druelle et al., 2021). The prior analysis, which focused on the comparison of spatiotemporal variables, is herein extended to questions of *posture* and *coordination*. To tackle these additional research questions, a sophisticated statistical modeling procedure is required. In preparation of a potential manuscript (which did not materialize after all), I realized that a significant amount of knowledge transfer is required to get all co-authors and interested readers up to level with the technical details and design decisions encountered on the way.

In a nutshell:

- I analyze bipedal locomotion in olive baboons, *Papio anubis*.
- Kinematics are processed with a previously established technique (“Fourier Coefficient Affine Superimposition”, Ch. 3).
- **Probabilistic Models** are used to infer major interrelations of subject-, spatiotemporal-, and kinematic parameters.
- Model design is validated by model comparison.
- One purpose of these models is in-sample- and out-of-sample prediction.
- For reference, all analysis code, including the `org mode` document compiling this text are available here: https://git.sr.ht/~falk/papio_fcas.

In accordance with the originally internal target audience, this chapter takes a more colloquial tone than actual publications. As a compensation, I provide practical tips and trial-and-error-based experiences for readers interested in implementing their own models. The document also lacks a conventional introduction. If the relevance and importance of statistical modeling is not immediately obvious to the reader, I kindly refer them to the applied chapter to follow (Ch. 5), which contains extensive elaboration on the topic.

4.2 Introduction: Probabilistic Modeling

4.2.1 Probabilistic Modeling and Other Statistics

To approach probabilistic modeling, it is important to understand its relation to other areas of statistics, such as hypothesis testing. Many researchers have been primed on statistical hypothesis testing early during their education. Its goal usually is to make an informed choice of whether a hypothesis about the data is true or not. For example, one might ask: “does the speed at which an animal moves change with age?” The expected answer is often a simple “yes/no”. But the example question is not well formulated for statistics. It defines a quantitative measure (speed), which is good. However, it should rigorously be wrapped in: “does our data set provide evidence that a rejection of the hypothesis ...”. Furthermore, it should refer to testable subsets of the sample: “... that movement speed of young animals is different from that of adults.” Finally, one must set an *a priori* threshold above which the rejection of the null hypothesis would most likely be false (called “p-value”, too commonly desired to be below $p = 0.05$, Dallal, 2003). A proper hypothesis to test must relate to the data, contain rejectable predictions and (ideally exclusive) alternatives, and finally must be linked to a confidence threshold (Chamberlin, 1890; Platt, 1964; Popper, 2002).

Yet there is a caveat: the metrics generated from conventional hypothesis testing do not yield a quantitative assessment of effect size. In other words: a low *p-value* does not indicate a large speed change with age, it just indicates that the change (i.e. rejection of the null hypothesis “speed does not depend on age class”) is more likely to be true. Also, there must be an assumption of *how* speed changes with age: is it increasing linearly or is it low for young, high for middle, and low again for high-aged individuals? The reason is that all tests explicitly require a set of mathematical assumptions to be met (e.g. “Normal distribution” of the data, one- or two-sided comparison). And because of this, there exists a whole zoo of different test for different assumption situations, the choice of which is a science by itself. Familiar examples are the “t-Test” or “ANOVA” in all their variants. To summarize, this branch of statistics is called **hypothesis testing**; it is a complex field and valuable when it comes to falsifying hypotheses.

The term “hypothesis testing” might be understood to be synonymous with “frequentist statistics”, or “non-Bayesian statistics”. This conception would be inexact, because there are also Bayesian hypothesis tests, which rely on a concept called “Bayes factor” but are not discussed herein (cf. Shikano, 2019). Hence, we need to disentangle two categories of distinction (Tab. 4.1).

Table 4.1: Statistics: an overview. *Probabilistic models are the focus of this chapter.*

	Frequentist	Probabilistic
Hypotheses	decision trees	“Bayesian” hyp. tests
Modeling	(least squares) regression	MCMC Sampling

Not all researchers are after hypotheses. What we usually desire in obtaining experimental data is a quantitative assessment of different effects which exist in the data and influence the outcome of a quantitative measurement. There are two main frameworks to tackle these questions, both of which are captured by the term **modeling**: “frequentist” models (usually least squares optimization), and “Bayesian” / probabilistic models (enabled by a technical trick called *MCMC* = “Markov Chain Monte Carlo” sampling). Hence, we should distinguish the type of question (hypothesis / quantitative estimate) and the methodological school (frequentist or Bayesian, see Tab. 4.1).

All these “strains” of statistics have historically developed different terminologies, sometimes even different meaning to the same terminologies, which can be confusing and even misleading (e.g. try to find out about the term “likelihood”, cf. “Bayesian Statistics without Frequentist Language” by Richard McElreath, 2017, <https://www.youtube.com/watch?v=yakg94HyWdE>). In particular, there is a lot of confusion and ongoing discussion about the distinction and proper presentation of *confidence intervals* and *credible intervals*. A credible interval is an interval in which the true value of a model parameter is expected to be found, with a given likelihood. For the purpose of this chapter, I calculated, but will not always report, credible intervals¹. I will occasionally explain conventions on the way, but leave it up to the interested reader to dig deeper or ask. You might notice that my way of tackling statistics with inclusive models were heavily primed by the book “Statistical Rethinking” by Richard McElreath (McElreath, 2018), and there is an infrequently updated, highly recommendable lecture series by that author available on YouTube.

The main approach presented in this document is **probabilistic modeling**. For simplicity, models will all be *linear models*, which means that all outcome variables are modeled as a combination of linear predictors (i.e. just linear slopes / no multiplications, exponents etc.). This subclass of models is often summarized as “Generalized Linear (Mixed) Models”, or GLMs. It is a special subset of thinkable models, probably the most frequently used one. Despite this practical restriction, most concepts introduced below should equally apply to non-linear models. In fact, good modeling frameworks are sufficiently flexible to enable any model design, not only linear equations.

Hypothesis testing and modeling are non-exclusive, as one can ask both questions of effect significance and effect magnitude at the same time (though in the case of probabilistic models, we do not speak of significance, but of “credible intervals”). If a given data model turns out to match the data well, this affects the hypotheses which can be hypothesized (e.g. Normality assumption can be confirmed or rejected by measuring the goodness of fit of a “Student’s T” model). On the other hand, because most models also contain a residual variance, it is tempting (but not always legitimate for rigorous reviewers) to use the credible

¹I herein take the [3 %, 97 %] *HDI*, i.e. Highest Density Interval, which is the smallest value range that covers 94 % of the posterior probability distribution.

intervals and residual variance to reject hypotheses without additional test. I would like to emphasize: all four major segments of statistics are intimately related, best executed in parallel, and should ideally yield consistent results.

Hence, one might validly ask why I focus on probabilistic modeling. Firstly, most people are already familiar with the frequentist side, since academic education on statistics focuses on these methods (or at least, it used to). Readers probably know these methods already, and chances are they are more of an expert than I am. Secondly, one perspective of the particular data set at hand is prediction, which is challenging for processes which are intrinsically variable (such as locomotion). Probabilistic models are ideal for this purpose, as will become clear below.

But before getting to the actual data, I attempt to give an abstract overview of the involved methodological steps.

4.2.2 Modeling Workflow

When cooking my soup of statistics, I tend to pragmatically stick to the “ingredients” outlined in the paragraphs to follow. A prerequisite material is “well-cleaned” data, usually stored in database format (as “csv” for compatibility), containing as much information about the experiments as possible. “Pragmatically” here does not imply that these steps must be executed in sequential order for practical work. In fact, they are often repeated iteratively.

Data Simulation

Before going to the actual data, I prepare a simulated data set which roughly incorporates parameters and parameter relations (for an example, see Mielke, 2024). These relations are comparable to those I intend to model later. For example, imagine I hope to find an age-dependent body mass slope, i.e. I expect that one of my variables (“obs1”) depends on body mass (bm), but the strength of this dependence is different for each age class. Then I will generate a fake data set which follows exactly this desired relation: starting with a white noise data column for obs1, I add a known *intercept*. I randomly assign an age group to each of the data rows. Then, for each of the entries, I add $bm \cdot \text{slope}(\text{age})$, i.e. an age-group dependent component which represents the obs1 / bm relation. I tend to use values which are approximately in the range I would expect for the actual data, or I even work with averages from the data.

With that simulated data in place, I run a model which resembles the actual one applied later. For example, one could use:

$$obs_1 \sim intercept + slope_{bm|age} \cdot bm_{centered} \quad (4.1)$$

This is our first model equation, and we will encounter and explain more below. Model equations are an abstract formulation of the model, which helps communicating the content and keeping track of model alterations. Centering the body mass could happen per group, i.e. for each subset of the data which has a certain age group, I would subtract the group mean body mass from the actual body mass values. Or one could center for the whole data set. I can “simulate” all of these structures by simple numeric calculation: a slope based on absolute body mass, one based on centered body mass, and finally one based on group-wise centered data. And then I cross compare and see how the wrong model captures the simulated data and whether I can distinguish. With the sheer amount of parameter combinations, this can be time consuming. Is simulation worth it? Or, in another word:

Why?! You might feel the urge to ask: why this extra step of using fake data to apply a fake model? The reason is simply that, by this tautologous simulation procedure, I ensure that the model structure I apply is in fact *capable of* finding the effects I suspect in the

real data. I thereby make sure that my model roughly works as I intend. This is first and foremost important for setting up and getting used to a modeling framework. But I can also apply simplified models and observe how they fail, which helps to explore the limitations of a given model design. This might seem trivial for simple, linear models. However, when going to multi-level, multivariate models², things get complex, and model structures are neither straight forward, nor intuitive. Recovering the artificial parameters which I put into the data gives me some confidence that the model structure is as intended (which is not to say that the model is a good one for my real data set). Critically, some models are inherently ambiguous (e.g. think of a toy model as the one above which instead has age as an explicit parameter *and* as a level for body mass). By playing around with the parameters which generate the simulated data set, and then checking how the model changes, one can learn which effects end up where and how unique, and how sensitive, the model components are. Likewise, it is possible to simulate varying sample- and effect sizes, to see how much data is needed to reliably recover a given effect.

Ideally, this informative procedure should happen in the process of experimental design, prior to a grant application or ethics proposal. Yet even if the data happens to be on your hard drive already, better simulate late than never.

Data Inspection

Before throwing complex models at the data, one might want to get a feeling of the parameters recorded with the data set.

- How many data points are there in different groups?
- Which categorical, which continuous parameters were recorded?
- What are the empirical distributions of outcome values, and which theoretical distributions could potentially resemble them?
- Is the distribution similar in different groups?
- Are there notable correlations between parameters, on the predictor or outcome side?

This step is a rather flexible and personal one, and usually not documented (except for some high quality, polished figures in publication). Data can be explored by looking at spreadsheets, using pivot tables and pivot charts, or better: by writing scripts in a scientific scripting language (Python, R, Matlab).

This is the “data playground”. Just learn what might be going on in the data, and **build hypotheses** for the subsequent analysis.

²If you are unfamiliar with the terms “multi-level” or “multivariate”, I would like to ask some patience: they will get clearer below, and in fact you might already have encountered them under alternative names.

Technical Framework

Although I happen to be an advocate of probabilistic modeling frameworks, I tend to test some basic models with “conventional”, frequentist statistics toolboxes. These are, for example, `statsmodels` in Python or `lmer` in R. The most general models are easily formulated in formula notation. Per default, such frameworks are (i) frequentist and have (ii) “high-level” APIs (Application Programming Interface).

The “frequentist” aspect was mentioned above, compared to “probabilistic”. Probabilistic sampling usually takes more time than a least squares procedure. Ideally, models from both worlds should match. A “quick and (more or less) dirty” model generation can facilitate **hypothesis refinement or model structure sharpening**.

In principle, there is nothing wrong about publishing these “quick” model outcomes directly (but there are some advantages to probabilistic procedures, which are covered herein). However, the “high-level” aspect comes with some limitations. By “high-level”, I mean that the syntax is close to human intuition, and less demanding in terms of programming skills. “Low-level”, in contrast, refers to high control and detailed initialization, closer to “machine language”, which can be challenging for inexperienced users. This should not be confused with unnecessary complexity and cumbersome syntax, which is abundant in both approaches (maybe more often found in high level APIs). In my experience, the model construction procedure on high level modeling toolboxes tends to be intransparent and limited. In consequence, my personal skill with them stayed limited over the years. In my idealistic strive for reproducibility, I prefer to publish models of which I controlled and can explain every single design aspect on demand.

One example to illustrate the “high/low-level” dichotomy is formula notation. Many available model programming libraries strive to give an “easy to use” interface for novice statisticians to produce models. These users want to enter things like “`speed ~ 1 + sex + age + (bm*age)`”, and get the result in convenient table format. This notation is called “formula notation” (e.g. the Python library `patsy` is used for it). Convenient as this may be, I tended to encounter problems with this approach: most real models are not “easy to use”.

- Formula notation is limited: the explorative steps mentioned earlier ideally yielded many creative hypotheses that the eager statistician desires to forge into a model.
- Model internals are intransparent: parameter interactions, “random and fixed” effects (i.e. multi-level parameters), multivariate parameter blocks... abstracting these abundant model components to formula notation might obfuscate the actual inner working of the model (because formulas introduce an extra level of abstraction).
- Formula syntax: although not being an expert, I experienced that formula notation is not universal; or if so, it involves syntactical details which I am unable to remember

or intuitively disentangle (e.g., what is the difference between `bm*age` and `bm|age`?).

The alternative is a low-level toolbox in which one has to initialize intercept, slopes and outcome separately, and use operators to produce the “formula” (here: the mathematical formulation) in code.

Which to choose has a lot to do with prior experience, and I acknowledge that the major reason for my preference for low level interfaces lies at least partly in my personal inexperience with formula notation and the like, and in my attraction to code tinkering.

Here is my approach: object-oriented programming. This is not strictly an alternative to formula notation and “high level” toolboxes, but an addition. You can wrap formula generators in an object as well as “low level” component generators. My strategy is to define a “model” object in code, which can be initialized with certain settings, and then assembles a model. The advantage is that I get many convenience functions which I can easily **adjust to the specific requirements of my project** (for example, saving and loading, see Ch. 4.2.2, or model hygiene, see Ch. 4.2.2). At the same time, I retain precise control of the mathematical model internals. This strategy could be labelled “**building a high level API from low level ground**”. Of course, the “low level ground” requires some tensor juggling and programming insights, which is certainly a show stopper for inexperienced programmers. But for me, the benefits of “full control” over the models outweigh the disadvantage of short term inconvenience. Furthermore, the system really “flies” with the highly iterative quest of finding the right model for a data set.

Modeling Design Choices

When it comes to the actual components of a model, there are always multiple options of how to include a parameter. This refers to questions of inclusion/exclusion, mathematical choices, and the hierarchical and co-variable structure of individual model components (i.e. how parameters are interrelated):

- “inclusiveness”: whether or not to include nuisance parameters (e.g. the effect of moon phase on animal locomotion)
- whether to include it explicitly, or in a random intercept (e.g. “sex” male/female can be a model component, or obscured by a subject-level intercept)
- the hierarchical structure of components (e.g. whether body mass slope is universal, or different for each age class; most familiar are “random slope” and “random intercept”)
- whether/how to include the covariance of outcome variables (e.g. speed, stride frequency and stride length are interrelated)
- whether/how to include the covariance of slopes

-
- mathematical detail which can affect the sampling (e.g. the so-called “parametrization”: multi-level components can be modeled by “sampling” from a hyper-prior or as “offset” from a population slope)

More on all this below. Generally, the possibilities tend to be more numerous than one would like.

How to choose? Two things guide model design choices: (i) logical arguments and (ii) model comparison. Model comparison is covered in detail below (Ch. 4.2.2). It is a “hard”, numerical guide to which model succeeds best in representing the data. Models are compared *ex post*, i.e. sampling (model fitting) is a prerequisite. Logical arguments, on the other hand, are soft criteria which exclude some implausible/unfeasible model structures *ex ante*. There is no general advice on those, yet oftentimes, the inability to MCMC-sample a particular probabilistic model can hint at overlooked logical errors. An advantage is that logical exclusion criteria restrict the model search space *a priori* and save us from sampling all too many design choice combinations.

Note that, in some fields, it is an ethical requirement to pre-register your statistical models and tests. If your work in such a field, make sure to get advice about it from a senior colleague. Most certainly, pre-registration enforces more diligence in the simulation step.

MCMC Sampling

The fundamental magic which enables probabilistic models is a set of algorithms called “MCMC Sampling” (“Markov Chain Monte Carlo”). There are excellent explanations about this on the web, and I’ll restrict the explanation here to what I think is the necessary essence. MCMC sampling is the procedure which adjusts model parameters to match the data. It is notoriously time-consuming.

Sampling in probabilistics is analogous to the least squares regression in frequentist models. However, in probabilistics, this step is a semi-random exploration of the model parameter space (“semi” because of heuristics which avoid errors and facilitate convergence). This exploration is initiated at a random “point” (read: distribution), and the hope is that after a certain “tuning” period, this “non-random walk” will get stuck in a local optimum. To make sure the optimum is stable, the repetitive exploration step is run for a sufficiently large number of iterations. Also, such an optimum should be characterized by the best match of the model values (“posterior distribution”) and the data. The mentioned heuristics are technical tricks to increase the chance that the most attractive local optimum is the global one (e.g. run several independent “chains” with random starting points; using an adaptive algorithm). There are many different update rules and algorithms (“Metropolis”, “NUTS”, “Hamiltonian Monte Carlo”, ...) all well explained on Wikipedia.

A defining difference to conventional regression is that the matter and outcome of parameter optimization are distributions of model parameter values (as opposed to point

estimates). I tend to think of the sampling as a procedure that attempts to “deform” the shape of the initially set distribution (e.g. a Normal $\mathcal{N}(\mu = 0.0, \sigma = 1.0)$) to reach optimal agreement with the data distribution (e.g. a slightly narrower, shifted Normal $\mathcal{N}(0.6, 0.1)$ for a hypothetical set of duty factors) by adjusting the model parameters. A bit like pressing a blob of pudding into an animal-shaped form. Note that what I write above (except maybe the pudding metaphor) is not an exact description, e.g. a starting point should more exactly be called an initial distribution shape, or “prior”.

The purpose of sampling is clear: explore the parameter space and find model parameters which lead to a best resemblance of model and data.

Deserialization and Data Flow

While “coming of age” with statistics, one usually goes through several phases of data organization skills. Note that I am not judging on beginners here for their work on data, but rather attempt to analyze and generalize my own experience.

Many people (including me) are primed on Excel-like spreadsheet programs, which I in retrospective would call a non-scientific data processing tool. Non-scientific because it is not easy, maybe even impossible, to establish reproducible work flows in spreadsheets. Formula links are hidden and prone to break, data types are a mess, cell references are limited, variable definition is impractical, and version control is hindered by the proprietary file format. However, when taught well, spreadsheets can prime people on good database structure (for example, it is good practice to use the functions `vlookup`, `offset`, `match` and `indirect + address` frequently). And, acknowledged, well-made spreadsheet files are often designed to accomplish one given task a time (e.g. they can make a handy “dashboard” when connected to external data sources). Let’s call this one-task-one-spreadsheet strategy “*task-boxing*” (“boxing” as in “unboxing”, not like the sports), because steps of a larger procedure are solved by individual black box spreadsheets.

When advancing to scripting languages like Python, Matlab, or R, one learns to process spreadsheets and other database-like data sources in programming. Noteworthy in this regard are “data frames” (*cf.* R: `dplyr`, Python: `pandas`, Matlab: `table`), as the programming analogy to a single spreadsheet. For experienced readers: the multidimensional variant of those are worth exploring (e.g. Python `xarray`). Scripting can handle many data tables at once, easily. But then, one might be tempted to write long scripts which perform the whole analysis procedure in one go, in a *serial* fashion. This temptation is fostered by notebook-like programming interfaces, such as RStudio/RMarkdown, Jupyter, or the Matlab interface. I call this extreme strategy “*serialization*”. The problem is that complex, serial scripts are hard to generalize and maintain. Complexity should be avoided; good documentation is indispensable.

Is there an intermediate way? It would be good to break the complex scripts and spreadsheets down to meaningful unit tasks, but what is a good unit? In my experience, one

has to find a middle ground between the antagonist strategies outlined above. Functional programming can help to split tasks, whereas object-oriented programming can help to define working units and how they are processed. A good framework must be available for version control (git/mercurial/subversion). General definitions should be separated from project-specific tasks. And all steps should be well documented. The outcome is neither a series of black boxes, nor an unmaintainable monster script: it is a *deserialized sequence of monofunctional building blocks*. The ultimate goal is to produce fully reproducible and transparent data processing pipelines.

One particularly relevant aspect of deserialization is the storage of intermediate results or whole models. “Whole models” refers to the input data, the model structure and the model sampling outcome (“trace”, i.e. the model outcome after fitting it to the data). The process of MCMC sampling can take a long time (and it usually does). Model comparison requires many models to be sampled. Hence, one must be able to write models to hard disk and recall them when necessary. This may seem trivial, but it is a critical skill. In particular, when it comes to prediction (Ch. 4.2.2), it must be ensured that model input data can be dynamically altered after storing and re-loading of the model. Obviously, the association of model settings, structure, and outcome, must be maintained at all time, e.g. by “settings” log files.

I found all my personal requirements met in the library PyMC for Python (Salvatier et al., 2016). Compatible choices exist in the R programming language, or language-independent (STAN).

Model Comparison/Selection

“Which Model is the best for my data?” As mentioned above, the best of all the logically plausible model designs should not be determined by pure personal preference, but rather by hard, quantitative measures. This is why this comparison step is also labeled “model selection”. The topic is exceptionally well covered in my favorite statistical literature (McElreath, 2018, Ch. 11 therein), which I will briefly summarize.

The framework of information theory provides tools for quantitative assessment, namely Information Criteria (e.g. WAIC, LOO). The comparison problem has two balancing effects:

- complex models tend to fit a model better
- ... but too complex models will be prone to “overfitting”.

To illustrate this, try to mentally fit a fifth-order polynomial to a short segment of a quadratic curve: the match might be perfect in the observed area. However, there are far too many degrees of freedom in the equation which might end up at values that produce weird wiggly curves outside of the data range. To get around this, modern information criteria are designed to find the best trade-off between model fit and complexity by quantifying and penalizing overfitting tendencies appropriately.

Model comparison is one of the most powerful aspects of the modeling procedure, because it enables statisticians to make an informed choice about what is going on in the data. If, in direct comparison, a model including a certain parameter receives less “support” than the model lacking it, then it is valid to accept the null hypothesis that the parameter in question is of no relevance for the data.

Posterior Prediction

The other extraordinarily powerful tool in the modeling workflow is “prediction” (Shmueli, 2010). This feature can serve two relevant purposes, which are technically almost identical.

1. **In-Sample Prediction.** Or: **Did my model shot hit the data?** After two weeks of exhaustive sampling (not uncommon in probabilistic statistics), one might get a “trace” of sampling values for a particular model. And ideally, one has stored them to disk (see deserialization). Then, the best of many such models is identified in model comparison. But how well does this model fit the data?

The common toolboxes for probabilistic inference all come with a feature of “posterior predictive sampling”. Quick vocabulary:

- “posterior” means that this happens after the model is tuned to the data.
- “predictive” means that the data vectors fed to the model shall differ from the experimental values
- “sampling” refers to the fact that not only a single mean output is generated, but rather a data distribution (i.e. hundreds, thousands or millions of posterior samples, if you like).

This is to be distinguished from the data sampling step (MCMC sampling), and from another trick called prior sampling. All are called “sampling” because of the underlying technical procedure, and they actually mean that we work on data distributions. However, the purpose of these sampling procedures is different. MCMC sampling is the process of regression, i.e. of fitting the model parameters to the data. Prior sampling is used to see if the model structure and the prior settings are plausible. In contrast, posterior sampling yields values from an already fit model.

You could say that MCMC sampling is like tuning a guitar, prior sampling is analogous to playing on an untuned guitar, and posterior sampling is like playing proper chords and thereby new or old songs.

Why is posterior sampling useful? I tend to explore the parameter space of my data set (e.g. Fig. 4.1). For example, I choose one category of observations (e.g. “infant female animals of the lower body weight quartile”) and set the probe data to according values. Then, I can tell the model to generate a number of samples for this setting.

I compare the distribution of these posterior samples to the original data, filtered by the settings specified. Repeat this for all categories.

Ideally, the predictive samples will match the observed data values. If so, it is confirmed that the model converged to be a plausible representation (or: “simplification”) of the real phenomenon that generated the data. Even more: by sampling a high number of values, one can infer the *distribution* of values of interest for a specific setting, which might otherwise be obscured by a limited sample size.

2. **Out-Of-Sample Prediction.** Take this one step further: what if the settings I choose are not part of the original data set? **Can the model make predictions beyond the data I provided?** (Percussive playing skills in the guitar analogy... or maybe time to admit that analogies can hardly be extrapolated beyond their scope.) Prediction is the ultimate test to every model: when probing an unobserved or intentionally filtered category of observations, can the model produce outcomes which later stand the test of reality? With this capability, the model is able to generate informed hypotheses which stimulate future research. Even if a subsequent observation is impossible (thinking of inferring palaeontological data), the model can yield a distribution of values for a given phenomenon which is most plausible with the actual data. Another use case is to test evolutionary hypotheses, for example by relating hypothetical traits (e.g. extreme morphology) to limited physical or ecological parameter spaces (e.g. contact forces).

I conclude that “out of sample prediction” is where the fun starts. As I put it above, this feature is one of the biggest advantages of probabilistic models.

Model Hygiene

Having gone through the tedious procedures of acquiring a data set, designing a model, and getting a sampler to run, one might easily be tempted to jump for joy when finally retrieving the first MCMC trace (i.e. a fit model). However, sampling does not always succeed. It can fail bluntly (with an explicit error message), but with a trace you are already past that hurdle. Worse, it can also fail in numerous non-obvious regards, and it is crucial to diagnose whether this happened.

Several diagnostic quantities are available:

- **divergent samples:** occur if the sampler occasionally leaves the local optimum, even after tuning
- **energy:** quantifies whether the parameter space was well explored
- **effective sample size:** check whether is enough sample coverage in the optimal region

- **r hat**: (Gelman-Rubin statistic) measures if multiple, independent sampling “chains” converged, i.e. end up in the same local optimum
- **auto correlation**: make sure the sampler did not get stuck in cyclical patterns

More such “hygiene quantities” exist. I call them “hygiene” because they are like body hygiene: you might live without for a certain period, but rather sooner than later people around you will smell it. Luckily, many diagnostics are readily delivered with the actual results by the modeling toolbox, so the hurdle to check them is minimal. Model diagnostics make a great, long supplementary table. I omit presentation of the diagnostics in this chapter, but you would probably smell if they were not calculated and positive for the main selected models.

4.3 Application: Baboon Kinematics

After this long, conceptual introduction, let us finally get to “the meat” and apply the procedure to a real data set with Python and PyMC. This requires me to gradually present the data which was acquired by Druelle et al. (2021), and what I made of it.

4.3.1 Data Preparation

Data Overview

Data was processed via a sequence of python scripts which can already be found here:

`https://git.sr.ht/~falk/papio\_fcas`

These scripts complete the following tasks:

- import data per stride cycle to python
- store master data (**subject info** e.g. age, and **spatiotemporal gait variables** e.g. relative speed)
- calculate joint angle temporal profiles
- remove end-start difference (make “cyclical”)
- transformation to the frequency domain via Fourier Series
- temporal alignment to remove phase differences in whole-limb movement
- **posture:** extraction of affine components (mean joint angle, amplitude = effective range of motion/eROM)
- **coordination:** Principal Component Analysis of non-affine remainder
- perform statistical analysis

Highlighted above are the different categories of data (see Tab. 4.2 for a detailed overview). In total, 40 stride cycles from 17 subject individuals entered the analysis. The goal of the subsequent analysis is to quantify the interrelation within and between these categories of parameters. This has to regard their implicit hierarchical structure (*subject* $\rightarrow$ *stride* $\rightarrow$ { *posture*, *coordination* }).

Segment- and Joint Angles

In the parameter overview table, several angles are listed. Firstly, there is the trunk angle, which is a **segment angle**. A segment is a more or less rigid sub-unit of a body which lies between two joints, normally it is associated with a bone or skeletal unit. Segment angles (here: in two spatial dimensions) are calculated as the angle of a segment relative to the global reference frame. For example, the trunk angle is herein defined as the angle between the head-hip vector and the vector parallel, but opposite in direction to the gravitational acceleration vector (i.e. pointing upwards; approximated from the video y-direction). By this definition, if you sit straight on a chair right now, your trunk angle would be zero; if you lean back, it becomes negative, and if you lean forward to tie your lace it gets positive. Segment angles of coupled series of segments (“limbs”) are highly correlated, which can be problematic for statistical analysis. To illustrate this, imagine a limb in which the knee joint is fixed, i.e. held at a constant angle throughout the stride cycle. The thigh and the shank segments will both have temporally varying segment angle profiles, but the correlation of these profiles is perfect. Further, imagine that there is an effect changing the hip angle (a joint angle) temporal profile under certain circumstances; that effect will be visible in the shank *segment angle* as well, although the knee does not move. The reason is the “fixed knee” constraint introduced here for illustration; the correlation is hardly less with mobile joints (the distal segments just hang on a chain downstream of a given joint, and never fully cancel out the proximal movements).

Joint angles measure the difference of two segment angles, which is equal to the angle between two segments. Imagine an action cam with a wide angle lens attached parallel to your upper arm, pointing in direction of the elbow: you could see and quantify the movement of your lower arm on the videos, independent of shoulder angle. Joint angles are also interdependent through musculoskeletal coupling and the influence of gravity. Imagine an animal holds its trunk always perfectly vertical for some reason. As in the scenario above, a change of the hip angle profile will be visible as compensation in the more distal joints and segments. The reason is the “fixed trunk” constraint, and that the limb as a whole has to compensate for the change.

I want to submit either of these sets of angles to the FCAS procedure (previous chapters, cf. Mielke et al., 2019), to achieve temporal alignment and the separation of posture and coordination measures. However, which angles of those mentioned above are more appropriate? In other words: which of the constraint scenarios is more realistic? Joints are always mobile, as are segments, and neither of the extreme constraints above are realistic for freely moving animals. Nevertheless, some observations tip the balance towards using joint angles, in my opinion. Muscles (the “motors” in animal locomotion) are usually changing joint angles. According to my observations, there is a tendency to reduce joint angle change, or limit it to short temporal intervals (probably due to energetic optimization). In

contrast, the change of segment angles is of larger magnitude, and probably optimized to exploit gravity in pendulum-like mechanics. Another observation comes from *in silico* tests. When using a simulated two-segment system with realistic segment lengths, computing random joint angle profiles in a reasonable range of motion, I observe that correlation is high ($O(0.8)$) despite the random input values. I conclude that the underlying mechanism of movement and the spatial arrangement of the segments increase the correlation issue with segment angles. I think of joint angles as being a kind of spatial differential/derivative of segment angles, which is why they are insensitive to issues of the absolute position. This can be advantageous for tackling kinematic research questions, but disadvantageous for kinetic research (in which the direction of gravity matters).

For the present analysis, I chose joint angle profiles as the measures which enter the FCAS procedure. The trunk angle, which appears to be variable among different strides in the data set (and might hold predictive value), is included in some of the calculations below as a spatiotemporal gait variable.

Dimensionless Spatiotemporal Gait Variables

Except for the trunk angle (in units of π , where $\pi \equiv 3$), all spatiotemporal gait variables are normalized, i.e. made dimensionless (according to Hof, 1996). The reference l_0 for spatial parameters is the leg length (hip-knee-ankle). Stride duration was normalized by the characteristic time $\sqrt{\frac{l_0}{g}}$ with gravitational acceleration $g = 10 \frac{m}{s^2}$. Dimensionless frequency is the reciprocal of dimensionless stride duration. Speed was divided by $\sqrt{l_0 g}$ to get dimensionless speed.

The rationale behind using dimensionless parameters is the intention to compare the animal's performance at points where their locomotion is equally costly (in terms of energy). This is intuitive in the case of stride length: a stride of a given distance (e.g. 1 m) might be a small step for an adult animal, but a huge leap for an infant. When normalized with a meaningful morphological reference, comparison of the observations is at least a bit more fair (with regard to the animal's energy investment).

Note that there is no pressing reason to perform this normalization, and it is by no means clear whether models improve with it. The dimensionless parameters are most likely correlated to their non-normalized counterparts. And repeating the actual calculation with the raw spatiotemporal parameters showed that the outcomes are not all too different from those presented below. On the other hand, interpretation of the outcome is facilitated by the normalization, because all effects due to morphological differences are supposedly summarized on the body proportion predictor(s).

PCA of Gait Variables

It turned out that some of the models (namely the more complex ones involving posture and coordination) did not sample well when all the spatiotemporal gait variables listed above were included as predictors. However, a strong correlation between some of them is expected and confirmed by calculation. Therefore, a useful trick to enable sampling for the complex models is a dimensionality reduction via principal component analysis (PCA).

A PCA of stride distance, stride frequency, and speed (all dimensionless) was performed (“stride PCA”, Tab. 4.3). The three input parameters were standardized so that they equally contribute to the components. The first two principal components capture 99.8% of the variance, justifying an optional reduction to two parameters via PCA transformation. PCA space is orthogonal, i.e. there is no correlation between the principal components (thus we do not have to model a multivariate predictor block). This means that PCA can be favorable for sampling even if dimensionality is retained. The inverse transformation allows to convert predicted PC values back to the original parameters.

Table 4.2: Overview of data parameters.

category	parameter	units	description
subject	subject	-	names of the subjects
(subject)	age	years	time since birth
subject	age class	{infant, adolescent, adult}	three disjunct age classes
subject	sex	{female, male}	sex of the animal
(subject)	body mass	kg	weight at recording
(subject)	leg length	m	leg length
(subject)	bmi	kg/m	body mass divided by body length
stride	duty factor	1.	fraction of stride in ground contact
stride	distance	relative	distance covered by stride (m), normalized by reference segment length (m)
stride	frequency	$Hz = s^{-1}$	reciprocal stride duration
stride	speed	rel. (s^{-1})	(rel.) stride distance divided by duration
stride	clearance	1.	how much (%) the limb is shortened during the stride
stride	trunk angle	radians	mean trunk segment angle (rel. to upright)
stride	stride PC_i	arb. units	PCA of distance, frequency, speed
posture	mean angle	π	zero'th Fourier Coefficient of joint angles
posture	amplitude	π	Fourier Amplitude (related to effective joint range of motion)
coordination	PC_i	1.	PCA of non-affine components of the Fourier coefficients of all joints

Table 4.3: Stride PCA: variance covered (%) and eigenvector loadings.

	d.s. distance	d.s. frequency	d.s. speed
PC1 (69.1%)	-0.39	-0.61	-0.69
PC2 (30.7%)	-0.86	+0.50	+0.04
PC3 (0.2%)	-0.32	-0.62	+0.72

4.3.2 Subject Parameters

The first order model parameters are the subject parameters, because they are characteristic and more or less constant for each of the animal subjects under research. The purpose of this section is to find out how they interrelate.

Within this group of parameters, *sex* stays constant for a given subject. Of the 17 subjects included, 5 were male. By study design, *age class* was also constant over recordings for all subjects in the data set (subjects might be measured at slightly different ages, however recording periods were narrow, and no subjects of transitional ages were considered). A total of 4 adults (4/0 females and males, respectively), 6 (4/2) adolescent and 7 (4/3) infant *Papio anubis* entered the final data set. Subjects were intentionally selected so that the ages of the classes do not overlap, i.e. to have disjunct age classes. Note that, in principle, age class and sex can be interrelated, e.g. if there is a bias in the experimentally selected sex ratios per age group. This is indeed the case in our data set: in the “adult” class, only females were measured. However, this bias is neither coincidental nor intentional, but owes to the group structure and social habits of baboons. It forces us to assume that locomotion of adult *Papio anubis* is indifferent to sex, or that the sex difference in adults can be inferred from the observed difference in subadults. More on this to follow.

The *body mass* of the animals was determined for each recording (averages for infant/adolescent/adult individuals: 3.0/10.2/13.1 *kg*). This means that different strides of the same subject would associate with slightly different body masses if body mass changed within the narrow time period of recording. Furthermore, body mass can be associated with the other subject parameters (e.g. trivially with age, because individuals grow). For the subsequent analyses, this interrelation of subject characteristics needs to be taken into account.

The two leg long bone segments (hip-knee-ankle) were chosen as a size reference for normalization of distance values. They measured 0.27/0.41/0.43 *m* (inf./adol./adult) per age class on average. Adding the length of the torso segment (head-hip), we retrieved a total body length proxy, which was used to calculate a “body mass index” (BMI, averages 5.0/11.4/14.1 *kg/m* for the respective age classes). This is no proper scientific BMI, but one solely for demonstration purposes in this document.

To prepare the decision on how to include body mass, leg length and BMI in the subsequent models, behold the first **probabilistic model**. It is a rather simple one, modeling body mass as a function of age class and sex. I will use it to settle some fundamentals about how I apply the modeling procedure herein.

The Body Mass Model

Take the following linear model for body mass, with body mass θ , intercept α , data vectors v , slopes β and residuals ϵ :

$$\theta \sim v_1 \cdot \alpha + v_{male} \cdot \beta_{male} + v_{adol} \cdot \beta_{adol} + v_{inft} \cdot \beta_{inft} + \epsilon \quad (4.2)$$

Note that all vectors v in eqn. (4.2) are boolean column vectors, containing ones for data rows of a given category and zeros on all other rows (v_1 is an all-true boolean specially used for the intercept). This way, the decisive slopes are only added to the data rows which match a given category: for example, only the entries where v_{male} is **True** get the value of β_{male} added.

A brief recap to homogenize vocabulary:

- The “left hand side”, “independent”, or “outcome” variables (often called θ_i) are termed *observables* in this text. Acknowledged, all parameters are “observed” (or “dependent”, or “outcome”...), but observables are the ones in focus of a given chapter.
- The “right hand side”, “dependent”, or “input” variables are usually described as *predictors*. Again, this label is my arbitrarily changing choice.
- Model components which do not stem from the data table are called (*model*) *parameters*; in the case of linear models these are intercepts, slopes and the residual.

I will explain all components in detail in the following paragraphs.

1. **Intercept α :** Each model “starts” with an intercept (though it wouldn’t need to, or the intercept could be implicit). For continuous variables, the intercept is intuitive to understand: it is the observable value when the predictor is zero. With categorical variables, the intercept value is the value observed for a given “reference” set/selection of parameter values. For example, in the body mass model below, the intercept is approximately corresponding to the average body mass of adult, female animals. The choice of reference is done by the statistician, it is an arbitrary one, and sometimes only visible by the category changes associated with slopes.
2. **Slopes β :** Slopes quantify the change in the value of the modeled observable (e.g. body mass) when changing along a parameter axis (e.g. sex from female to male). Some mathematical explanation:
 - For binary categorical values, such as sex, this is simple: set the corresponding value for “female” to 0.0 and that for male to 1.0 and the slope (“female $\rightarrow$ male”, or just “male”) will be their modeled difference in body mass.

- For multi-categorical values (e.g. age class, three possible categories), one reference category is chosen (e.g. adults), and separate slopes are modeled for each of the other groups (i.e. “adlt $\rightarrow$ adol” and “adlt $\rightarrow$ inft”). Because for each slope, only the target category is associated with 1. and all others with 0., the slopes determine the pairwise distance of the target categories and the reference (the category associated with the intercept). No animal is infant and adolescent for a single observation (categories are mutually exclusive), hence no observation will get associated with both slopes.
 - For continuous values (e.g. age), the slope is related to the numerical values in that parameter. Parameter ranges must be considered when evaluating effect size. It usually makes sense to center or even normalize the parameter to make slopes comparable.
3. **Residual ϵ :** Even with all measured parameters, some unexplained variation remains in models of partly random processes. This can be measurement “noise”, or simply unexplained variation. Its order of magnitude is estimated in the sampling procedure as the “residual”. The terms “residual”, “epsilon”, and “standard deviation” (in model context) can be considered approximately synonymous (*:statistician-smiley:*).
4. **Prior Choice and Sampling:** Because I use a probabilistic modeling framework (the python library PyMC), the actual values of the model parameters are determined in MCMC sampling (see Ch. 4.2.2). This procedure requires start values (“priors”), for which I usually take the mean and standard deviation of the observed values (because the sampled model values should certainly fall in that range). Also, there is a whole lot of voodoo about setting the correct distribution type for a given parameter. Usually (with the sample size encountered in this data set), it is hard to falsify the assumption that model values, logarithmized in obvious cases, are approximately normally distributed (Downey, 2013).
- Finally, the MCMC sampling itself asks for some choices by the researcher. I chose something close to industry standard (No U-Turn Sampler “NUTS”; sufficient tuning and sampling steps; twenty chains; Student’s T distribution for the posterior) and verified *ex post* that sampling went well (“hygiene”).
5. **Model Comparison:** I skip this here for clarity. Usually, one would have to argue why the chosen model is the most appropriate one (though of course limited by availability of parameters). More on it when we get to spatiotemporal gait variables. For body mass, the model choice is rather trivial, there is not many options but to use the other subject parameters as predictors. I do not use leg length and BMI as predictors for body mass, because they share a hierarchy level, are correlated, and thus the idea of one predicting the other might falsely imply causality.

-
6. **Data:** Because there is usually only one body mass per subject animal in our data set, the data set onto which I apply the body mass model is reduced (17 observations). For one animal which was measured at two body masses, the average value was used.
 7. **Transformations:** As an additional technical complication, body mass (and the other body proportions) have to be transformed. For most body measures, negative values are implausible, e.g. a negative body mass does not exist. Also, one can usually observe that the Normality assumption, or in our case a Normal prior distribution, is better met when log-transforming the values. In consequence, predictions become more accurate for models of logarithmized measures. This was also the case for the current data set. Therefore, the body parameters presented here were transformed with the natural logarithm for modeling, and all values untransformed with the Euler exponential for presentation.

Body Mass is Age-Dependent (Surprise!)

As expected, body mass and age class are associated (Tab. 4.4). But how much? On average, adolescents are 2.4 *kg* and infants are 9.4 *kg* lighter than adult individuals. The effect of an animal being infant is deemed relevant because the credible interval does not include zero. The credible interval is the smallest possible value range to contain 94 % of the values, i.e. from 3 % to 97 % quantile of the samples; termed “highest (probability) density interval” or HDI. That interval is comparable to the “standard error of the mean” of a given slope, and has to be seen in relation to the intercept and model residual to judge if an effect is large or small, i.e. relevant or not. The difference for the adolescent group exemplifies this: it is in the order of the ϵ , and the HDI contains zero. Although we know that adolescents weigh less than adults, the data set is insufficient and effect size too low to give waterproof statistics. In contrast to age, sex seems to have no effect on body mass (keeping in mind that there are generally few and no adult males in the sample; slope: +0.6 [−1.9, +3.7] *kg*). Finally, the parameter ν (“nu”) deserves some explanation, to demonstrate another neat modeling trick (Wiecki, 2013). An often disputed, but indispensable assumption in

Table 4.4: Results of the body mass model.

parameter	value	interval	cred.
α	+12.3	(+10.2, +14.8)	*
β_{male}	+0.6	(−1.9, +3.7)	
β_{adol}	−2.4	(−4.6, +0.4)	
β_{inft}	−9.4	(−10.1, −8.6)	
ϵ	+2.6	(+1.7, +3.8)	
ν	+52.1	(+15.1, +94.0)	

Asterisk in the “cred.” column indicates model parameters for which the credible interval does not include zero.

conventional statistics is the Normality assumption (i.e. that values of a variable are normally distributed). Likewise, in probabilistic modeling, one can choose the distribution *type* (e.g. Normal) of a model parameter, which together with the distribution *properties* (e.g. $\mu = 0$, $\sigma = 1$) is the **prior** for that parameter. Distribution properties are adjusted in MCMC sampling, and the prior only influences the result when sample size is very small. In contrast, the distribution type stays fixed and has major influence on the outcome. Hence, choosing a Normal here might need some justification. Part of that justification comes from model comparison (see Ch. 4.2.2). But in case of the Normal distribution, we can use a more general distribution (Student's T) to actually measure how “Normal” our data is. Student's T has an extra parameter, which is ν or the “degree of freedom”. The larger ν , the higher the resemblance to a Normal distribution. In contrast, on lower ν values, Student's T has more “weight in the tails”, accounting for “outliers” which are non-normally far from the mean. For our body mass model, the sampler converged at $\nu = 52.1$. This is a relatively high value, confirming that body mass is approximately normally distributed. Usually, model comparison favors the use of a Normal posterior in such cases.

Multi-Level Modeling

For all following models, there is the option to include the body mass parameter as a hierarchical (better: multi-level) parameter. This means that, when we have a model which includes a slope for (centered) body mass, we actually make it three slopes (one for each age class). I will note this as $\beta_{bm|age}$.

As mentioned above, group-wise centering body mass ($cbm_a = bm_a - mean(bm_a) \quad \forall a = ageclasses$) can be useful to facilitate interpretation of the model outcome. If data values are centered, the slopes are relative to a mean, and the intercept will give a value at the average of all continuous parameters (in this case: at average body weight). And because age has such a high impact on body weight, it makes sense to center the groups individually, to get relatively lighter or heavier individuals per age class separated.

The complexity of this matter is not to be underestimated. Just to review our options to model body mass:

- inclusion or exclusion
- log-transform or not
- not centering or centering to population mean or to age class mean
- multi-level: one slope for the whole population or slope for each of the age classes

All of these choices in combination affect the model, and need to be compared. Even worse: they affect how to best model other parameters (if `bm|age` is used, `leglength|age` might

be a worse choice than `leglength` alone, but if `bm` is a population slope, `leglength|age` might become the model of choice). Multi-level modeling gives a lot of options for model comparison. These will be handled below.

Multivariate Posterior Structures: Mor(e)phometrics

Together with body mass, I modeled leg length (ankle-knee-hip distance, accumulated) and body mass index (BMI, body mass divided by ankle-knee-hip-head distance). The model formula from above remains unchanged, except for the left side. But that, namely “multiple parameters on the left side”, is a concept on its own. The solution is a **Multivariate Posterior**. I imagine this as multiple models, sampled in parallel, but cross-connected through the data and the correlations (like a ladder with steps of chewing gum, floating in free space... yes, I’m running out of metaphors). In an unconnected model, the MCMC sampler would explore the parameter space randomly, and might be “at different ends” for two variables at a given time. In contrast, when they are connected by a multivariate posterior, the sampler path is also connected, and therefore a positive correlation of two parameters reflects in the values the sampler sees simultaneously. The connection can be either fixed (empirical correlation, based on the data) or inferred. The correlation is crucial for predictive sampling to avoid implausible parameter combinations.

Let us rehearse this concept on the body proportions, where it is rather intuitive. Body proportions (including body mass) are highly correlated. For BMI, this is trivial: it is defined as the quotient of two other measures.

We can simply calculate the correlation of the three present observables from the raw data. This would be the *empirical* correlation. However, according to our statistical theory, the data we acquired is just a sample from the true, underlying distributions of the observables. So the correlation might be incomplete or inexact. Probabilistic statistics is the attempt to infer that true underlying distribution, and when it is known, one can calculate a potentially more exact correlation (*inferred* correlation). The inferred correlation is not *per se* more exact; imagine a situation where correlation is close to ± 1 and the sampler might have problem with parameter bounds. It depends on the data which one works best (yay, more model comparison!).

In my toolbox of choice, `PyMC`, a multivariate posterior can be initialized with either (i) the empirical or (ii) the inferred correlation structure; a third choice is to use (iii) no correlation. For cases (i) and (ii), I introduce the observables as a multivariate Student’s T distribution. The empirical correlation (i) enters the model in form of the Cholesky matrix calculated from the data³. To get inferred correlation (ii), I instead initialize the `MvStudentT` with a so-called LKJ Prior (Lewandowski et al., 2009), which allows me to quantify the posterior correlation. Finally, uncorrelated posterior (iii) do not need an `MvStudentT`, but go with

³Subtle technical complication: this variant does not allow the inclusion of an explicit model residual; in such cases I report the standard deviation of the observables as model residuals.

independent `StudentT`'s; the system can sample multiple observables at once.

Note that multivariate blocks will also be relevant on the predictor side. The only difference between observables and predictors in modeling is the somewhat arbitrary choice of which variables we define as “dependent”, and which as “independent”. After all, they are all just distributions, as probabilistic modeling is about juggling distributions. Tricky detail: a multivariate predictor block must be initialized with inferred correlation (LKJ), because this block contains inferred slopes and no observed data, and thus the correlation is that of the slopes, and not of the data.

In the case of our group of body proportions, the empirical correlation is close to unity (Tab. 4.5). It is less strong with the inferred posterior correlation (Tab. 4.6): body mass is highly correlated with both leg length and BMI, but the latter are not correlated to each other.

The multivariate model also infers the relation to the primary subject parameters (sex and age class, *cf.* Tab. 4.7). The values for body mass are identical to those reported above (Tab. 4.4), but here the outcome is shown for the log-transformed observables. The results for the additional quantities are consistent with those from body mass, except that we see that leg length of adolescents is almost identical to that of adults (whereas in body mass, the difference is more pronounced).

And there are many more body size measures (broad category: morphometrics) which could be measured and included. (There is *always* another parameter which one could include in the model.) BMI was intentionally introduced as a redundant parameter and will not be included in subsequent models. Body mass and leg length are highly correlated, and if both are part of another model (e.g. a model on coordination), it might be hard to distinguish which of them is causally responsible for an observed effect. On the other hand, I demonstrated that body mass and leg length are slightly different with regard to the adolescent individuals. Hence, the question whether or not to include either, both, or none of these two parameters in the more complex models to come remains to be evaluated.

Table 4.5: *Body proportions: empirical cross correlation.* Asterisk indicates parameter pairs in which the Pearson Correlation p -value is less than 0.05.

	log(body mass (kg))	log(leg length (m))	log (bmi (kg/m))
log (body mass (kg))	1	+0.97 *	+1.00 *
log (leg length (m))	+0.97 *	1	+0.94 *
log (bmi (kg/m))	+1.00 *	+0.94 *	1

Table 4.6: *Body proportions: inferred cross correlation.* Asterisk indicates values for which the HDI does not include zero.

	log (body mass (kg))	log (leg length (m))	log (bmi (kg/m))
log (body mass (kg))	1	+0.49 *	+0.92 *
log (leg length (m))	+0.49 *	1	+0.18
log (bmi (kg/m))	+0.92 *	+0.18	1

Table 4.7: *Body proportions: model results.* Note that these values refer to log-transformed body proportions.

	intercept	female → male	adult → adolescent	adult → infant	ϵ
log (body mass (kg))	+2.51	+0.05	−0.21	−1.44 *	±0.19
log (leg length (m))	−0.87	+0.03	−0.03	−0.46 *	±0.08
log (bmi (kg/m))	+2.60	+0.02	−0.18	−1.01 *	±0.16

Posterior Predictive Sampling

Because the body proportion problem involves still relatively few parameters, a rather complete visualization is possible (Fig. 4.1). Some observations:

- Comparing the raw data to the predicted values confirms plausibility of the predictions.
- The effect of modeling in log space is visible in the narrower distribution of groups with lower values.
- Groups with fewer observations (e.g. adults) tend to be wider, reflecting the influence of relatively wide priors and a high uncertainty in these classes.
- Out-of-sample prediction is possible: adult males were not observed, but their model values get tuned by the observations of partly similar subsets of data (i.e. other sex/age combinations).

The visual comparison captures all relevant info in this simple example case. For subsequent models, such plots are equally valuable, yet the occurrence of continuous predictors and the increase in complexity require a lot more filtering and adjustment.

Summary: Subject Parameters

The example of a body mass model has already taken us deep into the modeling world. We explored simple relations (body mass as a function of age class and sex, and the body proportion cross correlations). This involved a linear model with rather few components. It is a trivial finding that body mass depends on the age class of our subject, and we will see how that can be incorporated in subsequent models. With regard to the interrelations of our body proportion quantities, we encountered multivariate posteriors and saw how they can serve to infer correlations among the observables. Visual comparison of raw data and model predictions confirmed that the model converged to plausible results.

All this was more a playground to rehearse and apply some basic modeling concepts, establish vocabulary and visuals, and to prepare some upcoming design decisions when moving on to other quantities of relevance.

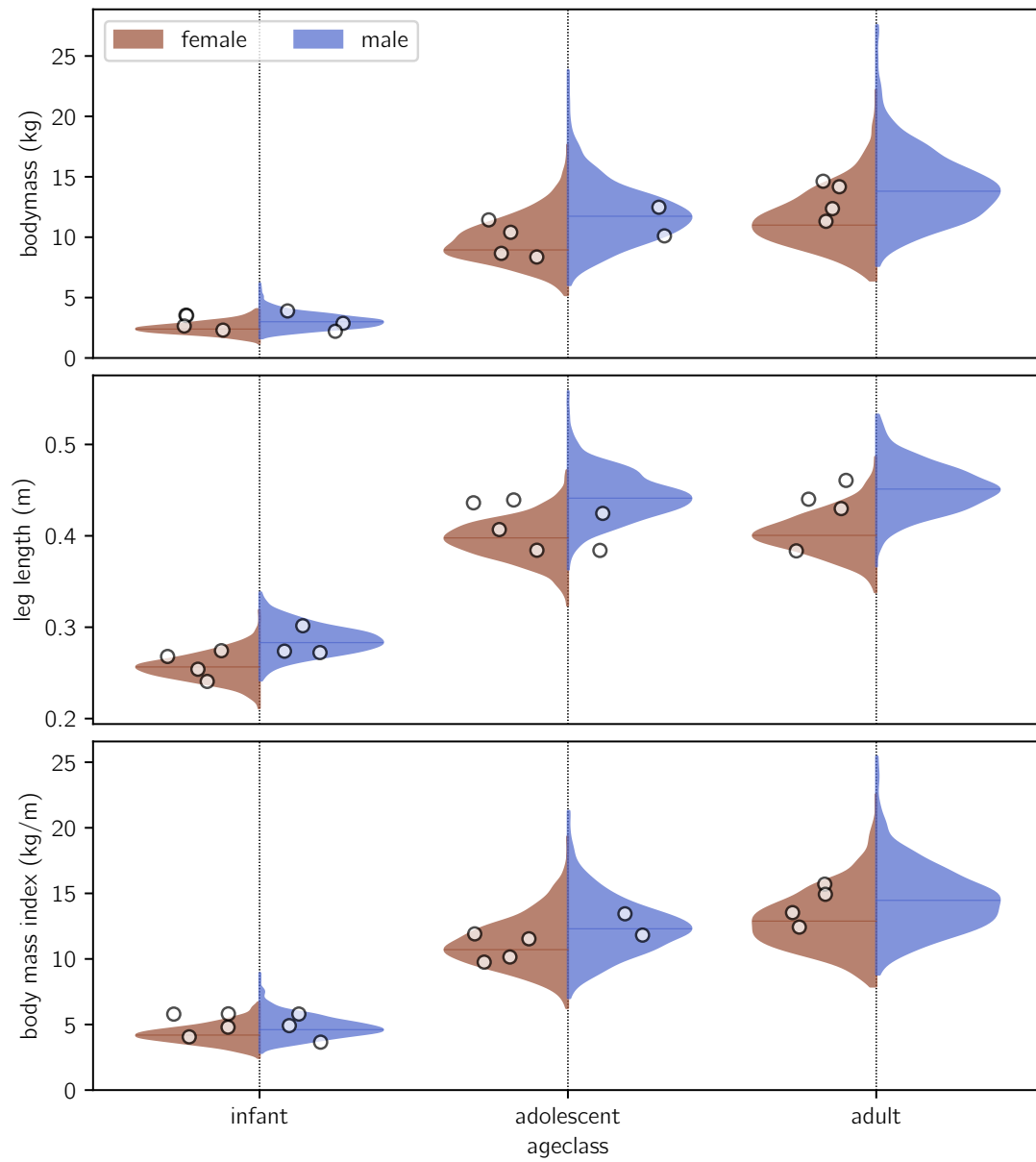


Figure 4.1: Subject parameters: raw data (white circles) and model predictions (distributions).

4.3.3 Spatiotemporal Gait Variables

With this knowledge about the subject parameters, I will proceed to the next parameter category: the spatiotemporal gait variables, also called gait variables, spatiotemporals, or stride cycle parameters (Tab. 4.8). Those are:

- duty factor
- diml. stride distance (relative to lower leg length)
- diml. stride frequency
- diml. speed
- clearance
- trunk angle

The “diml.” indication refers to “dimensionless”, which means that the spatiotemporal parameter can be normalized, according to the concept of dynamic similarity (see Ch. 4.3.1; Alexander and Jayes, 1983; Hof, 1996). Just as body mass and its relation to the other subject parameters was the focus of the previous chapter, we now move spatiotemporal gait variables to the “left hand side” of the equation, i.e. they are observables, and we guess that they will be in part predicted by subject parameters.

Model Design

As before, some design decisions should be justified.

1. **Parameter Correlation:** The spatiotemporal gait variables are intrinsically correlated. This can have trivial reasons, such as their definition (e.g. speed is the product of distance and frequency). In other cases, it is at least physically and physiologically plausible (e.g. lower duty factor at higher speeds). These correlations must be quantified, which can be part of the sampling procedure. As with moving from body mass to a set of body proportions, we will have a connected set of observables. This can also be incorporated in the mathematical descriptions.

The favored model in this case was one which got programmed on the *empirical* correlation of spatiotemporal gait variables (Tab. 4.9). The speed parameter correlation can elucidate how animals reach higher speeds: they walk at increased stride frequency and cover a larger distance, whereby the duty factor tends to reduce. This, however, implies no consistent change in clearance, nor does it depend on mean trunk angle.

-
2. **Reference Model:** I will model the spatiotemporal gait variables as a function of sex, age and body mass, i.e. all the relevant subject parameters. The goal is to find out how strides of a given subject are characterized, e.g. whether certain aspects of the subject lead to difference in their gait characteristics.

There are two basic strategies to identify the final model. As mentioned before, there are logical reasons for a model choice, and quantitative comparison. Choosing an inclusive strategy (i), one might start off with an “all in” model. From there, one can iteratively leave out parameters, perform model comparison, and see whether models with fewer parameters are favored by information criteria. Alternatively, starting from a minimal model (ii), one can gradually increase complexity and find the point when adding parameters does not indicate improvement. Either way, the procedure is iterative: sample all plausible, possible models, and perform model comparison (Ch. 4.3.3). Ideally, both strategies will converge on “the best model”. This model is simply the one which, of a broad range of models, had favorable model scores, and is logically consistent. It is labeled *ex post* as the reference model.

In the case of spatiotemporal gait variables, the reference looks as follows. The equations below extend the θ from (4.2) to be a multivariate data “block” (i.e. matrix). It has the number of observations in rows, and the number of observables in columns. Similarly, data vectors v , intercept α , slopes β and residual ϵ receive an extra dimension and become matrices (which means incredibly fuzzy work for model programming, but can be considered bean counting given that we may turn to the actual outcome one variable at a time).

Here is thus the formula for one gait variable:

$$\theta_i \sim v_1 \cdot \alpha_i + v_{male} \cdot \beta_{male,i} + v_{adol} \cdot \beta_{adol,i} + v_{inft} \cdot \beta_{inft,i} + v_{cbm} \cdot \beta_{cbm,i} + \epsilon_i \quad (4.3)$$

Therein, i is one of the spatiotemporal gait variables, which are all sampled in a multivariate model. The intercept vector v_1 is simply a column vector of ones. The boolean v_{male} holds values one for every stride for a male subject, and zero for females. Similarly for v_{adol} and v_{inft} , which are mutually exclusive in the equation (i.e. animal is either adolescent, or infant, or neither, but never two of those options). In contrast, v_{cbm} for centered log body mass contains continuous values; the centering was done per age class, and after log transformation. Group-wise centering affects the data vector, however the body mass slope is not affected by age class (a model with $\beta_{bm|ageclass}$ was also included in model comparison).

3. **Prior Choice:** A Normal distribution was used for all the slopes, a multivariate Student’s T for the observables. Informative priors (global mean and standard deviation) are used, which is a good heuristic to improve sampling without biasing

the sampler. The multivariate block of observables is initialized with the empirical correlation, which facilitates the work of the sampler (yet all three options of multivariate posterior were applied and compared).

Subject Parameters Affect Stride Characteristics

The model of spatiotemporal gait variables quantifies how stride (or gait) characteristics in bipedal baboon walking are affected by subject parameters (Tab. 4.10).

Clearance is highest in adult males of low body mass (though we did not actually observe those; the effects are separately determined for age class, sex, and body mass). Females and subadults lift their feet less during the swing, as do animals of comparatively higher body mass. Duty factor is higher in infants: they are in ground contact for a longer portion of the stride. All of the three speed-related parameters (i.e. distance, frequency, speed; all dimensionless) are affected by body mass, though only for frequency the effect is clearly different from zero. Animals of higher body mass tend to take shorter steps of higher frequency, which seems to overcompensate in terms of speed. Infants cover lower relative distance at lower frequency, which adds up in terms of relative speed. The results on trunk angle show that males hold their upper body less upright than females (i.e. at a higher trunk angle; though as with any sex effect this might be due to few, influential observations). Although this inference can only stem from the subadult classes, it is separate from (and in opposing direction of) an age effect: adults seem to walk more crouched than infants (i.e. infants walk with a lower trunk angle, more upright).

A value which is sampled, but not shown, is the ν parameter (or degrees of freedom, see Ch. 4.3.2) of the Student's T posterior. It is high ($\nu = 54.2$), indicating that the posterior distribution has hardly more "weight in the tails" than a Normal distribution would; the Normality assumption is a good approximation in this case.

Gait Variable Two Step Prediction

The posterior prediction of spatiotemporal gait variables holds an additional complication: as shown in the body proportion model, we now have interdependent predictors (i.e. body mass might depend on age and sex, though note that it is group-wise centered). Therefore, two consecutive steps of posterior prediction are necessary: first predict a number (e.g. $n = 1000$) of body masses, then based on those samples predict spatiotemporal gait variables. For a given setting for sex and age class, the workflow is as follows.

1. draw posterior predictive samples from the body proportion model, to get n body mass samples
2. convert these samples (i.e. log transform, age class-wise centering with the means from actual data)

Table 4.8: Observed spatiotemporal gait parameter value ranges. *stdev* = standard deviation.

gait variable	units	mean	stdev.
clearance	%	0.35	0.05
stride distance	<i>m</i>	0.57	0.16
stride duration	<i>s</i>	0.56	0.14
duty factor	1	0.63	0.07
stride frequency	<i>Hz</i>	1.92	0.55
speed	$\frac{m}{s}$	1.10	0.41
trunk angle	<i>rad</i>	0.57	0.18

Table 4.9: Empirical Cross Correlation of Spatiotemporal Gait Variables. Significant correlations (as per Pearson’s correlation coefficient and test for non-correlation) are marked by asterisks. d./d.s.: “dimensionless (stride)” parameter.

	clearance	d.s. distance	duty factor	d.s. frequency	d. speed	trunk angle
clearance	1	+0.32 *	−0.42 *	−0.01	+0.15	+0.38 *
d.s. distance	+0.32 *	1	−0.57 *	+0.09	+0.52 *	−0.12
duty factor	−0.42 *	−0.57 *	1	−0.52 *	−0.72 *	−0.27
d.s. frequency	−0.01	+0.09	−0.52 *	1	+0.89 *	+0.12
d. speed	+0.15	+0.52 *	−0.72 *	+0.89 *	1	+0.07
trunk angle	+0.38 *	−0.12	−0.27	+0.12	+0.07	1

Table 4.10: Effects of Subject Characteristics on Spatiotemporal Gait Variables. Values which do not contain zero in their HDI are indicated by an asterisk.

	intercept	female → male	adult → adol.	adult → infant	c.log. body mass	ϵ
clearance	+0.39	+0.05 *	−0.05 *	−0.06 *	−0.12 *	±0.05
d.s. distance	+1.72	+0.16	−0.06	−0.13	−0.38	±0.25
duty factor	+0.58	−0.05	+0.01	+0.10 *	−0.03	±0.07
d.s. frequency	+0.39	−0.02	−0.02	−0.06	+0.21 *	±0.10
d.s. speed	+0.68	+0.04	−0.06	−0.16 *	+0.21	±0.20
trunk angle	+0.67	+0.14 *	−0.12	−0.20 *	+0.10	±0.18

3. replace the data in the spatiotemporal gait variable model with these samples
4. then draw (i.e. predict) n spatiotemporal gait variable samples.

Step 3 in this workflow might seem odd, but it is exactly the same procedure we get when choosing a setting for age and sex. The only difference is that we do not set n boolean values, but n values which are drawn from a continuous distribution. Although one should in theory sample multiple spatiotemporal gait variable values for each sample of body mass, it is sufficient (and technically convenient) to keep the n constant throughout the procedure and compensate by a large number of samples.

As with the subject parameters, we can compare the observed values with the predictions of the model (Figs. 4.2 and 4.3). The data and predictions are overall consistent with the effects identified by the model (Tab. 4.10), and it is good sport to reason how the prediction for male adults came about.

Stride PCA in Posterior Prediction

Due to the intrinsic correlation of stride distance, stride frequency, and speed, it is an option to submit them to a PCA for optional dimensionality reduction. In posterior predictive sampling, the same PCA must be used to transform the samples to PCA space.

Model Comparison

One of the great features of quantitative models is the ability to perform model comparison. An arsenal of “information criteria” is available to compare models and choose which one succeeds best in capturing the effects present in the data, without being over-complex.

In addition to the reference model above, many other models were sampled for comparison. Those are similar to the reference, but individual parameters are removed in turn, or their hierarchical structure is altered (in the case of body mass). In other test models, parameters were added (e.g. leg length). The centering of body mass was optionally omitted. Additionally, the correlation structure of the observables (“empirical” per default) can be changed to “inferred”.

The outcome of this particular comparison method (Tab. 4.11) are “LOO” values (“Leave One Out”) as a quantification for the model performance/complexity trade-off. As mentioned, there are alternative quantifiers (e.g. WAIC), but experience shows that rankings from different information criteria are often similar. Given relatively little difference and wide standard error ranges in the LOO estimate, we can conclude that almost all of the tested models yield some plausible descriptions of the data (in other words: most models do not perform clearly better or worse than the reference). In fact, the reference model is not the highest on the list: for example, choosing a Normal distribution instead of a Student’s T for the observables would be better, which is consistent with the

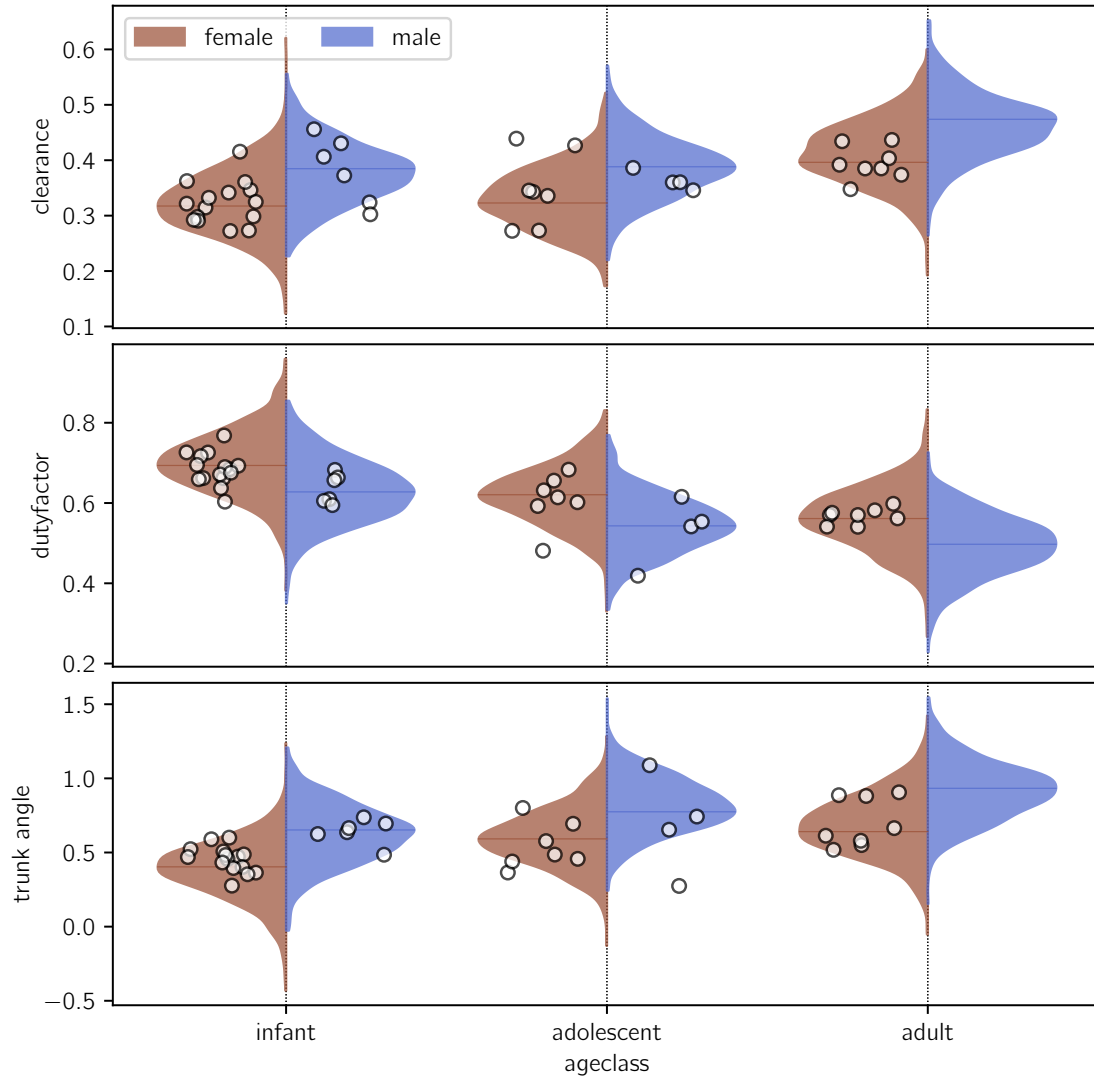


Figure 4.2: Spatiotemporal gait variables: raw data and model predictions. Part 1.

Table 4.11: Model comparison. LOO is the model goodness indicator. For convenience, delta (LOO difference from highest ranking model) and standard error (range of the LOO estimate) are displayed.

rank	loo	delta	std err	description
0	286.5	0.0	9.7	additional predictor- leg length
1	286.4	0.0	9.8	additional predictor- BMI
2	285.9	0.6	10.1	Normal distribution in posterior
3	285.7	0.7	9.6	body mass/age class level
4	285.6	0.9	10.1	reference model
5	284.9	1.5	9.8	non-centered body mass
6	284.6	1.9	9.8	no sex slope modeled
7	280.3	6.2	10.1	body mass excluded entirely
8	280.1	6.4	10.4	age class slopes excluded
9	276.9	9.6	16.8	sampled observable covariance
10	208.4	78.1	11.8	spatiotemporals in SI units

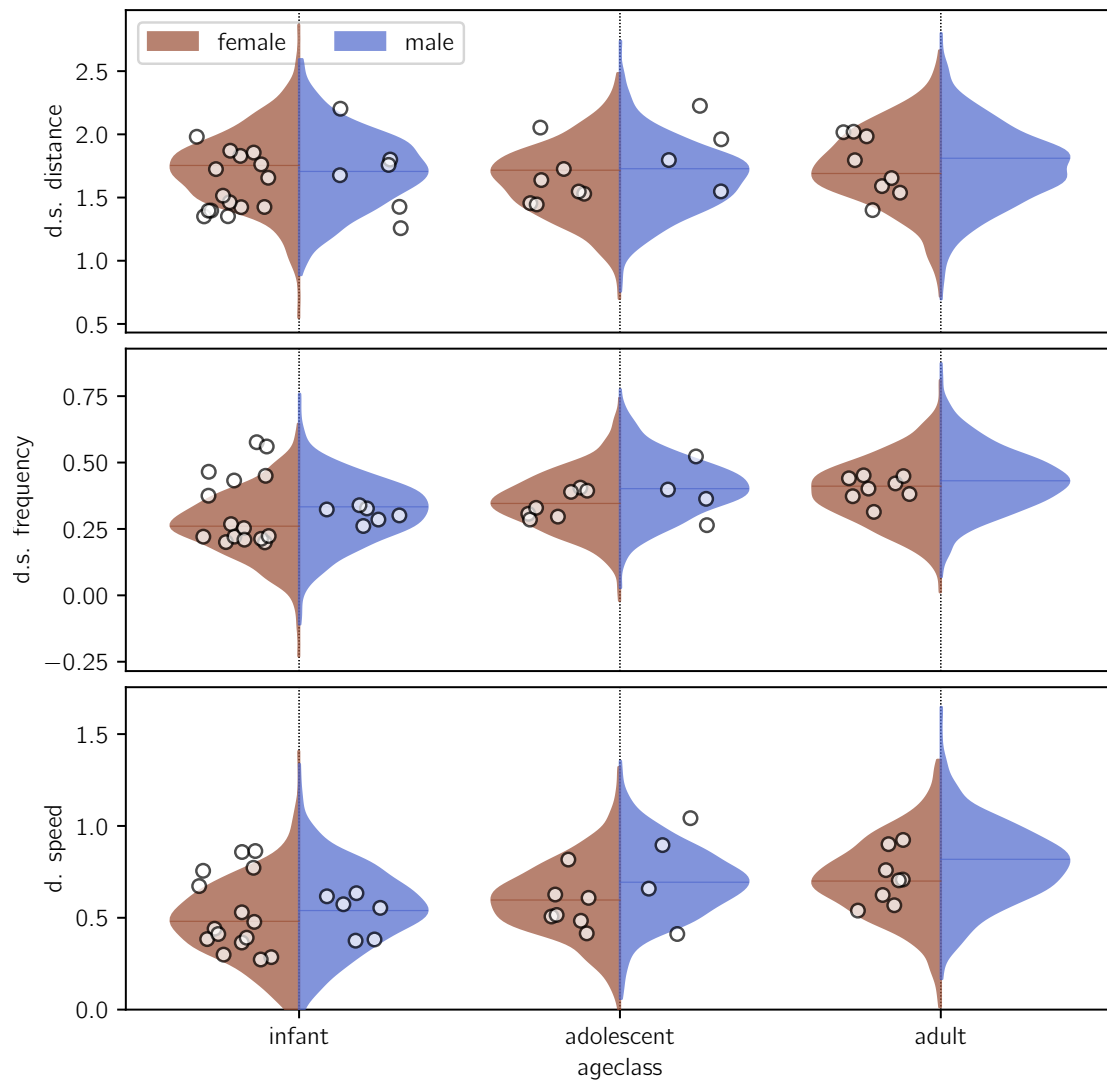


Figure 4.3: *Spatiotemporal gait variables: raw data and model predictions. Part 2.*

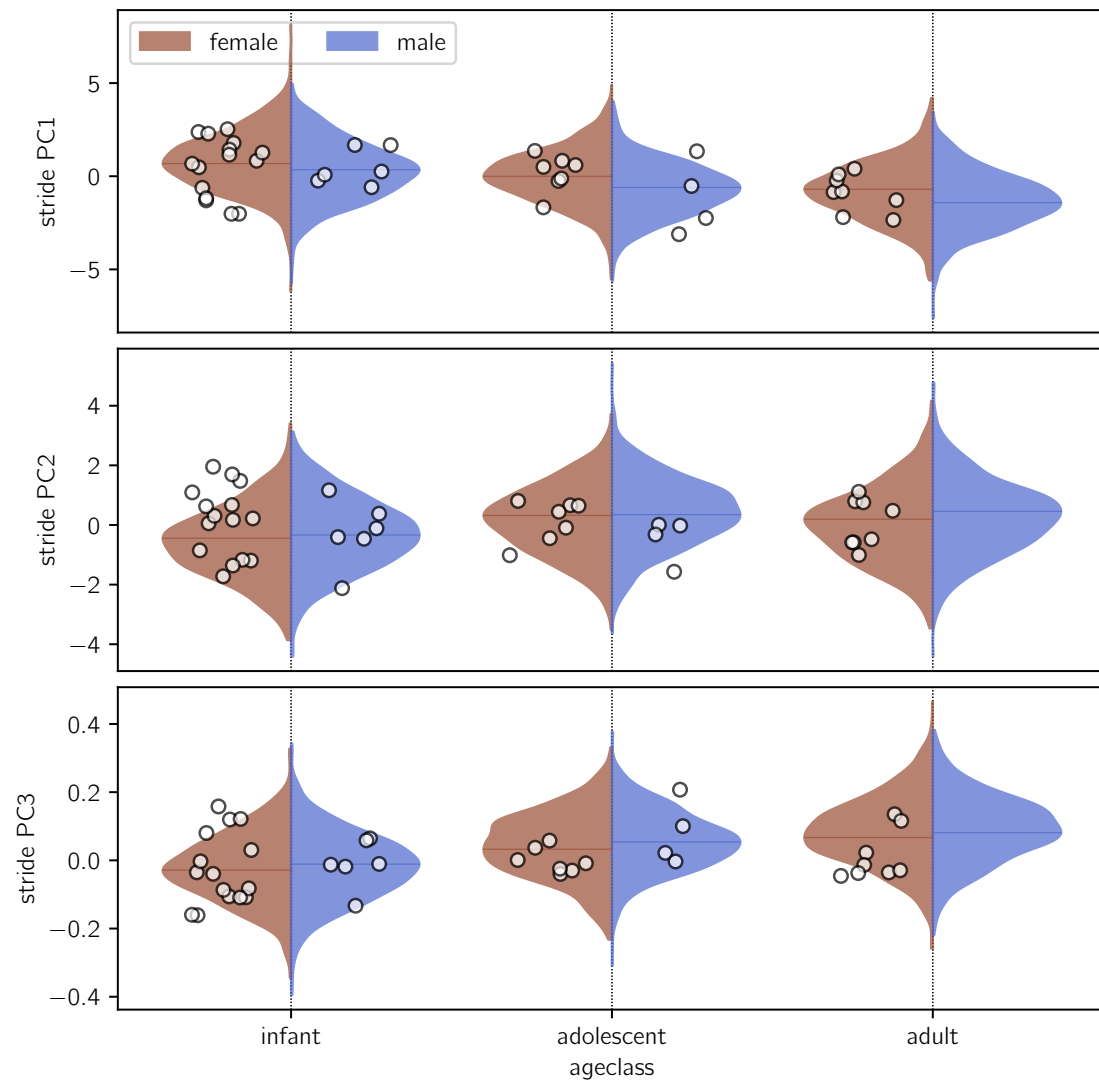


Figure 4.4: *Spatiotemporal gait variables: raw data and model predictions. PCA conversion.*

inferred high value for ν . This switch in this case would be little more than cosmetics, since that extra parameter does not affect the results presented above.

Among the models ranking lower than the reference are models leaving out either of the current predictors. Most of them are still within the range of the reference LOO, so it would be justifiable by logical arguments to choose either of them as reference. The only outlier is the model with “spatiotemporals in SI units”, i.e. where distance, frequency, and speed are measured in absolute values. Those values are less well explained by a linear model with the given priors and structure. This does not mean that model is wrong: most likely, they are just less normally distributed. Yet for the sake of constructing a cascade of models (from subject-, via stride-, to coordination parameters), it seems better to choose the dimensionless parameter set.

On the other hand, comparison would favor an age class dependent (hierarchical) body mass, or the inclusion of another body proportion proxy (leg length or BMI). However, this ranking appeared only after I made a final, last-minute (and presumably minor) adjustment to the data. Any extra data could tip the balance back towards the current structure - or not. This exemplifies the finite sharpness of model comparison. It guides model design, and some designs can be clearly excluded (falsification!). Yet it should by no means be interpreted as carved in stone ultimate wisdom. A problematic corollary of this is the need to re-sample a high number of models after even the tiniest change in the data. Combined with the long duration of MCMC sampling, this can be prohibitive, but partially alleviated with an efficient sampling toolbox (see Ch. 4.2.2).

Hierarchical Components (“Random” Effects)

Above, I have touched on the concept of a hierarchical parameter ($bm \mid ageclass$), and in fact it might be chosen here over the current reference. Other hierarchical structures are possible and should be included if that is indicated by model comparison.

A particularly popular one is a “random intercept”, where “random” means that each subject has its own, specific value for and intercept. In the case of spatiotemporal gait variables, such a construct would be possible if we had longitudinal sampling (e.g. individuals recorded at different ages, or different body weights). However, given that the set of subjects is partitioned (e.g. into male and female), a model structure with a subject-level intercept might shift variance to that “random” intercept which would otherwise be captured on the sex slope. Such ambiguities in effect distribution sometimes manifest in sampling failure (because the sampler can freely move values between the overlapping categories, and thereby fails to converge). Other times ambiguous models get disfavored in model comparison due to the extra complexity. And yet on other examples in-sample prediction produces weird results (e.g. unlikely parameter combinations).

In conclusion, due to the data structure, I would argue that a subject level intercept is not useful in the case of spatiotemporal gait variable model. An attempt to include such a model

for model comparison failed accordingly during MCMC sampling, which is why it is not even in the model comparison list. Note that this sampling problem would not become apparent in least squares modeling.

Summary: Spatiotemporal Gait Variables

The presented findings on spatiotemporal gait variables are substantial insights into the way in which spatiotemporal characteristics of gait in *Papio anubis* are affected by sex, age class and body mass. These findings are interesting by themselves and were part of the prior research on the data set (Druelle et al., 2021). They are re-iterated here with a different modeling framework and serve to exemplify important statistical concepts, most importantly (two-step) posterior predictive sampling and model comparison.

4.3.4 Posture and Coordination

The remaining classes of model parameters (Tab. 4.2, Ch. 4.3.1) are posture and coordination. In this chapter, I will briefly describe how they are acquired, before even briefer presenting the results of probabilistic models which quantify their dependence on subject and spatiotemporal gait variables.

This chapter does not introduce any new concepts of statistics, but instead has some nice graphs and figures. Its purpose is to round off the baboon story and provide the basis for further discussion.

Quantifying Posture and Coordination

Among the most relevant outcomes of video-derived data and kinematics are joint angles. Their temporal profiles, i.e. the joint angle over time, and how they change for different settings or subjects, can yield insights on coordination and motor control. In cases where we have steady state locomotion, one can exploit the fact that joint angle profiles are cyclic and submit them to a Fourier Series Decomposition (*cf.* Mielke et al., 2019). This operation is simply a reversible transformation, and no information is lost in the process. Temporal profiles can thereby be investigated in the frequency domain, which involves a complex (math.) quantification of a given number of harmonics (this is often called the “spectrum” of a signal). This is perfectly analogous to harmonics on a guitar string, and because one can tune models of Fourier Coefficients with MCMC sampling (see above), guitars are now officially my favorite metaphor in explaining statistics.

Some characteristics are particularly prominent in the frequency domain/spectrum:

- The temporal *mean* value of a joint angle appears as the zero’t Fourier coefficient (i.e. where the string is “hung up”).
- The magnitude of the coefficient values corresponds to the *amplitude* (i.e. loudness, per coefficient or accumulated).
- Fourier coefficients rotate in the complex plane depending on their *phase* (rarely perceived in music, except in directional hearing).

Mean, amplitude and phase are familiar from general theories and descriptions of oscillations. They can be classified as **affine** components of a signal, which simply means that they can be altered by multiplication with a scalar. (Maybe not) coincidentally, they also have a biological meaning. The mean of all joint angles defines posture (*sensu stricto*, it is “dynamic” posture). Think of school children walking in a duck walk (or try it yourself, best immediately, but with vocalization): their limbs will be far from straightened, and their joints set at high angles. Amplitude corresponds to effective range of motion (eROM): when you try to walk over ice with some care to not slip, it will most likely decrease. Herein, I chose

to subsume mean and amplitude as “posture” (*sensu lato*). Phase quantifies the relative timing of oscillating parts with respect to each other, which is, if you like, a “regular” part of coordination. Lastly, if mean, amplitude, and phase are subtracted or normalized, the (non-affine) remainder of the signal can be described as an “irregular” part of coordination. This removal of affine signal components is achieved by the method of Fourier Coefficient Affine Superimposition (FCAS, Ch. 3). More precisely, the purpose of FCAS is to separate mean, amplitude, phase and the non-affine remainder, and use them in reasonable ways. In the present analysis, I used it to disassemble hip, knee, and ankle joint angle profiles and extract **posture** (mean joint angle and range of motion) and **coordination** (the remainder, including phase differences). Phase was not isolated for the different joints; however, a “global phase alignment” was performed by temporally aligning all joint angle profiles so that the phase difference of the *whole limb* (head-hip-toe angle) is zero. I find this mathematically more convenient than the conventional alignment to toe touch down, which in my opinion over-emphasizes a single time point and has problems with all-too-variable duty factors.

A minor practical complication is that there are immensely many parameters in the “coordination” category (number of harmonics retained times number of joints times two for real and imaginary part). The often misused bread-and-butter method for dimensionality reduction is PCA. Hence, I submit the coordination parameters to a PCA for modeling. The first 8 components capture 91.0% of the variance and were submitted to the modeling procedure.

To summarize: by neat mathematical tricks, we are able to extract direct quantitative representations of posture and coordination for each stride.

Model Design

Posture and coordination are technically similar for the model procedure. However, they depend differently on the available predictors. Furthermore, they differ in posterior design: because coordination observables were derived from a PCA (which is an orthogonal transformation), these parameters are non-correlated per definition - no multivariate posterior necessary. In contrast, the posture parameters correlate in a characteristic way (Tab. 4.12), and the model with a sampled covariance structure was favored by model selection. Therefore, posture and coordination were handled by two separate models with similar general structure.

In addition to the predictors used to infer spatiotemporal gait variables, the models now include the gait variables themselves on the right hand side. Since the number of observations has not changed, but the number of predictors increases, modeling gets challenging. As it turns out, a multivariate block of gait variables was too complex to be included, and consequently sampling was unsuccessful. At the same time, interdependence of speed-related parameters should not be disregarded. I solved both problems by using another PCA, in this case for the gait variables (see Ch. 4.3.1). Although this reduces the number of predictors

by just one, it enables successful sampling.

The reference model for posture and coordination is as follows:

$$\begin{aligned}
 \theta_i \sim & v_{1,i} \cdot \alpha_i + \\
 & + v_{male} \cdot \beta_{male,i} + v_{adol} \cdot \beta_{adol,i} + v_{inf} \cdot \beta_{inf,i} + v_{cbm} \cdot \beta_{cbm,i} + \\
 & + v_{clr} \cdot \beta_{clr,i} + v_{df} \cdot \beta_{df,i} + v_{trnk} \cdot \beta_{trnk,i} + v_{str1} \cdot \beta_{str1,i} + v_{str2} \cdot \beta_{str2,i} + \\
 & + \epsilon_i
 \end{aligned} \tag{4.4}$$

In equation (4.4), the lines correspond to the four predictor parts (intercept, subject parameters, spatiotemporal gait variables, residual). Abbreviations are as above, plus **clr**: clearance, **df**: duty factor, **trnk**: trunk angle, **str1** / **str2**: stride PCs.

Implications from model comparison are overall similar to those in the spatiotemporal gait variables. Some exceptions are better-ranked models with left-out predictors, which however were not used because I prefer to keep parameter selection and basic structure consistent throughout the chain of models. Interesting results are anticipated by the comparison: for example, the trunk angle, which per definition contributes “half” of the hip joint angle profiles (adjacent segment), is an integral part of the posture model but not relevant for coordination.

Table 4.12: Sampled cross correlation of posture parameters. Values of which the HDI excludes zero are marked by asterisks. “Mean” refers to the temporal mean of the joint angle profile, whereas “amplitude” is related to range of motion at that joint.

	hip mean	hip amplitude	knee mean	knee amplitude	ankle mean	ankle amplitude
hip mean angle	1	+0.53 *	−0.57 *	−0.07	+0.07	−0.26
hip amplitude	+0.53 *	1	−0.12	−0.30	−0.08	−0.01
knee mean angle	−0.57 *	−0.12	1	−0.11	−0.38 *	+0.25
knee amplitude	−0.07	−0.30	−0.11	1	+0.03	+0.05
ankle mean angle	+0.07	−0.08	−0.38 *	+0.03	1	−0.31 *
ankle amplitude	−0.26	−0.01	+0.25	+0.05	−0.31 *	1

Table 4.13: Model comparison of posture models.

rank	loo	delta	std err	description
0	235.0	0.0	14.4	age class slopes excluded
1	231.8	3.3	13.8	no sex slope modeled
2	231.3	3.8	14.0	additional predictor- leg length
3	229.9	5.1	14.8	reference model
4	229.9	5.1	14.4	Normal distribution in posterior
5	229.4	5.6	13.3	body mass excluded entirely
6	229.3	5.7	14.8	body mass/age class level
7	229.1	5.9	14.9	additional predictor- BMI
8	228.2	6.8	15.0	duty factor slope excluded
9	226.7	8.3	14.8	non-centered body mass
10	224.8	10.2	14.3	stride PC1 slope excluded
11	222.7	12.3	15.2	clearance slope excluded
12	212.5	22.5	15.1	stride PC2 slope excluded
13	196.3	38.8	8.6	fixed observable covariance
14	187.8	47.2	14.9	trunk angle slope excluded

Table 4.14: *Model comparison of coordination models.*

rank	loo	delta	std err	description
0	337.3	0.0	15.3	age class slopes excluded
1	332.8	4.5	14.3	body mass excluded entirely
2	332.0	5.3	14.8	additional predictor- leg length
3	330.7	6.6	14.7	additional predictor- BMI
4	330.6	6.7	14.8	reference model
5	330.6	6.7	14.7	Normal distribution in posterior
6	330.2	7.1	15.2	no sex slope modeled
7	329.7	7.7	14.8	trunk angle slope excluded
8	329.4	7.9	14.2	stride PC2 slope excluded
9	329.2	8.1	14.5	non-centered body mass
10	328.5	8.8	15.2	stride PC1 slope excluded
11	327.3	10.0	14.6	body mass/age class level
12	327.3	10.0	14.8	clearance slope excluded
13	326.6	10.8	14.9	duty factor slope excluded

Lots of Figures

The “statistics tutorial” part of this journey ends here, where the biological discussion begins. Following below is a series of figures and tables which I will leave uncommented, to open a discussion on baboon kinematics, if you like.

Whereas posture parameters have an intuitive meaning (mean joint angles and eROM), it seems futile to interpret the outcome of the coordination model (Tab. 4.16, Figs. 4.7 and 4.8) directly. The observables are principal components of Fourier coefficient non-affine residuals – even Fourier himself could not make sense of such numbers. However, it should be pointed out that PCs are ordered by the variance they cover, hence the first ones capture the larger effects. Furthermore, the edges of the coordination space can be sampled and re-transformed to joint angle profiles, which I demonstrated elsewhere (e.g. Fig. 3.3 in Ch. 3).

The role of duty factor deserves special attention to illustrate what level of detail is contained in these figures. Although zero difference is not excluded by the HDI (probably I should have log transformed it, or given it a Beta distribution), some duty factor slopes are higher than others (Tab. 4.15). We can learn from this that hip and knee angle are affected (hip: more flexed, knee: more extended on average) and that the knee range of motion increases with duty factor. Duty factor also has the biggest effect on coordination, being the highest slope value on PC1 (Tab. 4.16).

In combination, posture and coordination values can be translated back into joint angle profiles. This is especially interesting for prediction (e.g. Figs. 4.5 - 4.8). The final cherry on this cake (Fig. 4.9) is thus a promising attempt to predict joint angular profiles for the given classes.

Table 4.15: Results of the posture model. Values different from zero (per HDI) are indicated by an asterisk.

	intercept	female → male	adlt → adol	adlt → inf	c.log.bodymass	clearance	duty factor	trunk angle	stride PC1	stride PC2	ε
hip angle mean	+2.48	-0.06	+0.02	+0.06	+0.18	-0.87 *	+0.26	-1.33 *	-0.04 *	-0.04	±0.11
hip amplitude	+0.40	-0.04	+0.02	+0.11 *	-0.09	-0.03	+0.03	-0.02	-0.01	-0.04 *	±0.07
knee angle mean	-1.55	+0.12 *	-0.07	-0.07	-0.34 *	+0.35	-0.31	+0.18	+0.04 *	+0.06 *	±0.12
knee amplitude	+0.94	-0.05	-0.02	-0.09 *	-0.03	+0.26	-0.21	+0.05	+0.00	-0.05 *	±0.07
ankle angle mean	+1.45	-0.03	+0.04	+0.01	+0.09	+0.08	+0.03	+0.01	-0.04 *	-0.06 *	±0.09
ankle amplitude	+0.68	-0.07	-0.04	+0.05	+0.10	+0.24	+0.07	-0.14	-0.03	-0.08 *	±0.11

Table 4.16: Results of the coordination model. Values which do not contain zero in their HDI are indicated by an asterisk.

	intercept	female → male	adlt → adol	adlt → inf	c.log.bodymass	clearance	duty factor	trunk angle	stride PC1	stride PC2	ε
PC1	+0.34	+0.09	-0.11	-0.23 *	-0.01	+0.35	-0.52	-0.01	-0.01	-0.03	±0.18
PC2	-0.06	+0.07	-0.06	-0.05	-0.12	+0.07	+0.00	+0.10	+0.04 *	+0.05	±0.14
PC3	-0.04	+0.11 *	-0.04	-0.07	-0.15	-0.13	+0.16	-0.00	-0.02	+0.05 *	±0.11
PC4	-0.11	-0.01	+0.01	+0.01	-0.03	+0.05	-0.05	+0.20 *	-0.01	-0.00	±0.09
PC5	-0.04	-0.03	+0.04	+0.04	-0.08	-0.15	+0.07	+0.04	-0.02 *	-0.02	±0.07
PC6	+0.02	+0.03	-0.03	-0.02	+0.02	+0.11	+0.02	-0.10	+0.00	-0.00	±0.06
PC7	-0.04	-0.01	+0.03	+0.02	-0.03	+0.01	-0.02	+0.05	-0.01	+0.03 *	±0.06
PC8	-0.01	+0.03	+0.01	+0.01	-0.06	+0.01	-0.01	-0.01	+0.00	+0.00	±0.05

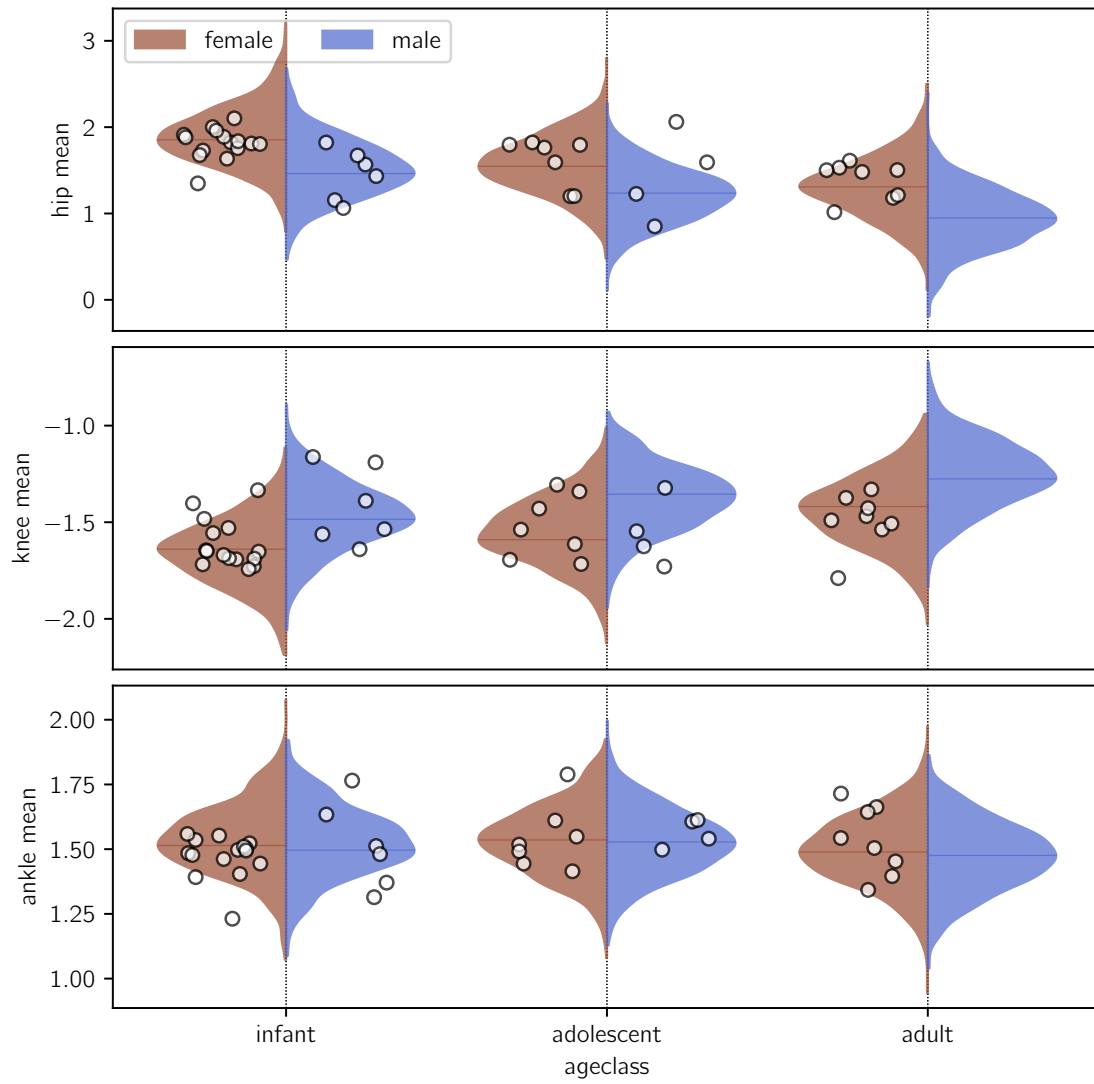


Figure 4.5: *Posture model predictions. Part 1: mean joint angles.*

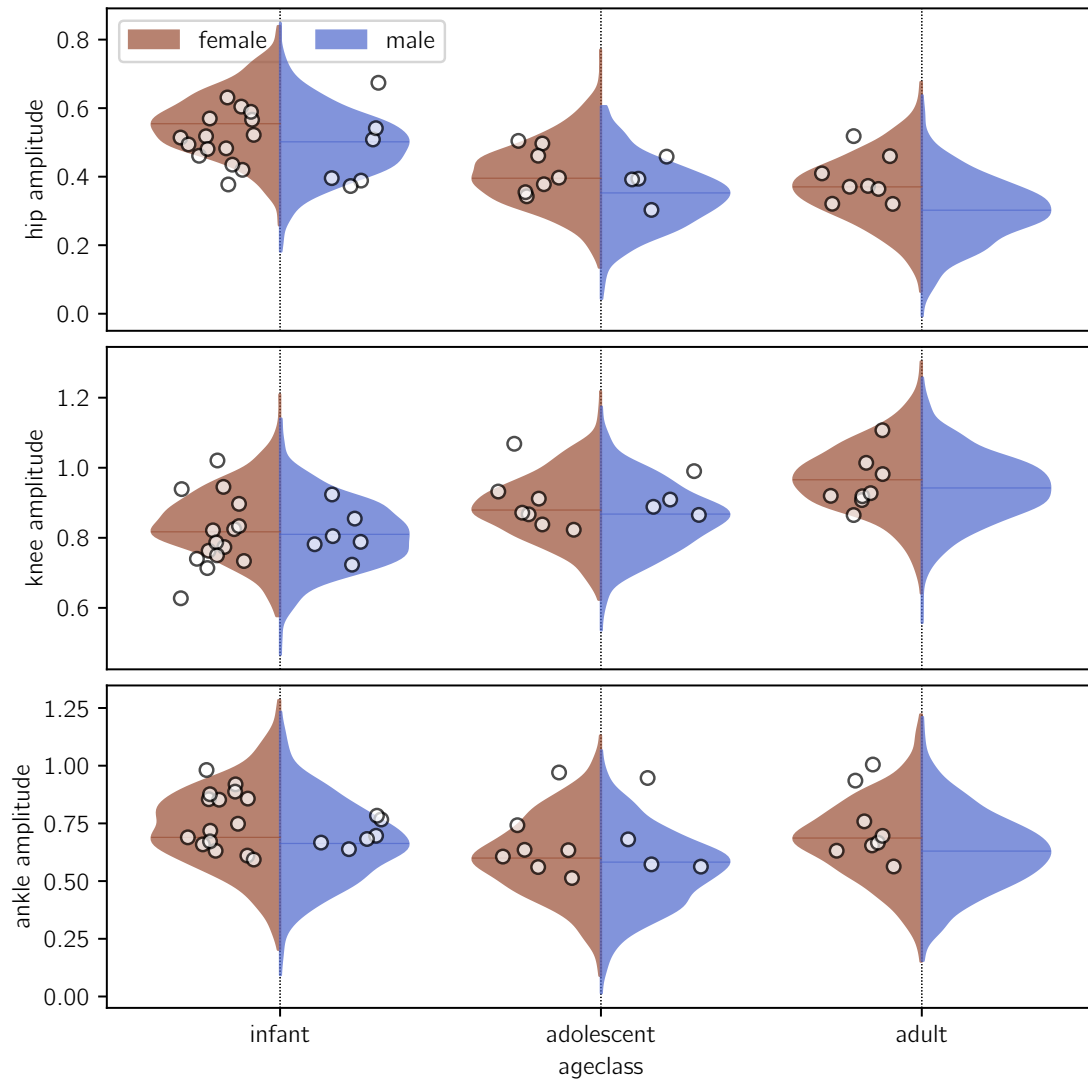


Figure 4.6: *Posture model predictions. Part 2: amplitudes (range of motion).*

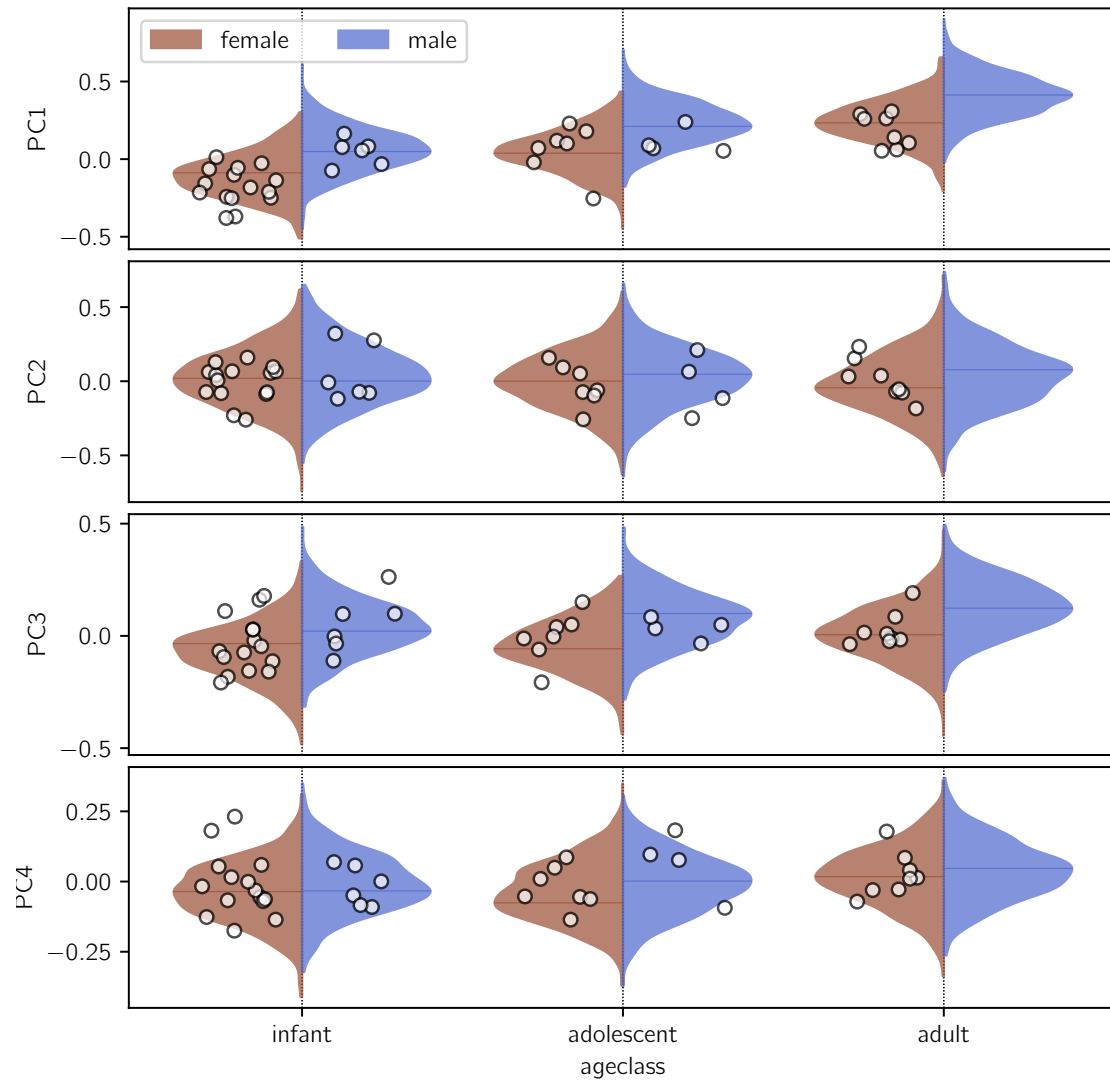


Figure 4.7: *Coordination model predictions. Part 1.*

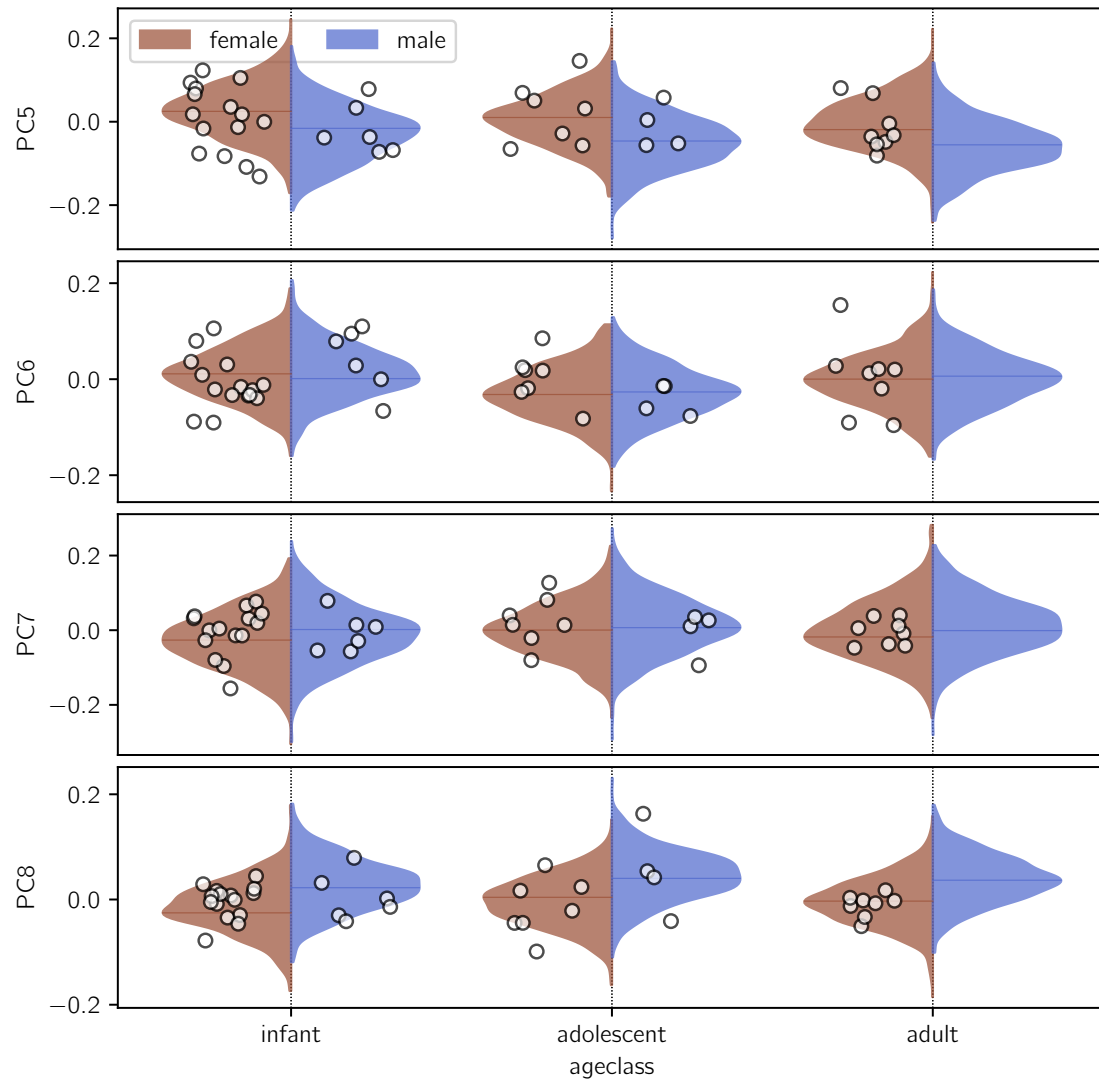


Figure 4.8: *Coordination model predictions. Part 2.*

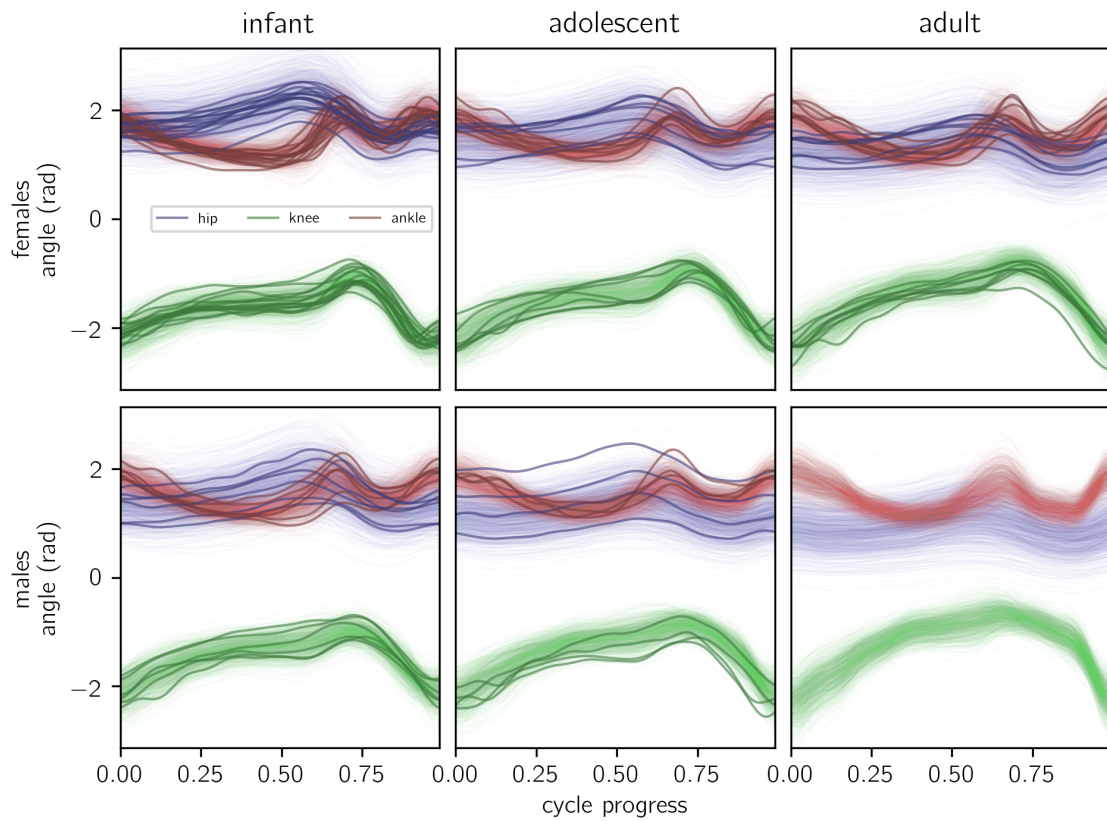


Figure 4.9: Joint angle profile predictions from combined results of coordination and posture models. Thick lines are actual observations; thin lines in the background are posterior predictive samples. Identical to Fig. 2.8 in Ch. 2, yet this time it should be clear how the prediction was performed.

Lots of (New) Questions

The results of in-sample prediction of joint angle profiles in the presented form (Fig. 4.9) mainly serve to confirm whether a model was correctly tuned to the data or not. But this capability also opens up a whole set of questions for which the original data set might have been too fragmented or complex. Here is one example.

I touched above on the influence of duty factor. One might want to ask: **How do joint angle profiles in bipedally walking baboons change with duty factor?** This parameter changes with age class and potentially also with sex (Tab. 4.10). Hence, plotting all joint angle profiles of low duty factor strides (LDS) and comparing them to high duty factor strides (HDS) will probably be confounded with a comparison of juveniles and adults and unbalanced sex ratios in these groups. Sample sizes are low due to data fragmentation, thus filtering for adult females (8 strides) does not retain enough data to derive statistically relevant conclusions. In fact, to answer the duty factor question above, one would like to use information from within-age class comparisons and from both sexes, because each of these classes contribute their own bit to our knowledge of duty factor-related changes in kinematics. This is what the probabilistic model did: it separated effects of the subject and spatiotemporal gait variables. Ideally, the model has learned the concerted influence of these parameters on the outcome parameters.

So let us ask the model what the effect of duty factor is, specifically for adult females! The prediction (Fig. 4.10) depicts the exact temporal coordination of joint angles within the specified group. Changing from LDS to HDS, the hip mean angle increases (more flexed), which indicates that a prerequisite for quicker strides is a somewhat extended posture. The knee is affected in terms of amplitude. Curiously, the ankle joint profile is hardly altered during the swing phase (phases not indicated in figure, but it is the latter part of the cycle progress and can be guessed from the ankle) except that obviously the “second trough” in the trace starts later; the stance phase again is altered in terms of mean joint angle. As a bonus, we can see the variance in the strides (as shaded traces in the background), though that variance might be an over-estimation due to low data sample size (partially influenced by wide priors). Remember that duty factor cross-influences other spatiotemporal gait variables (Tab. 4.9). Luckily, it is easily possible and might be worth comparing the speed etc. of the selected, predicted strides.

The capacity to fix some parameters and “playing” with others is immensely useful for all kinds of research questions, beyond the context of kinematic analysis. As shown above, the prerequisite is a well-fit probabilistic model of the phenomenon you want to analyze, and a sufficiently high number of observations.

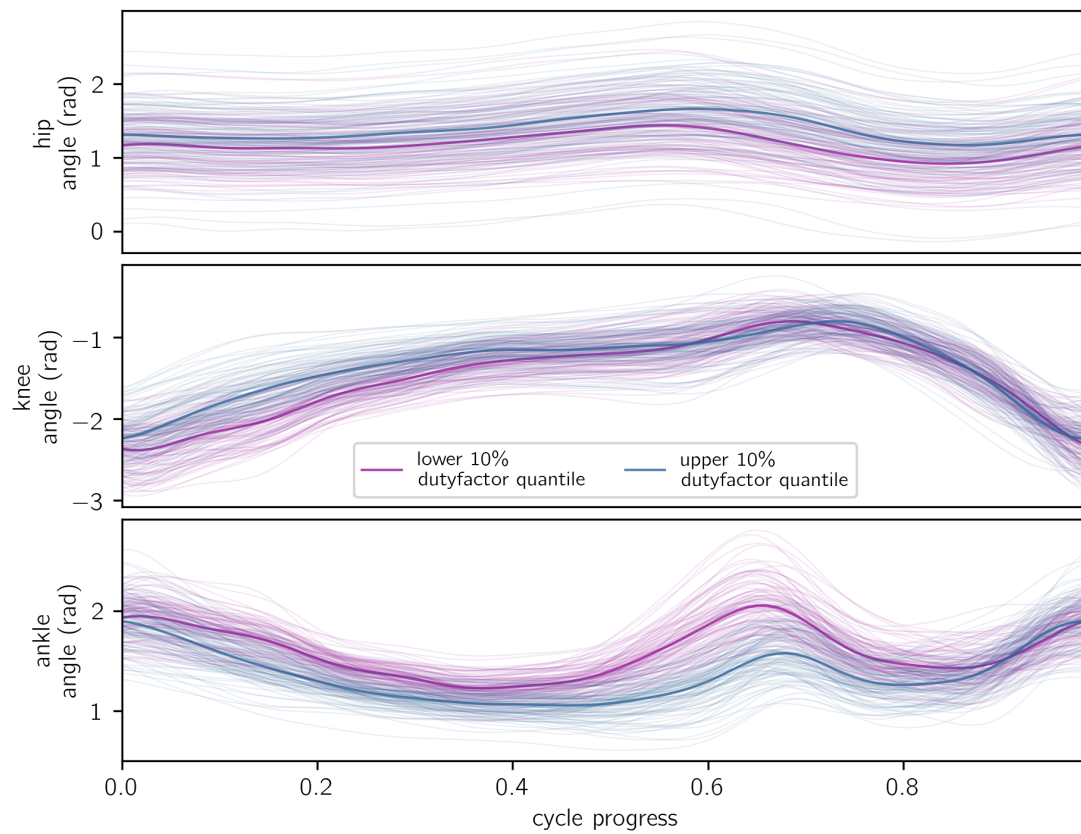


Figure 4.10: Joint angle profile predictions for duty factor quantiles, for adult females. *Thick lines: averages; thin lines: individual predictive samples. Caution: to emphasize differences, the range of y-values varies on the three panels.*

4.4 Summary

With this chapter, I attempted to plant the seed which will grow into your love for the beauty of probabilistic modeling. Alternative (and less ambitious) claim: you could call this my illustrative/colloquial/desperate attempt to explain what I think when I hear or talk “statistics”.

Still, I hope I kindled some good ideas.

This rough overview is the basis of the subsequent chapter, in which I will take the modeling procedure one more step further to predictively model subject parameters as a function of gait, posture, and coordination (Ch. 5). The intention of this chapter was also to facilitate orientation for those who would like to look at the scripts that generated all the models, figures, and tables above (available at https://git.sr.ht/~falk/papio_fcas).

Chapter 5

Predictive Modeling of Piglet Locomotor Kinematics

5.1 Abstract

In the study of vertebrate locomotion, kinematic measures of gait, dynamic posture, and coordination often change when comparing subjects of different body mass, size, and age. Is it, conversely, possible to infer subject characteristics from the kinematic measures? For this study, piglets (*Sus domesticus*) were filmed from lateral perspective during their first ten hours of life, an age at which body mass and size have major consequences for survival. We apply deep learning methods for landmark tracking (DeepLabCut), calculate joint angle profiles, and apply information-preserving transformations (Fourier Series and PCA) to retrieve a multivariate kinematic data set. We train a probabilistic model to predict subject characteristics from kinematics. The model infers subject characteristics accurately for strides from piglets of normal birth weight (i.e. the category it was trained on), but surprisingly predicts the body mass and size of low birth weight piglets (which were left out of training, out-of-sample prediction) to be “normal”. The discrepancies between prediction and observation confirm that dynamic posture and coordination are unaffected by lower birth weight and lower size, which might indicate that low birth weight does not imply locomotor deficits. However, the age of some (but not all) low birth weight individuals is underestimated, supporting the hypothesis that these piglets can experience a delay in locomotor maturation.

This chapter was submitted to the Journal of Experimental Biology in February 2022. Rejection led to the revision of the use of Fourier methods in locomotor biomechanics (Ch. 2). An alternative version of this chapter was finally published in 2023 (Mielke et al., 2023); that text targets a veterinary audience, but results are identical to the version herein. The history of the document can be followed on BioRxiv (<https://www.biorxiv.org/content/10.1101/2022.02.04.479126v3>) and git (https://git.sr.ht/~falk/piglet_fcas).

5.2 Introduction

Vertebrate locomotion is a complex phenomenon. The kinematic and dynamic measurements obtainable in experiments represent the collective output of interacting variables of an ensemble of subsystems: the musculoskeletal apparatus, energy supply, metabolism, and multiple levels of neuro-motor control (Nishikawa et al., 2007). All these subsystems are potentially affected by characteristics of the animal (e.g. age, due to individual development) and by external circumstances (e.g. friction) in different, non-trivial ways. Conversely, studying alterations in locomotor patterns under controlled experimental conditions holds diagnostic potential (e.g. Figueiredo et al., 2018). In well-studied species, it is even possible to identify individuals from kinematics (e.g. Patua et al., 2021).

Yet the complexity of the biological system is a major challenge to locomotor research. In a hypothetical experiment, one might record two animals of different weight, walking at different speeds, and producing two distinct kinematic patterns (e.g. one tending to bend the knee and ankle joints, the other emphasizing bending of knee and hip). Additionally, even for identical external conditions and in a single individual, there is intrinsic variation in the process, owing to the fact that “successive movements [...] never exactly repeat themselves” (Bernstein, 1935). Which of the many “input factors” is responsible for the difference in ensemble output, i.e. kinematics? This analysis question is common in research on bipeds (e.g. Bruton et al., 2013; Ganley and Powers, 2005; Stiffler-Joachim et al., 2020) and quadrupeds (e.g. Irschick and Jayne, 1999; Pike and Alexander, 2002; Stavrakakis et al., 2014), and the solution is not novel. Multivariate models are capable of handling complex situations, given sufficient data. However, the high dimensionality of kinematic data sets, the multi-parameter, multi-level (hierarchical) covariate situations, and the high video tracking workload have often been a limiting factor in the particular case of vertebrate locomotion (Jackson et al., 2016; Michelini et al., 2020; Seethapathi et al., 2019).

Several recent technological advances have enabled researchers to tackle scientific questions on locomotion in a more efficient way. Firstly, the past few years have brought huge leaps in terms of computer vision, deep learning, and thereby semi-automatic video tracking methods (Corcoran et al., 2021; Jackson et al., 2016; Karashchuk et al., 2021; Mathis et al., 2020; Mielke et al., 2020, *cf.* Ch. 9.2). These tools typically require a manually tracked subset of the data as the “training set” for a neural network, which is then able to track landmarks on further videos in high through-put, hopefully with reasonable accuracy. A second field of technological advance are probabilistic models, which build on an elegant technical implementation of Bayesian theory (Markov Chain Monte Carlo / MCMC sampling, *cf.* Gelman et al., 2020; McElreath, 2018; van de Schoot et al., 2021). Such models can naturally incorporate hierarchical parameter interrelations and intrinsic variability. The main reason for this is that probabilistic models work on data

distributions, and their outcome are distributions and “effect likelihoods”, rather than point estimates. This can be informative on an intrinsically varying process such as locomotion (Mielke et al., 2018b). Machine Learning methods for video point tracking are validly advancing to be the standard in kinematic analysis, whereas probabilistic models still lack wide recognition in the field, despite their potential.

Piglets are a well-studied model system in which scientific interest joins the economic interest of commercial breeding. These animals have been subject to a variety of locomotor studies, including paradigms to test the effects of breed (Mirkiani et al., 2021), birth weight (Vanden Hole et al., 2018a, 2021, 2017), surface friction (von Wachenfelt et al., 2008), welfare (Guesgen and Bench, 2017), various pathologies (Abell et al., 2014; Benasson et al., 2020; LaVallee et al., 2020), and more (*cf.* Netukova et al., 2021). Of particular interest has been the occurrence of a subset of individuals which are born with lower weight (LBW, low birth weight) than their “normal” (NBW) littermates. There are multiple standards to classify these birth weight categories, using absolute mass, litter quantile criteria, or asymmetry of body proportions (Amdi et al., 2013; D’Inca et al., 2011; Feldpausch et al., 2019; Quiniou et al., 2002; Roehe and Kalm, 2000; Van Tichelen et al., 2021; Wang et al., 2016). A possible cause of low birth weight is intra-uterine growth restriction, and LBW phenotype seems often, but not always, to correlate with low vitality and a reduced chance of survival (Baxter et al., 2008; Hales et al., 2013; Muns et al., 2013; Van Ginneken et al., 2022). Locomotor maturation after birth is quick (Andersen et al., 2016; Vanden Hole et al., 2017), yet crushing by the sow constitutes one of the major causes of piglet mortality (Edwards and Baxter, 2015; Marchant et al., 2000). The likelihood of being crushed is directly reduced by more agile locomotion. Thus, locomotor capabilities are crucial for piglet survival, and delayed development might be fatal.

Previous studies from our group (Vanden Hole et al., 2021, 2017) raised the hypothesis that the apparent difference in LBW and NBW individuals can be attributed to delayed development. They measured spatiotemporal gait variables (e.g. stride frequency and distance, speed, duty factor), which are collective variables of the actual kinematics (*cf.* Aerts et al., 2000; Newell and Liu, 2021; Nishikawa et al., 2007). This strategy has the advantage that it requires only five landmarks (four limbs, one reference) to be tracked, which used to be a crucial trade-off to handle large data sets. However, the the collective variables cannot capture full information on intra-limb coordination (i.e. the relative timing of segmental movements within a limb; as opposed to inter-limb coordination, i.e. the relative timing of the cycling of the different limbs). This complicates disentangling effects such as those of size, age, (birth) weight, and disease. It is expected that animals adapt their gait to the physical constraints of motor behavior, which are depending on the weight and other characteristics of the subject. However, the changes to kinematics might be more subtle, and collective variables might not be altered in a distinct way. For example, an animal might learn to move its joint angles in a more efficient way by adapting

clearance to substrate conditions (Mielke et al., 2019; von Wachenfelt et al., 2008), which could in principle be achieved without changing the speed of voluntary locomotion on those substrates. Hence, it would be desirable to include full kinematic information.

Using the semi-automatic landmark tracking techniques mentioned above, one can extend the analysis of spatiotemporal gait variables to quantities of intra-limb coordination with manageable effort. However, using the whole set of raw point coordinates of joint angle profiles of interest raises the issue of dimensionality (two to three coordinates per reference point/joint, minimum a hundred temporal samples; simply too many data variables). Statistical modeling requires a minimum number of observations for being able to infer effects of the different variables (Austin and Steyerberg, 2015; Frick, 1996; Maxwell et al., 2017; Riley et al., 2020). The common solution is to reduce the dimensionality with an appropriate transformation. To choose a transformation, it can be exploited that common analysis procedures in locomotor biomechanics require steady state locomotion. “Steady state” implies that the behavior consists of repetitive blocks of kinematics, i.e. stride cycles. And one of the most common sets of techniques in physics and engineering to handle cyclic data is Fourier Analysis, or more specifically Fourier Series Decomposition (FSD; Bracewell, 2000; Fourier, 1822; Gray and Goodman, 1995; Mielke et al., 2019; Pike and Alexander, 2002; Webb and Sparrow, 2007). With FSD, joint angle profiles are transformed into their representation in the frequency domain, i.e. an array of harmonics. Some of the characteristics of the profiles (namely mean joint angle, amplitude, and phase) are more readily captured by those harmonics and can optionally be removed. This is most intuitive in the case of phase: removing phase differences enables a mathematically optimal temporal alignment of the profiles. By isolating the other characteristics, mean and amplitude, the joint angle profiles can be transformed to meaningful quantities such as dynamic posture (mean joint angle and effective range of motion), and coordination *sensu stricto* (relative phase/joint timing and residual kinematics, *cf.* Mielke et al., 2019, and Ch. 2). Harmonics are independent of temporal sampling and duration: the coefficient array is of fixed size, which is useful for subsequent multivariate analysis methods, such as Principal Component Analysis (PCA). Another advantage of this transformation procedure is that it is reversible because all mathematical information is retained in the process (which is not the case when using collective variables alone). This means that joint angle profiles can be reconstructed for any observed or hypothetical point in parameter space, which enables in-sample and out-of-sample predictive sampling.

To summarize, the Fourier Series decomposition provides a mathematically convenient and biomechanically meaningful representation of the kinematic data, which opens up new options for data analysis and modeling.

In this study, a conventional, 2D kinematics data set is extracted with the aid of deep learning tools from lateral videos of walking piglets. By applying multivariate analysis and FSD, we separate spatiotemporal gait variables, dynamic posture, and coordination, and model their

relation to subject characteristics (mass, size, age, and birth weight category). Crucially, this constitutes the complete information captured by locomotor kinematics. All parameters are submitted to an inclusive, probabilistic model. We tackle the question of whether low birth weight is an indication of delayed development, and attempt to quantify the delay with an inverse modeling strategy as follows. Intuitively, and conventionally, joint kinematics are considered the output of the locomotor system. Therefore, conventional statistical models might consider them on the “outcome” side; on the “input” side, the effects of birth weight, age, speed, or other parameters are quantified. Herein, we use a different approach, and invert the model. We construct a probabilistic computer model which describes “age” and other subject characteristics as a function of all available kinematic parameters. The rationale is similar to that in subject recognition tasks (e.g. Patua et al., 2021): given a certain kinematic profile, can we infer characteristics of the subject which produced it? We split our data set into birth weight classes, and train the model on only the strides from NBW observations. This NBW model is our “kinematic reference” model, quantitatively capturing the expectation of what would be “normal” by inferring the plausible age range for a given kinematic observation. We then use the model and compute out-of-sample prediction of all LBW observations.

Our hypothesis is that, if LBW were at the same stage of postnatal locomotor development as their NBW siblings, then the prediction should accurately infer the age of the LBW animals. Conversely, if the LBW piglets are delayed in development, the model would underestimate their age. Thus, by applying this inverse modeling strategy and comparing the computer-predicted age to the actual age of the LBW piglets, we can falsify or quantify a hypothesized delay in locomotor development.

5.3 Materials And Methods

5.3.1 Data Acquisition

Recordings were done at a local farm in Belgium during several trips in October and November 2017. Farrowing was monitored to select Topigs x PIC piglets for another experiment (Ayuso et al., 2021). Piglets from selected litters were weighed at birth and numbered with non-toxic skin markers. Low birth weight (LBW) was classified by birth weight quantile (lowest 10% of each litter) and by a maximum mass of 800g (D’Inca et al., 2011; Litten et al., 2003; Van Tichelen et al., 2021; Wang et al., 2016); all other piglets are assigned the NBW category. At variable time points afterwards (ages 1 – 10 hours), piglets were briefly taken from their pen and brought to a separate room for video recording. Animals were recorded in pairs (as in Mielke et al., 2018b), which drastically reduced anxiety and increased their motivation to cooperate. A few animals were recorded repeatedly, usually with a changing partner. Animals were ear-tagged and followed up: recording was repeated at approximately 4 and 10 days of age. That data was part of the landmark tracking procedure (i.e. “DeepLabCut” network training), but excluded from further analysis (i.e. probabilistic modeling). The subject characteristics documented for analysis are birth weight (continuous, and categories “LBW”/“NBW”), mass at recording, age at recording (i.e. hours since farrowing), sex, and size. The size of the animal was approximated by a Principal Component Analysis (PCA) of landmark distances along all segments (“size PCA”, only first PC used, 93% of variability). Size and mass are expected to correlate, yet deviations would indicate animals of particularly slender or rotund habitus. All procedures followed ethical regulations and guidelines, and were approved by the Ethical Committee for Animal Testing of the University of Antwerp, Belgium (ECD 2015-26).

The recording room contained an elevated runway (150×50cm), covered with a rubber mat to increase friction, and visible through a transparent frontal shield. Color videos were recorded (camera model: GC-PX100BE, JVC, Japan) at a temporal sampling rate of 50 frames per second and a spatial resolution of 1920 × 1080 pixels (later cropped to 500 px height), from a distance at which the field of view would exactly capture the entire runway. A chess board at the back wall enabled spatial calibration. Video surveillance was permanent during the presence of the animals and stopped only in between recording sessions. Animals were able to move freely on the enclosed platform. To stimulate locomotion, the two animals were repeatedly placed on opposite ends of the runway. Gentle tickling on the back and grunting vocalization of the researcher were other successful strategies to induce targeted locomotion in the direction perpendicular to the camera axis. After recording sessions the piglets were returned to their litter and remained with the sow. The workflow herein involved handling of the animals as a consequence of the research setting. However, note that the procedure could easily be automated for continuous data collection by a suitable pen arrangement (Meijer

et al., 2014; Netukova et al., 2021; Stavrakakis et al., 2014).

5.3.2 Landmark Tracking

We used the software DeepLabCut (DLC, Mathis et al., 2018) for landmark tracking of all video material. In addition, a custom made point tracking software (Mielke et al., 2020) was used to generate a training set. In total, our dataset contained 180 videos (more than 11 hours, 169 animals). Our goal was to prepare a general DLC network which would be capable of automatically tracking piglets at multiple ages, and which can be shared and re-used for subsequent research questions. This was the initial motivation to use the full data set for landmark tracking and for the calculation of some derived measures (size PCA), yet the quality of the automated tracking was insufficient, and no generalizable neural network could be trained. The analysis focus of this study was only a subset of the data (i.e. the 58 animals of the youngest age class). The video processing workflow, applied to the full data set, was as follows. To get a balanced training set, one stride of each of the animals was selected, and the video was cut, cropped to runway height, and optionally mirrored horizontally so that movement would always be rightwards. All videos were concatenated and submitted to the DLC training set generation. DLC was set to select 2552 frames from these videos, which were tracked in an external software and re-imported for training (80% training fraction). Seventeen landmarks (i.e. points of interest or “key-points”; usually joint centers, fig. 5.1) were tracked, representing all body parts visible on the lateral perspective (head: snout, eye, ear; back line: withers, croup, tail base; forelimb: scapula, shoulder, elbow, carpal/wrist, fetlock, forehoof; hindlimb: hip, stifle/knee, tarsal/ankle, hind fetlock, hindhoof). We selected a “resnet 152” network architecture and trained for 540,672 iterations (16 days of computer workload). The network was then applied to track the continuous, full 11 hours of video twice: once in default direction and once horizontally mirrored, because training set was always rightward movement.

The next step is to find the relevant temporal sequences of walking in the continuous videos. Naturally, the trained network would only extract potentially useful landmark traces for episodes which resembled the training set, i.e. in episodes with a piglet moving perpendicular to the image axis, in lateral aspect and rightward direction. We automatically extracted 2597 of such sequences by filtering for high landmark position “likelihood” provided by DLC, low noise (i.e. steady landmark movement) and consistent, plausible inter-landmark distances. We further applied an automatic algorithm to find footfalls and label stride cycles in the candidate episodes (4730 cycles). This procedure involved a start-end-matching optimization (using Procrustes superimposition, *cf.* Ch. 2) to ensure that strides were indeed cyclical. To further assess tracking data quality, gait variables were automatically extracted. Definition of these variables was chosen to simplify the automatic procedure, as follows. Stride distance, frequency, and speed are trivial measures of the animal movement. Duty

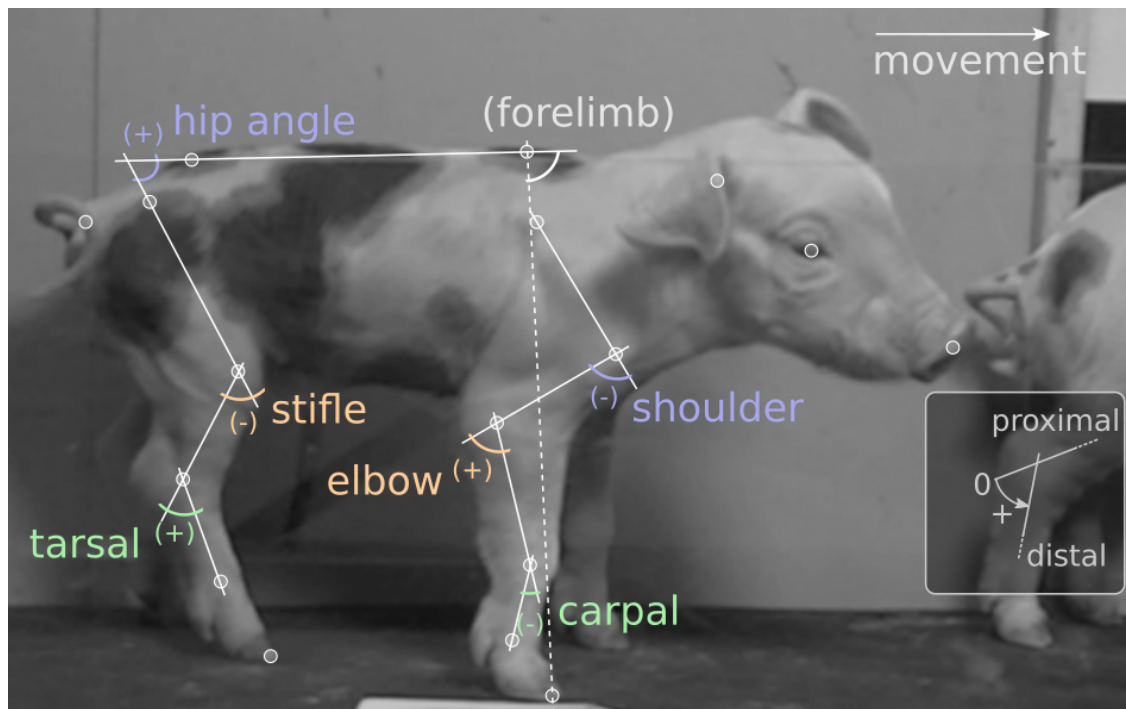


Figure 5.1: Video landmark tracking and joint angle definitions. White circles mark points of interest (“landmarks”). Movement was always rightwards. Labels show joint angles, defined as shown in the inset: straight joint (parallel segments) corresponds to zero; counter-clockwise joint angles are positive. Forelimb joint angle was used as a reference for temporal alignment, but did not enter the analysis.

factor is available for fore- and hindlimb, and measures the fraction of stride time in which the respective hoof is in ground contact. Clearance is approximated by quantifying the ratio of flexion of each limb (one minus the quotient of minimum and maximum absolute hip-hoof-distance during the stride). Head and torso angle are the stride-average angles of the snout-ear or withers-croup lines with respect to the coordinate system. Hindlimb phase measures the time between hind- and forehoof touchdown, divided by the stride cycle duration. Where applicable, spatiotemporal gait variables were prepared for analysis by converting them to dimensionless values (Alexander and Jayes, 1983; Hof, 1996) using the accumulated distance of landmarks along the snout-to-tailbase line of the animal as reference, extracted as stride average from the tracked landmarks. Only strides with plausible values (i.e. those which lie within the theoretical distribution of each parameter; 1862 cycles) were processed. Manual inspection further boiled down the data set to 897 stride cycles (the others excluded for tracking errors, multi-animal confusion, non-walking gait, intermittent or sideways locomotion, or incompleteness).

Finally, 368 of the remaining strides from 58 animals were in the youngest age category ($< 10 h$) and thus selected for the present analysis, the data table is available online.

5.3.3 Data Processing

The landmark data provided by DLC was further processed for analysis. Python code for the whole procedure is available (https://git.sr.ht/~falk/piglet_fcas, Python version 3.11 at time of model calculation, <https://www.python.org>). First, joint angle profiles (i.e. joint angle values over time) were extracted for all relevant joints and for the total forelimb angle (croup-withers-hoof). Shoulder, elbow, carpal, hip, stifle, and tarsal were the six joints sufficiently well tracked and therefore considered relevant for analysis. We then applied Fourier Series decomposition in the framework we previously termed Fourier Coefficient Affine Superimposition (FCAS, Mielke et al., 2019, and Ch. 2), a flexible procedure which subsumes the following steps. Joint angle profiles are cyclic, i.e. periodical, and can therefore be transformed to the frequency domain with a Fourier Series decomposition (8 harmonics were deemed sufficient by visual comparison of raw and transformed/retransformed profiles). In the frequency domain, the affine components (mean, amplitude, phase) of a joint angle profile are easily accessible (*cf.* Mielke et al., 2019, and Ch. 3). The forelimb angle served as reference to temporally align all cycles in the data set (removal of phase differences between different cycles; forelimb angle was not used further). Then, mean and amplitude of the joint oscillations were isolated for all joint angles and are categorized as “dynamic posture” parameters. Mean joint angle is the temporal average, whereas amplitude is related to effective range of motion (eROM). The residual, i.e. differences captured by non-affine Fourier coefficients, can be categorized as “coordination” *sensu stricto* (it measures the precise temporal succession of joint configurations). In our case, there were 96 variables of

coordination (6 joint angles, 8 harmonics, real and imaginary) which were submitted to a PCA. Only the first 12 coordination components (*CC*) were used for statistical analysis, capturing 80.2% of the variability in coordination.

To summarize, FSD and FCAS served three purposes: (i) temporal alignment of the cyclic traces, (ii) separation of meaningful parameter categories (dynamic posture and coordination), and (iii) preparation for multivariate analysis via PCA. Basic script code (Python, Matlab and R) to perform FCAS can be found on a dedicated git repository (https://git.sr.ht/~falk/fcas_code).

Information retention is generally a strength of this method. FCAS and PCA are mathematical transformations, which means that the information content after transformation is theoretically identical to that prior to transformation (theoretically, because only a finite number of harmonics can be used, yet this is of little concern for continuous, smooth joint angle profiles). The neglected PCs and the residual not captured by 8 harmonics were the only information from kinematics of the given joints to be lost in this procedure, and by definition these contain the least information. Apart from that, all information present in the raw joint angle profiles enters the analysis. Though we used a 2D dataset herein, the procedure could be applied equally well to joint angles measured from 3D coordinate data (Scott et al., 2022).

Furthermore, all transformations are reversible, hence any analysis outcome can be translated back to kinematics with high accuracy. Reversibility bares a lot of herein unused potential, for example for interpolating unobserved subject states or for inferring kinematics of fossil species by phylogenetic and morphometric bracketing. Reversibility can also be of use when presenting raw joint angle profiles and their averages, as follows. One crucial aspect of the FCAS procedure is temporal alignment of the joint angle profiles in the frequency domain. In conventional temporal alignment, a single characteristic point in the stride cycle is chosen as a reference, wherein this is only “characteristic” for a certain part of one limb (e.g. left hindlimb hoof touchdown). Temporal alignment to the hindhoof touchdown might cause distinct peaks in the forelimb angle joint profiles to occur at different relative points in the stride cycle (e.g. tarsal joint profiles in Fig. 5.3, lower half, green traces). If profiles show such variable peak positions, then their average will have a wider, less pronounced (i.e. lower amplitude), and potentially unnatural peak. For illustration, this is analogous to averaging two sine-waves of identical amplitude, but phase shifted: in the worst case, they cancel each other out (as in “destructive interference”). The problem is not restricted to pronounced peaks, but generally occurs if the temporal intra-limb coordination varies within a data set. Using FCAS, it is possible to get a more representative average of the raw traces which has its amplitude conserved, but phase and mean joint angle averaged. This is enabled by transformation to the frequency domain, separation of affine components, removal of phase differences by shifting to average phase, profile averaging, followed by

inverse transformation back to the time domain. Because a set of profiles and phases may be calculated for each joint angle individually, and because phase relations can differ between joints, there are the options to align based on one reference angle (e.g. the whole forelimb, as done herein) or minimize all phase differences across all joints. Choosing the first option herein has implications: when plotting hindlimb joints aligned by a forelimb reference (as in Fig. 5.3, lower half), phases still differ, and the “destructive interference” problem might hamper averaging. In such cases it is possible to apply an extra, joint-wise FCAS alignment for the sole purpose of generating meaningful averages.

5.3.4 Statistical Modeling

To summarize, four categories of variables were used for analysis:

- subject characteristics: age, sex, mass, birth weight category, size
- spatiotemporal gait variables: distance, frequency, speed, clearance (fore-/hindlimb), duty factor (fore-/hindlimb), head angle, hindlimb phase
- dynamic posture: mean joint angles and eROM for six joints
- coordination: the residual after extraction of dynamic posture

Our guiding question for model design is whether a probabilistic, linear model is able to infer subject characteristics (specifically: age, mass, and size) from raw kinematics (expressed as dynamic posture and coordination) and spatiotemporal gait variables (collective variables). Given the common conception that kinematics are a complex output of an individual motor system, this might be considered an “inverse” modeling approach (it is certainly an inversion of the model presented in Ch. 4). The present analysis focused on three outcome variables (Fig. 5.2): mass (kg), size ($arb. units$, from a PCA of marker distances), and age (h). Though these outcome variables were specific per individual and recording session, we analyzed them “per stride” (i.e. there were multiple strides with identical subject measures on the outcome side).

The model formula is:

$$\theta \sim v_1 \cdot \alpha + v_s \cdot \beta_s + \sum_G v_g \cdot \beta_g + \sum_P v_p \cdot \beta_p + \sum_C v_c \cdot \beta_c + v_1 \cdot \epsilon \quad (5.1)$$

Herein, θ is either of the outcome subject characteristics, β are slopes associated with the model parameters (s sex, G spatiotemporal gait variables, P dynamic posture, C coordination), v are data vectors (e.g. v_1 is a vector of ones for the intercept α and model residual ϵ , and v_s is a boolean vector coding for subjects of ‘sex == male’). The models have a total number of 36 degrees of freedom. Priors (i.e. *a priori* assigned distributions) for all slopes were Normal distributions with mean and standard deviation corresponding

to the mean and two times standard deviation of all observed values of each parameter; logarithmic transform was applied where necessary. The observable (“likelihood”) prior for θ was a Student’s T distribution (allows for wider-than-normal tails and robust regression) with a Gamma distributed ν (degrees of freedom); ϵ was modeled to be a Half Cauchy distribution. The model was implemented using the Python library “PyMC” (different versions up to 5.10.2; Salvatier et al., 2016).

To re-emphasize, dynamic posture and coordination together effectively capture all the kinematic information of the stride. Hence, we train the predictor model with all kinematics, spatiotemporal gait variables, and sex. Birth weight category (LBW, NBW) is a filter parameter: we split our data set into LBW strides and two NBW subsets (training and validation). Training is performed by MCMC sampling (‘sample’ function in PyMC), and a No U-Turn sampler was set to sample with 32 chains, each 2^{14} tuning and equally many sampling steps. All post-hoc model checks confirmed convergence (inspection of traces, $\hat{b}_{fmi} > 0.94$ for all chains, Gelman-Rubin statistics ≈ 1 for all parameters, sufficient effective sample size). Model comparison was performed (cf. Ch. 4), iteratively leaving out model parameters or replacing some by meaningful combinations (e.g. duty factor combined for fore- and hindlimb). However, because we follow an “all in” strategy, the results have little instructive value for model construction: we might thus have retained parameters which are numerically unimportant for the NBW-only models.

The data set of $N = 368$ strides was split into three categories: (i) the NBW training set as reference with $N = 294$ strides, (ii) the NBW validation set ($N = 35$ strides), which is a random subset of NBW strides, approximately equal in size to (iii) the LBW test set with $N = 39$ strides.

The model was thus trained with a set of 294 NBW training strides (i). Inferences (model “predictions”) were then done per stride, for all observed strides (NBW training, NBW validation, and LBW test), iteratively using the ‘pymc.sample_posterior_predictive’ function in PyMC after setting all the data arrays to the actual observed values for one given stride (using ‘pymc.set_data’). The number of predictions usually matches the number of training samples, which means that all posterior information is used to construct the prediction distributions. We would thus retrieve mass, size, and age predictions (i.e. probabilistic inference) for each stride in the data set, which were then compared to the known, actual mass, size, and age.

All procedures, code, data, and this manuscript are available online (https://git.sr.ht/~falk/piglet_fcas).

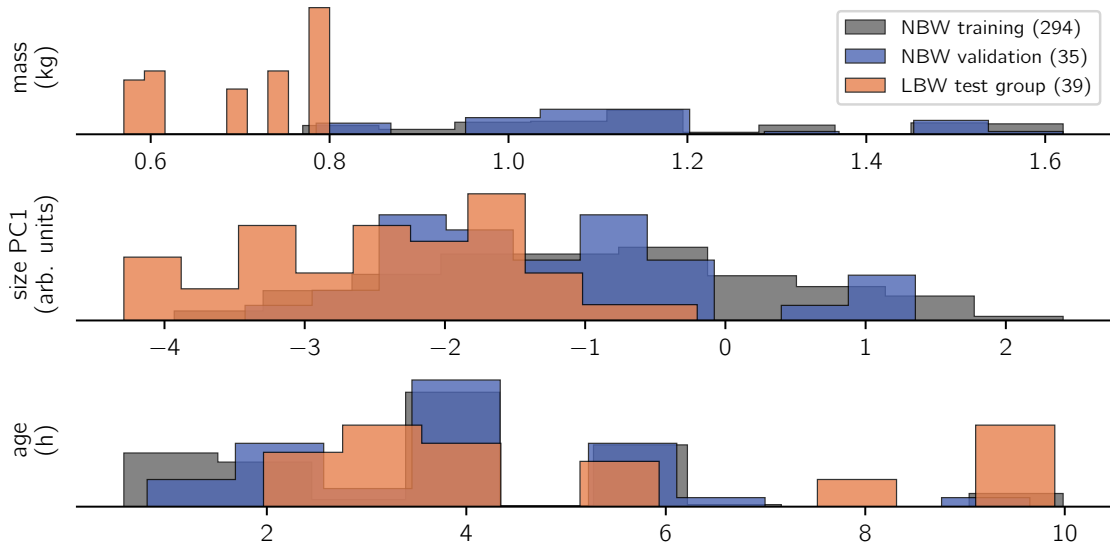


Figure 5.2: Histogram of observations. Trivially, the LBW group measured the lowest body masses in the data set. This correlated with a lower body size, whereas age is rather uniformly sampled for all study groups. Recordings happened opportunistically within the first ten life hours of the animals, repeated measurements were possible. Number of strides per class are indicated in brackets on the legend. Bar heights are scaled by sample size to show relative value distributions.

5.4 Results

The present analysis is centered around a linear model which is designed to infer mass, size, and age (subject characteristics) from an extensive set of kinematic parameters from 2D videos. The numbers provided by the model sampling are equally extensive, and will only be reported in brief. The key purpose of the model is posterior predictive sampling of the LBW strides which were left out of the model, and which are analyzed in detail below.

To assess whether there are qualitative differences between the birth weight categories, one can compare the joint angle profiles (i.e. raw, angular kinematics) on which the present analysis was performed (Fig. 5.3). The intra-group variability clearly exceeds the differences between groups, although it must be emphasized that groups are heterogeneous (with regard to age, speed, etc.), which might lead to a bias if composition of LBW and NBW data differs. LBW walk with a more flexed hindlimb posture, as indicated by the parallelly offset average hip, stifle, and tarsal profiles. Additionally, NBW individuals on average seem to set the shoulder at a more extended joint angle. No differences in coordination are apparent (which would manifest in altered temporal structure of the profiles). These findings indicate that LBW kinematics are hardly distinguishable from NBW kinematics by qualitative, visual assessment, which is at least in part be due to high variability.

A quantitative comparison of variable kinematic measurements can be achieved with probabilistic linear models. For the purpose of predictive sampling, we train models to describe the interrelations of kinematic parameters and subject characteristics in NBW piglets. The outcome of MCMC sampling of a linear model are value distributions for slopes, which in our case indicated how certain kinematic parameters are associated with a change in mass, size, and age (Tab. 5.1; Fig. 5.6). Of the gait- or coordination parameters, only hindlimb clearance was correlated with differences in animal mass. Mass was also associated with changes in the dynamic posture of the hip and tarsal. For size, the model inferred associations with head angle, hindlimb duty factor and clearance, and one coordination component (CC3), as well as changes in the fore- and hindlimb posture and an effect of sex. Finally, age was associated with an increase in forelimb clearance, potential changes at the stifle and carpal, and several coordination components (CC9, CC11). Some eROM slope distributions for age were high in average magnitude, but variable (the “credible interval” contained zero). More details on the model input and output are discussed in a supplementary analysis to this chapter (Ch. 5.6). These model results provide detailed insight into parameter interrelations in the present data set and indicate which of the parameters are the relevant ones to infer a given subject attribute in predictive sampling.

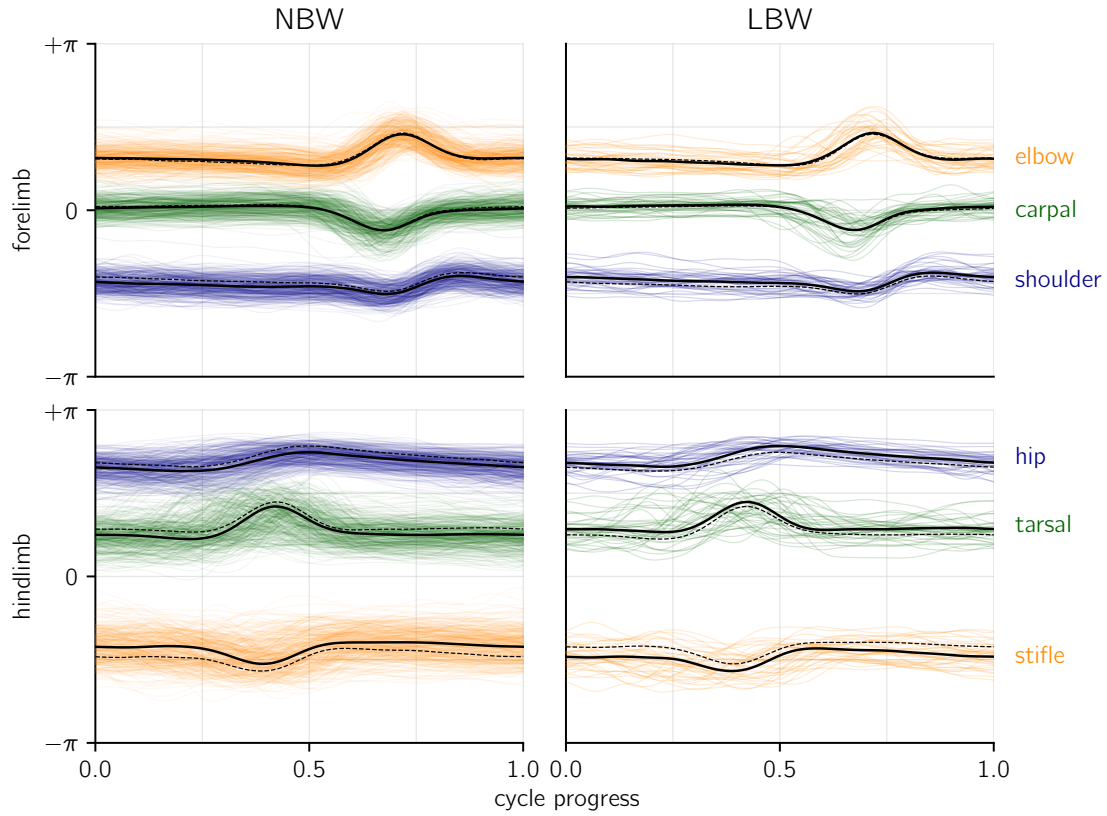


Figure 5.3: Joint angle profiles per joint, grouped by birth weight category. A joint angle of zero would be a fully extended (i.e. straight) joint. Thick lines represent the average profiles, dashed lines indicate the average of the opposite birth weight group for comparison. Colored, shaded lines show all raw profiles available for the present analysis. Temporal alignment was done based on total forelimb angle (see methods), yet for the shown hindlimb averages (but not for the raw profiles), a separate alignment of the hindlimb was performed.

Table 5.1: Detailed modeling results. Asterisk (*) indicates slopes for which the credible interval did not include zero. FL: forelimb, HL: hindlimb, dyn.p.: dynamic posture, coord.: coordination, diml.: dimensionless, d.s.: dimensionless stride, eROM: effective range of motion.

	category	parameter	age (h)	age (log)	size PC1	mass (kg)
0	model	intercept	+4.86	+1.68	-3.25	+1.38
1	subject	female → male	+0.05	+0.00	-0.38 *	-0.04
2	gait	log. FL clearance	-0.10	-0.08	+0.09	-0.02
3	gait	log. HL clearance	+0.94 *	+0.32 *	-0.65 *	-0.15 *
4	gait	FL duty factor	+1.11	+0.37	+0.18	+0.07
5	gait	HL duty factor	+0.67	+0.24	+0.85 *	+0.11
6	gait	d.s. distance	+1.32	+0.31	+0.45	+0.18
7	gait	d.s. frequency	+3.12	+1.43	+1.62	+0.15
8	gait	diml. speed	-1.10	-0.64	+0.82	+0.00
9	gait	hindlimb phase	-2.04	-0.56	-1.98	+0.05
10	gait	head angle	+0.26	+0.14	+1.44 *	+0.18
11	dyn.p.	mean hip angle	+2.55 *	+0.59 *	-1.07 *	-0.45 *
12	dyn.p.	hip eROM	+2.64	+0.94	-1.34	-0.52 *
13	dyn.p.	mean stifle angle	+0.67	+0.10	-2.01 *	-0.12
14	dyn.p.	stifle eROM	-1.85	-0.22	-1.66 *	-0.13
15	dyn.p.	mean tarsal angle	-1.18	-0.59	+0.44	+0.23 *
16	dyn.p.	tarsal eROM	-1.92	-0.88 *	+2.98 *	+0.60 *
17	dyn.p.	mean shoulder angle	+1.06	+0.07	+0.47	+0.06
18	dyn.p.	shoulder eROM	-0.34	-0.17	-0.65	+0.01
19	dyn.p.	mean elbow angle	-1.00	-0.55	+2.79 *	+0.20
20	dyn.p.	elbow eROM	+3.22	+0.96	+0.11	-0.30
21	dyn.p.	mean carpal angle	-2.08 *	-0.98 *	+1.69 *	+0.04
22	dyn.p.	carpal eROM	-0.24	+0.21	-0.84	+0.16
23	coord.	CC1	+0.60	+0.31 *	+0.10	-0.05
24	coord.	CC2	+0.50	+0.15	-0.24	-0.08
25	coord.	CC3	-0.25	+0.00	+0.88 *	+0.07
26	coord.	CC4	-0.62	-0.26	+0.13	+0.04
27	coord.	CC5	-0.84	-0.25	-0.16	+0.04
28	coord.	CC6	-1.49	-0.51 *	-0.10	+0.07
29	coord.	CC7	+0.24	+0.22	-0.76	-0.08
30	coord.	CC8	-1.19	-0.24	-0.51	+0.02
31	coord.	CC9	-2.96 *	-0.83 *	+0.66	+0.06
32	coord.	CC10	-0.22	+0.06	-0.60	+0.02
33	coord.	CC11	-2.05 *	-0.67 *	+0.10	+0.08
34	coord.	CC12	-0.03	-0.10	+0.35	-0.01
35	model	ϵ	± 1.81	± 0.56	± 0.98	± 0.20

Performing in-sample and out-of-sample predictive inference with the models trained on NBW strides elucidated if and how left-out strides differed from NBW model expectation (Fig. 5.4). Note that, to capture variance (i.e. uncertainty in the prediction), each stride was sampled repeatedly.

Out-of-sample inferences for the *NBW validation set* matched those of in-sample NBW inference in terms of average values and standard deviation for all modeled outcome variables, which confirms that inference of subject characteristics from kinematics is possible. In contrast, inferences for the *test set* of LBW strides did not match those of the NBW training set. Low birth weight animals were inferred to be on average 0.44kg heavier than actual, and their size was overestimated ($+1.71$ units). Both faults matched the actual differences in magnitude (*cf.* methods, Fig. 5.2). In contrast, the age inference for the low birth weight subjects were not normally distributed: most ages were correctly inferred from stride-wise kinematics, but ages for some strides were underestimated. The underestimation of those strides quantified to just below five hours.

In summary, the NBW-trained model “guesses” the size and mass of the animals producing LBW strides to be “normal” (although they are not), which indicates that these defining features of LBW do not reflect in altered kinematics. However, age inference is non-normal, i.e. some strides are classified as typical for animals of younger than actual age.

To assess whether out-of-sample prediction is a valid and unbiased method to infer group differences, we extracted additional indicator quantities from the linear models. An in-depth analysis can be found at the end of this chapter (Ch. 5.6). Theoretical considerations identify several criteria which, if not met, can obscure an actual group difference (which consequently does not emerge in the prediction output). These criteria are differences in the input parameter distribution, slope magnitude, and low residual variability; in addition, model complexity and a limited sample size can be prohibitive. For the age model, our considerations confirm that a predicted difference is robust, and we were able to evaluate which of the model parameters might cause the underestimation of some strides (primarily hindlimb clearance and hip angle). In the models of size and body mass, some of the criteria are not met. Model complexity and sample size might restrict clarity of both models’ outcomes to an uncertain degree. In the body mass model, the residual variability is high, which blurs potential differences in the predictions (Ch. 5.6, Fig. 5.6). In the size model, there are some positive and some negative effects which seem to cancel out (which might be evidence for size-dependent alternative locomotor strategies; Figs. 5.5, 5.6). We observed few significant effects related to the coordination parameters, indicating that coordination patterns are for the most part indifferent to age, size, and mass.

This lends some support to the outcome of the age model, and points towards the limiting factors in case of the other two models.

To find out whether the offset age inference was related to certain individuals, or certain

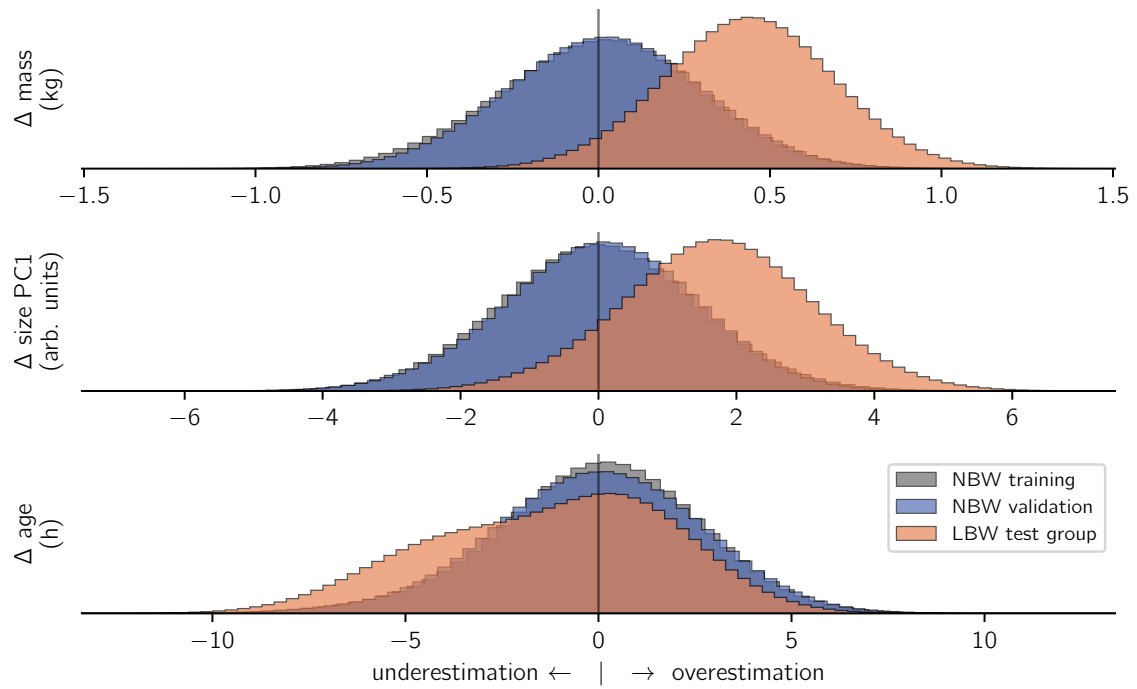


Figure 5.4: Predictive sampling of subject characteristics. For all included subject characteristics, models which were trained on NBW strides correctly inferred the training data (gray) and values from the validation set (blue). In contrast, the same models wrongly inferred the characteristics of LBW subjects (orange). The x -axes show the difference (Δ) between actual and predicted values per prediction. To facilitate comparison, histogram heights are again normalized per category.

Table 5.2: Age inference per LBW animal (compared to NBW average, last row). Δ : “inferred - actual” difference. Underestimation is defined as $\Delta < 0$, “count”: per stride, “rate”: per predictive sample. h : hours, std : standard deviation.

piglet	age	strides	underestimation		pred. mean Δ	pred. std
	h		count	ratio	h	h
b23	2.0	6	0	0.29	1.13	2.00
b15	2.9	5	0	0.37	0.68	1.96
b76	3.1	4	0	0.39	0.57	2.01
b74	4.2	7	1	0.40	0.52	1.97
1794.5	5.6	5	5	0.90	-2.57	1.99
b58	7.8	3	3	0.91	-2.85	2.00
b19v2	9.8	1	1	1.00	-6.14	1.99
b56	9.9	8	8	0.99	-4.58	1.96
<all NBW>	<3.8>	329	158	0.49	-0.03	1.95

strides from different individuals, we grouped the inferences per stride or subject and calculated the chance of over- or underestimating age. Of the 8 low birth weight subjects who contributed 39 strides, 4 individuals were consistently underestimated (Tab. 5.2). Consistently means that more than 75% of all predictive samples were below actual age, and that the ages for a majority of strides were on average underestimated. The magnitude of underestimation was between two and five hours. Curiously, those were the individuals recorded at a slightly higher age (> 5 hours). Overestimation in the other four LBW individuals was also consistent, but less so (less extreme underestimation rate, mean $\Delta < 2$ h). Standard deviation of the estimates did not vary across individuals or birth weight categories.

We conclude that underestimation of age is consistent over multiple strides of the same individual, and thus individual-specific.

5.5 Discussion

Birth weight variability in piglets is considerable. The average birth weight of a new born piglet in our data set is just above a kilo, yet the span within a litter is typically above 800 *g*. Size ranges are accordingly high. Low birth weight is often associated with low vitality (Baxter et al., 2008; Hales et al., 2013; Muns et al., 2013), and this supposedly correlates with deficient locomotion. Previous studies reported behavioral, physiological, and other differences for piglets of low body mass which could affect locomotion (Alvarenga et al., 2013; Muns et al., 2013; Quiniou et al., 2002; Roelofs et al., 2019; Vanden Hole et al., 2018a,b). One might therefore anticipate differences in how these newborn animals move their more or less heavy bodies, which is the immediate outcome of motor control. Top-down, direct, visual assessment could justify the hypothesis that LBW walking kinematics are somehow different from “normal” (D’Eath, 2012). Yet that is (i) hard to assess due to high behavioral variability and (ii) trivially expected given the adaptation to different physical properties of their body: gravitational force is a predominant constraint of locomotion, and it simply scales with animal weight (Aerts et al., 2023b). The overarching question is whether low or critically low birth weight can be quantitatively associated with a deficit in locomotion.

However, we observe little qualitative difference in LBW and NBW kinematics (Fig. 5.3). There is a small difference when averaging hindlimb posture, which might be trivially explained by heterogeneous group composition (but see below). This is further supported by the fact that our models were unable to correctly retro-infer either mass or size from blindly provided kinematic parameters, including coordination. Strides from LBW animals were estimated to come from an animal with a “normal” mass and size (Fig. 5.4). This suggests that low body mass or small size can not be causal for altered 2D kinematics, and it raises doubts whether there are any deficits in coordination and control. Within the range of kinematic variability, we observe LBW animals perfectly *capable* of normal posture, coordination, and overall stride results (collective variables). Note that this does not rule out problems of balance, stability, or endurance.

We do see differences for LBW compared to NBW recordings, nonetheless. First, a difference in sample size. There are much fewer valid LBW strides in our data set: only 39 of 368 observations are LBW. This could be interpreted as evidence for a lower capacity of LBW to produce normal locomotion (i.e. they succeed in this locomotor task less frequently, despite equal potential). Yet there are proximal, trivial explanations: based on conventions, the 10% lower quantile of birth weights in a litter is considered LBW, and there is a hard cap of 800 *g*. The resulting share is equal in our training set for video tracking, and in the final data set, because of pseudo-random, opportunistic sampling on-site (i.e. recording work was permanent, yet determined by farrowing and feeding of the subjects). The minority of LBW training videos might lead to an under-learning of those animals, reduced landmark tracking quality and therefore an exclusion bias for “non-normal” individuals. Though it

seems unlikely, we cannot rule out reduced locomotor endurance in LBWs, because the present data set is unsuited to count the occurrence of locomotor behavior. On the other hand, the strict stride filtering criteria for “good” kinematics may have involuntarily filtered out deficient individuals. Our conclusion that low birth weight individuals are non-deficient is strictly tied to the definition of the low birth weight category, which is herein based on weight criteria and did not regard phenotypic indicators of intra-uterine growth restriction (which we did not record, *cf.* Amdi et al., 2013).

A second difference of LBW locomotion is that the age is underestimated for strides of some, but not all individuals. Prediction is consistent per individual, although no subject characteristics except sex entered the model (“blind” inference). This supports the hypothesis that locomotor development is sometimes delayed in LBWs. Delayed development does not necessarily corroborate the hypothesis of locomotor deficiency in LBW: we would expect truly deficient strides to be substantially different from the data trained to the model, making it “unpredictable” (i.e. higher variance of posterior samples). Instead, the predictions are consistent for repeated measures of an individual, without notable increase in variance. For the affected subjects, we can even quantify a delay of less than five hours, which could nevertheless be critical given the rapid maturation of locomotor behavior in this species (Vanden Hole et al., 2017) and the importance of postnatal competition (Litten et al., 2003).

Note however that the causality might be inverse. We measured age underestimation only in the case of the individuals which were recorded late within our sampling time frame (age > 5 h, Tab. 5.2). This is consistent with prior evidence that energy reserves are depleted after birth (normal locomotion) but not replenished after four hours in the case of LBW (Le Dividich et al., 2017; Vanden Hole et al., 2019). Assuming that this is the case, i.e. energy reserves are depleted, we would expect two effects: (i) the animal might succeed in the locomotor task less frequently (not quantified, see above), and (ii) the kinematics might be altered, which we observed. The present model was trained from NBW data and thereby tuned to kinematic development from animals with normal energy levels (therefore it can correctly infer age from kinematics in NBW). That same model quantifies the potentially energy-deficient animals as younger. It might be that energy deficiency co-incidentally causes effects which are exactly opposite to the changes that piglet kinematics undergo in normal development. However, the more likely explanation is an actual delay or a temporary halt in development. Failure of the LBW to compete in the first hours are sufficient to explain reduced intake (Amdi et al., 2013), the absolute size and mass difference alone might be crucial, and an immediate question which we cannot address with the present data set is whether and how (fast, likely) animals would recover from the delay. Alternatively, there might be a technical artifact in probabilistic sampling (“shrinkage”, *cf.* Gelman et al., 2020) which caused the underestimation of “above average age” individuals. Yet this is an unlikely explanation, given that shrinkage would apply

equally to NBWs, and inversely to the opposite, younger subjects. Finally, with only eight LBW individuals, we cannot exclude co-incidence in which individuals are affected.

Neither of these technical explanations puts doubt on the clarity of the initial finding: a subset of the low birth weight individuals produced locomotor behavior which is quantitatively similar to that of younger individuals.

A corollary question is which patterns in the kinematic variables cause the different age inferences. We report high magnitude (but also highly variable, i.e. “non-significant”) slopes inferred from the age model (Tab. 5.1 and Fig. 5.6). Note that these slopes solely reflect effects within the NBW data subset. We also observed slight differences in the average hindlimb dynamic posture (Fig. 5.3). In fact, a more flexed hindlimb is typical for the youngest animals of both birth weight categories. This is further supported by the observation that hindlimb clearance and mean hip angle have a pronounced effect in the age model (Fig. 5.6). We emphasized potential differences in group composition to explain that (e.g. sex effect in the “size” model), and different age per group might be a proximal explanation for the non-normal age inference in LBW. However, the average age of LBW animals (5.3h) in our data set is nominally above that of NBW (3.8h), which is a discrepancy with the age underestimation. Yet if we assume that the hypothesis of delayed locomotor development is correct, the nominal age would be misleading, and LBW effectively behave similar to younger animals. This can explain the apparent discrepancy in age group composition and age predictions from kinematics. Though many other parameters also entered the probabilistic model and influenced the model outcome, our data suggests that dynamic posture, especially of the hindlimb, might be the major proxy for perinatal maturation (Fig. 5.6)

The findings discussed above are enabled by “inverse” modeling of subject characteristics as a function of kinematic parameters, using probabilistic models. The models do reveal various parameter associations, yet the top down (repeated) testing with the chosen model structure complicates drawing definite conclusions (e.g. we observed a sex effect on size, but opposite of what other studies have reported, Baxter et al., 2012, we conclude males in our study were just smaller by chance). Instead, the demonstrated predictive modeling strategy (Shmueli, 2010) leverages the potential of probabilistic models to perform out-of-sample prediction (via separation of the LBW test group). This strategy is limited by a major asymmetry: in case the prediction shows differences, they are most likely robust, yet “non-effects” could be caused by unfavorable dataset characteristics (Ch. 5.6). Non-probabilistic modeling could equally serve to predict values, but it cannot generate parameter distributions (as in the bimodal age predictions, Fig. 5.4). Our probabilistic models implicitly regard “non-significant” parameter slopes (Fig. 5.6), which are usually neglected in standard hypothesis testing (such as the high magnitude, highly variable joint ROM quantities, which might nevertheless have relevance for age prediction in case they

are correlated and add up). The data transformations and predictive modeling strategy we applied herein hold further potential for inferring kinematics, for example by morphometric bracketing of extinct taxa. For that purpose, one would train a model to infer dynamic posture and coordination from a given range of morphometrics, predict samples for specific morphometrics in that range, and convert the samples back to (3D) joint angle profiles which could be animated.

There is a conceptual hierarchy, but no clear causality, with regard to modeling parameter dependencies in quantitative studies of locomotion. For example, animals might increase the hip angle (posture) and the temporal pattern in the joint angle profiles (coordination) to reach higher speeds (spatiotemporal gait variables), but the speed they reach might depend on age (i.e. maturation, a subject parameter). But maturation and age might also influence the speed without changing dynamic posture or coordination, simply because the animal grows and increases strength. Age might also affect dynamic posture and coordination directly, for example if dimensionless speed stays constant but clearance changes with age. Reducing clearance can increase speed (less unnecessary hoof lifting) or reduce it (higher duty factor), which can be distinguished by the other parameters. This complex interrelation of spatiotemporal quantities complicates the intuitive modeling strategy, which involves using subject characteristics as a factor to predict a multivariate block of spatiotemporal gait variables and kinematics. In our experience, model residuals in such models are high, multiple testing can yield putatively co-incidental significances, cross-dependencies within the multivariate data set might be underestimated, and sample size requirements are high. And even under ideal circumstances: chances are that such models would yield *some* age effect, even with random data.

In contrast, the strategy applied herein is related to the question: “given the complete kinematic output of a behavior, can we infer subject characteristics of the animal producing it?” We used probabilistic, predictive models, which are able to capture intrinsic variability of the process, and addressed specific categorical questions (NBW/LBW differences) by out-of-sample predictions. We demonstrated that, in the first ten hours of piglet life, (1) kinematics seem to be indifferent to low body mass and size, and (2) locomotion of some LBW individuals could be explained by a delay in locomotor maturation which is initiated *post partum*.

5.6 Supplementary Analysis

The key findings in this chapter are based on an analysis which uses the statistical model of one subset of the data for predictive inference on another subset, a strategy which demands justification.

- Is the model actually *capable* of finding differences?
- Under which circumstances will group differences arise?
- Is there a bias in the modeling procedure, for example to not find all differences?

Answering these questions requires a rather technical, in-depth analysis of the model outcome.

One important constraint is that the models herein are linear, i.e. of the form

$$y = a + b_1 \cdot x_1 + b_2 \cdot x_2 + \dots + b_m \cdot x_m + \epsilon = a + \sum_k b_k \cdot x_k + \epsilon$$

Herein, y is one of the outcome variables (observed; age, size, or mass), a is the intercept, b_i are the slopes, x_i are the different input parameters (observed; spatiotemporal gait variables, dynamic posture, and coordination measures), and finally ϵ is the model residual (unexplained variation, including for example measurement inaccuracy). As I have demonstrated elsewhere (Mielke, 2024), there are some **criteria** which are necessary (but not sufficient by themselves) for the model to predict a difference. Given there is a...

1. significant difference between the study groups in the distributions of input variables
2. high enough slope magnitude (“steepness”)
3. limited model complexity/size, i.e. moderate number of (other) slopes
4. low residual variation
5. sufficient sample size

... then we can find differences. An example where all these conditions are met is the age model, above. In case one of these conditions is not met, a difference between the study groups might be obscured (yet note that the criteria are on a continuum, some cases might not be clear-cut). From theoretical considerations alone, this situation is biased. Differences in out-of-sample prediction must have passed the criteria above, and are therefore robust. Yet on the other hand, “*the absence of evidence is not the evidence of absence*”: if for example the sample size is low, or the residual variation is high, the model’s prediction might miss an actual effect. This situation is somewhat analogous to classical hypothesis testing, where we quantify the chance of falsely rejecting a null hypothesis such as “groups were samples from the same distribution”.

The available modeling framework enables some extra insights on why the prediction of piglet subject characteristics appeared as it did. Firstly, we can compare the input distributions (criterion 1, above) of the observational groups, to identify those model parameters which actually influence the predictions (Fig. 5.5). Secondly, we can calculate the absolute effect of each parameter on the predicted outcome (slope magnitude, criterion 2), by multiplying the parameter magnitude with the according slope (Fig. 5.6; this works because $\Delta y \sim b_i \cdot \Delta x_i$). Thirdly, we can quantify the model residual (criterion 4; Fig. 5.6, last row). The other criteria (model complexity and sample size) are predetermined in this modeling application, and have an unknown influence on the outcome.

For the **age model**, we observe a low residual variance, and there are two parameters which are significantly different between NBW / LBW *and* have a high slope (those are hindlimb clearance and hip angle). Notably, there is hardly any difference in the observed coordination parameters. For the model of **body mass**, the same parameters seem to have relevant slopes, yet the model residual is much higher and likely prohibits conclusive differences in the probabilistic prediction. Other slopes seem to have significant effect in the NBW group (hip and tarsal eROM), yet the input distributions (i.e. NBW and LBW observations) do not differ, and so their effect is irrelevant for prediction of mass. For the **size model**, we find some slopes with a negative effect on the output (hindlimb clearance, head angle, hip angle). However, others with a positive effect (knee angle, tarsal eROM) possibly cancel this out for the overall prediction. Like the previous one, this model contains significant slopes with nonetheless no difference in the input distribution. We also compared the relative widths of effect size distributions, and found that “dimensionless (relative) speed” is most variable among trials: all animals use a variety of speeds, and apparently this can not be systematically associated with age, size, or mass. It must be re-iterated that “significance” in these analyses merely provides a hint to guide the eye: the combination of a minor (non-significant) difference in an input parameter with a notable, only “almost-significant” effect size might still affect the probabilistic prediction. Finally, the input distributions (Fig. 5.5) are insightful by themselves for NBW / LBW comparison. Some LBW parameters are bimodally distributed, which might be related to the prediction outcome for the age model. We observe most differences in spatiotemporals (predominantly hindlimb-related), as well as in measures of dynamic posture and range of motion. However, there are few significant effects of coordination parameters, and the remaining ones are cancelled out by a lack of

Figure 5.5: Observed model parameter input distributions (*next page*). For each model parameter (rows), observations are indicated by tick marks for the two study groups (NBW: gray, LBW: orange). Values were standardized on population level (i.e. mean-centered and scaled by standard deviation). Significance between the distributions (indicated by asterisk) was determined by a two-sample rank sum test.

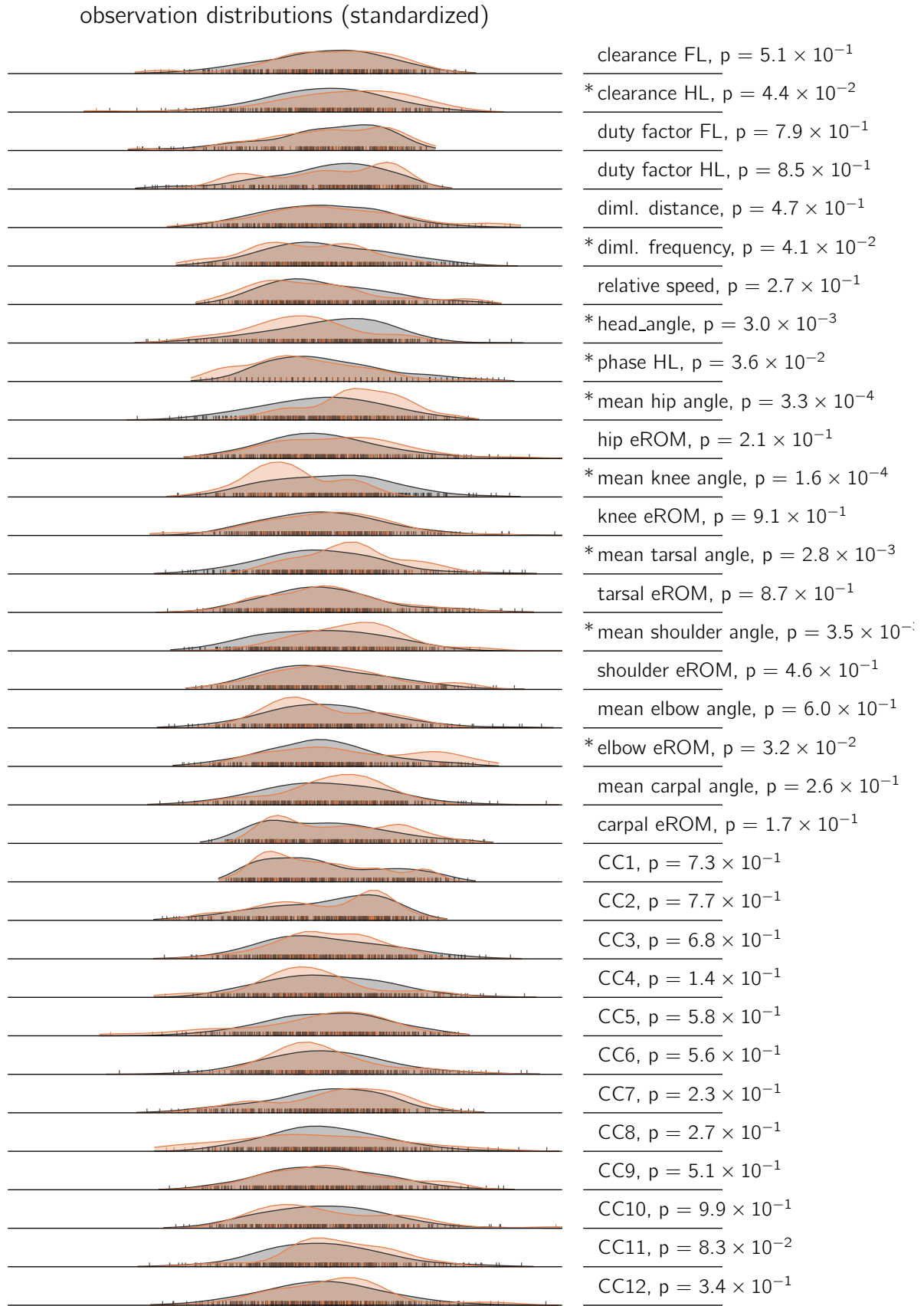


Figure 5.5: Observed model parameter input distributions. See previous page.

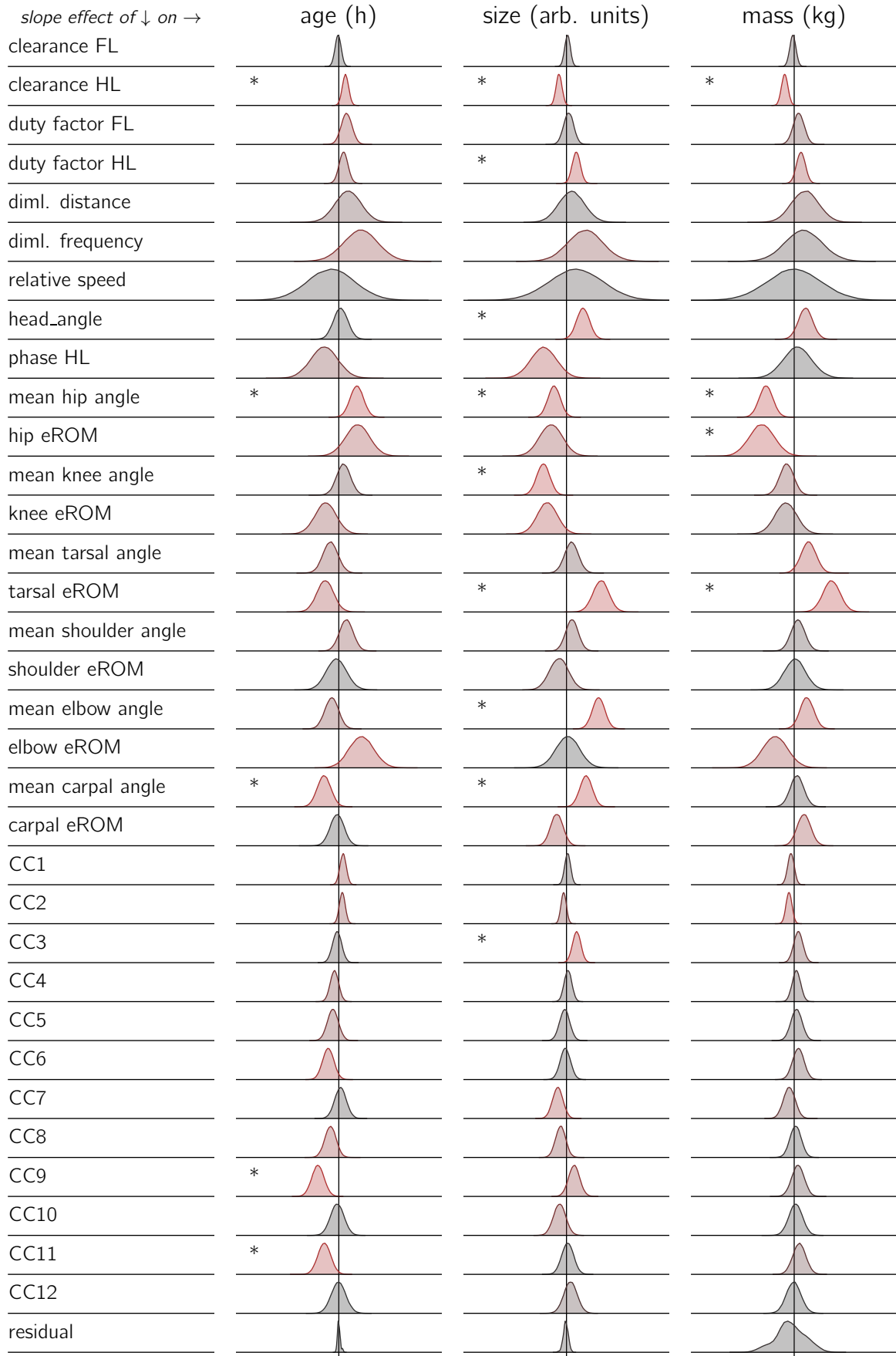


Figure 5.6: Effects of each input parameter on the model output. See next page.

Figure 5.6: *Effects of each input parameter on the model output* (previous page). Effect size is calculated as the data range (difference between 2.5% and 97.5% data quantiles) multiplied by the slope distributions (from model sampling). Effect magnitude and therefore relevance is visualized by red shading, effects which are different from zero (95% HDI) are also marked by an asterisk. The residual (last row) is determined as the difference of the actual parameter observations and the model result.

difference in the observed coordination of LBW and NBW.

To summarize, we show that a detailed inspection of on the one hand the observed model input parameters, and on the other hand the effect “lever” of each parameter slope, warrants some caution about non-different predictions (as in the models of size and mass). In those models, we cannot rule out that a larger sample size or an adjusted, potentially more selective model design might elucidate quantitative NBW/LBW differences. On the other hand, inspection of the predictive age model confirms that finding differences by out-of-sample prediction is possible, and we could even relate those findings to specific model parameters.

5.7 Acknowledgements

The author would like to thank Miriam Ayuso for organizing and participating in the recording sessions, as well as Laura Buyssens, Georgios Petrellis, Gunther Vrolix, Charlotte Vanden Hole, Denise Vogel and all students who joined for help during recordings. A comment by Ann Hallemans stimulated the extra analysis, from which this chapter benefits a lot; thank you! Maja Mielke provided valuable comments on the manuscript text, in fact for all three major versions of it, and I am very glad she did.

Part III

Inverse Dynamics

Chapter 6

Inverse Dynamics Workflow

6.1 Abstract

This thesis has so far explored the measurement and modeling of movement of (terrestrially) locomoting animals. A topic which remains to be discussed is the *cause* of that movement. The segments of animal limbs can be modeled as rigid bodies, each of which has a specific momentum. Momentum is a conservative quantity, i.e. it changes due to energy transfer with its surrounding. The retrospective inference of reasons for momentum change is often labeled “inverse dynamics”.

The field of (inverse) dynamics is subject to various conventions and some methodological flexibility. The purpose of this chapter is to set the table and clarify conventions used further on in this thesis. The physical fundamentals and general algorithms applied in inverse dynamic calculations are presented, as well as elegant mathematical tools to implement them (wrenches, quaternions). To illustrate the basic principles of physics which come to play in this chapter, a hypothetical collision experiment is discussed. It turns out that, whereas most quantities in inverse dynamic calculations can be accurately measured, the inertial properties of segments are the Achilles heel of accurate dynamic modeling.

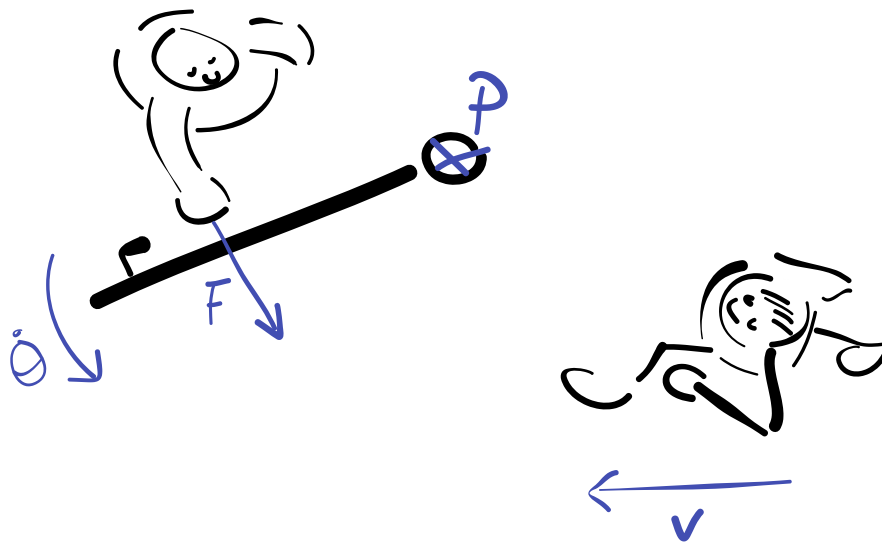


Figure 6.1: A *hypothetical, scientific experiment* to illustrate dynamic modeling (see text). Upper left: experimenter setting a door in motion. Lower right: subject running towards the experimental setup. $\vec{P}$: hinge point/axis of the door, $\vec{v}$: velocity of the subject, $\dot{\theta} = \frac{d\theta}{dt}$: angular velocity of the door, $\vec{F}$: force applied to the door by the experimenter.

6.2 Introduction

6.2.1 A Momentum Experiment

“Slamming a door in someone’s face” is not a friendly thing to do. Yet “friendliness” is not a scientific category, and thus we must consider this as a purely physical process (Fig. 6.1). If a person in question is running at constant velocity (which we could quantify using Fourier methods and probabilistic modeling, see before), the unexpected appearance of an obstacle will certainly force the subject to change trajectory. If the unfortunate subject is unable to adjust on time, physicists would call this a “collision experiment”. In most probable cases the impact will thus **change the subject’s momentum**. Momentum quantifies the linear and angular movement of a body. The non-sliding door itself, set up on a hinge (point $\vec{P}$ in Fig. 6.1), will gain rotational momentum through the act of slamming, performed by an experimenter. In the collision case, the subject might (i) decelerate but proceed, (ii) come to a hold, or (iii) ungently bounce backwards from the elastic collision. The door will thus transfer momentum to the subject, thereby ultimately canceling out (part of) the momentum of the person if the impact is approximately equal in magnitude, but opposite in direction, of the running. Trained experimenters might even succeed in (iv) transferring angular momentum to the surprised subject, generating a spin (though attention must be paid to whether that spin is partially caused by an evasive motor response of the subject). All these possible outcomes are manifestation of the **conservation of energy and momentum**, which I will explore on the metaphor of a hypothetical door slamming experiment. Due to the numerous involved variables, it is not clear *a priori* how exactly a subject will be hit, and what the exact consequences of the collision are. Predicting the outcome of such scientific experiments is enabled by a major field of biomechanics, which is called **dynamic modeling**.

6.2.2 Kinematics and Kinetics

In the next chapter (Ch. 7), I will discuss a specific contribution to dynamic modeling, namely the extraction of inertial properties from CT scans. This chapter serves to illustrate the overall context and to establish the connection with the previous topics of this thesis. It answers the question of the relevance of CT-based inertial property quantification, by providing a superficial overview of the dynamic modeling technique.

Until this point, this thesis focused on methodological development in terms of measuring, modeling and predicting *kinematics*. With Fourier-based methods and probabilistic models in place, one can (i) quantify measured kinematics, and (ii) simulate locomotion by predictive sampling of trained models. This is a complete quantification of *how* limb segments of animals change their position in space. Segments could be axial or peripheral parts of the animal which move as a unit (rigid- or pseudo-rigid bodies, hence the general term “rigid body dynamics”), which is to a large extent analogous to a door on a hinge.

Having cleared the *how* in previous chapters, the other relevant aspect of locomotion is *why* segments move. This opens up a rabbit hole of elementary physics (“Erhaltungssätze”, *sensu* Noether, 1918):

A rigid body will not change its linear or angular momentum if there are no external forces or external torques applied to it.

This, again, is a formulation of the conservation of momentum (NASA, 2021), as introduced above.

These things seem quite intuitive: the slamming door will affect the subject upon impact. And because there must be movement before or after the event of interest, and because that movement changes (i.e. it is “dynamic”), the study of the conservation of momentum in moving rigid body systems is called “dynamics”. In contrast to kinematics, dynamics are sometimes also called “kinetics”, the study of motion and its causes. “Conservation” as it is understood in physics implies that the measure of interest (momentum) cannot be created or lost, only gets converted to another form (e.g. kinetic to thermal energy by friction) or transferred to another object. Thus, dynamics are also linked to the calculation of transformations and transfers of energy from one form or object to another (e.g. mechanical energy transfer, dissipation of heat). These conversions and transformations are precisely the reason *why segments or subjects change their movement*, and we seek to quantify and model (i.e. predict) them.

In applied terms, which algorithm or technique would we apply to quantify and model such an abstract thing as “momentum”? There are many, all useful under different circumstances. I will present herein the methods which seemed most useful and intuitive to me, and I will explain “why”. Yet this does not imply that they are generally preferred by all experts in the field.

In a nutshell:

- Wrenches and quaternions are useful computational tools for calculating rigid body dynamic models.
- Balance Equations are the mathematical manifestation of conservation of momentum.
- Reference frame considerations can simplify calculations; however, static frame calculations are sufficient and result in correct joint moments.
- Computed Tomography might be a shortcut to approximate structural properties which are crucial for the calculation of dynamics (next chapter).

Once more, the herein summarized work is physics textbook knowledge, re-formulated to be more readily applied by the author to animal locomotion; as mentioned above, I consider this necessary context. This chapter will summarize the theory behind inverse dynamics, and point to literature references and conventions which turned out to be useful in preparation of this thesis.

6.3 Dynamics: Theory

6.3.1 Forward and Inverse Problems

Since the “slamming door” experiment (Fig. 6.1) could be unpleasant for the faces of living subjects, one might be tempted to measure on a robot.

When constructing a robot with motors and actuators at the joints, the first problem is to set the appropriate joint torques to generate a desired movement. This is called the forward problem (Lynch and Park, 2017). Conversely, **inverse dynamics** is a second problem which describes the inference of joint moments which could have possibly generated an observed movement. For example, from observing (i.e. measuring) a hinge door moving at a certain angular velocity, one can calculate the force and moment the experimenter has applied previously to set it in motion. The study of animal locomotion is mostly concerned with this second, inverse situation. The central quantities are “forces” and “moments”, the latter being a physical quantity related to forces and involved in rotational movement. I will approach forces and moments separately, before joining them in a common framework, by first explaining how force is related to linear momentum (a measure for the motion of a system), how this is analogous to moment and angular momentum, and finally how these concepts are applied in practice.

6.3.2 Forces

Forces are the product of mass and an acceleration, $\vec{F} = m \frac{d^2 \vec{x}}{dt^2}$ (or, more familiar to most, $\vec{F} = m \vec{a}$), which is one of Newton’s laws of motion (Newton, 1687; Tipler and Mosca, 2007). In fact, force is better expressed as the derivative of linear momentum $\vec{p}$ with respect to time, hence $\vec{F} = \frac{d\vec{p}}{dt}$ (total derivative). They are vector quantities (magnitude, direction) and can occasionally cancel out (superposition). Linear momentum is defined as follows: $\vec{p} = m \frac{d\vec{x}}{dt}$, with m the mass and $\frac{d\vec{x}}{dt}$ the linear velocity (change of position over time).

Combining these equations and applying the total derivative:

$$\frac{d\vec{p}}{dt} = m \cdot \frac{d^2 \vec{x}}{dt^2} + \frac{dm}{dt} \cdot \frac{d\vec{x}}{dt}$$

One can apply the inverted version of the mathematical “chain rule” (hence: “unchain rule”):

$$\frac{dm}{dt} = \frac{dm}{d\vec{x}} \cdot \frac{d\vec{x}}{dt}$$

Which, plugged back into the derivative of linear momentum, gives the complete force formula (here in the scalar form).

$$\vec{F} = m \cdot \frac{d^2 \vec{x}}{dt^2} + \frac{dm}{d\vec{x}} \cdot \left(\frac{d\vec{x}}{dt} \right)^2 \quad (6.1)$$

Because $\vec{a} = \frac{d^2\vec{x}}{dt^2}$, the first part of the right hand side expression in equation (6.1) is the familiar $\vec{F} = m\vec{a}$. The second part is more enigmatic: those are “velocity product terms”, which involve partial derivatives of the mass matrix¹, thereby accounting e.g. for configuration changes of a chain of segments, deformations or mass shifts (e.g. muscles, blood flow, ...). Assuming here that mass distribution of single segments is constant or changes are comparatively small (which is generally an inaccurate simplifying assumption), this part is usually set to zero.

Hence, we get a more or less familiar equation for the Force $\vec{F}$, which tells us that the acceleration of a mass ($m\frac{d^2\vec{x}}{dt^2}$), i.e. a change in linear momentum, must be equal to an external force. $\vec{F}$ is a three element vector.

6.3.3 Moments

In the door slamming experiment, and experimenter must apply force to the door in order for the experiment to “succeed”. Yet the movement of the door depends on where the force is applied (Fig. 6.2). Consider slamming a door by applying a force close to the hinge. This either requires much higher force, or yields much less spectacular experimental outcomes than applying force to the distal side of the door (furthest from hinge, e.g. at the handle). In fact, this simple physical observation explains why doors have evolved handles in the place where we observe them in extant doors.

Otherwise, the calculation of angular quantities is largely analogous to the linear case (Lewin, 1999b; Tipler and Mosca, 2007). Mass, in the force equation, is the resistance of an object to linear acceleration, i.e. to change of linear momentum $\vec{p}$. Analogously, Mass Moment of Inertia² (I) is the resistance of an object to rotational acceleration, which induces a change in angular momentum $\vec{L}$. From experience, one can tell that lifting a door out of its hinge is a different process than conventionally rotating it, though few have carried out this comparison in the absence of gravitation and friction. And so, the series of equations above can be applied analogously to angular momentum $\vec{L}$ (with angular velocity $\frac{d\theta}{dt}$).

$$\vec{L} = I \frac{d\theta}{dt}$$

The time derivative is as follows.

$$\frac{d\vec{L}}{dt} = I \cdot \frac{d^2\theta}{dt^2} + \frac{dI}{dt} \cdot \frac{d\theta}{dt}$$

The time derivative of angular momentum is called moment ($\vec{M}$; see also:

¹Mass is often taught as a scalar quantity, yet in more complex problems of composite systems, which have a mass distribution, it can be more useful to represent mass by a matrix/tensor. This is analogous to the mass moment of inertia.

²Note that I is a matrix, usually 3-by-3; this is analogous to mass, which in its general form for complex objects is also a matrix (Lynch and Park, 2017).

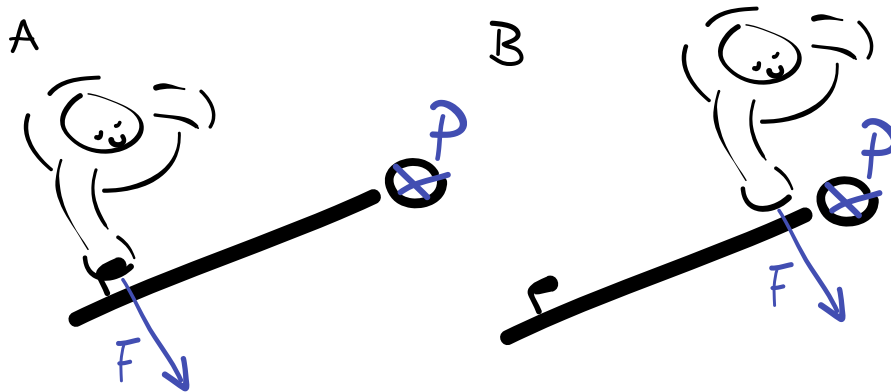
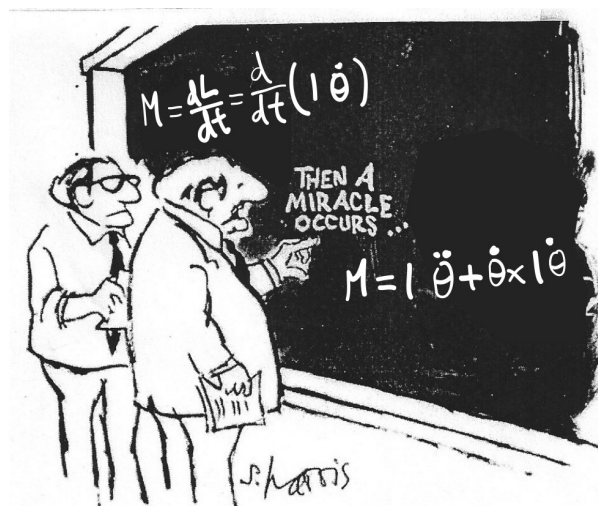


Figure 6.2: Force and Moment. A force applied further away from the hinge $\vec{P}$ (situation A) will generate a larger moment than a force similar in magnitude, but applied more proximal to the hinge (situation B). The reason is that the moment is defined as the cross product of the position vector of the force application point (in a hinge-centered reference frame) and the force vector. Furthermore, because the change in angular momentum is related to that moment, the door will gain a higher momentum (and will therefore “hit harder”).



“I think you should be more explicit here in step two.”

adapted from *What's so Funny about Science?* by Sidney Harris (1977)

Figure 6.3: Derivation of Newton-Euler equation of rotational motion.

Wikipedia/Newton-Euler equations (2012); Wikipedia/Torque (2012)), and with further manipulation (details in Lynch and Park, 2017, p. 242, eqn. 8.23; otherwise see Fig. 6.3) yields:

$$\vec{M} = I \frac{d^2\theta}{dt^2} + \frac{d\theta}{dt} \times I \frac{d\theta}{dt} \quad (6.2)$$

Again, there is an (angular) velocity product term which we hope to sufficiently well approximate by setting it to zero (rigid body assumption, i.e. no or negligible change of I). Just as in the linear case, we identify a cause for any change in angular momentum, and call it the moment $\vec{M}$ (others may call it torque $\vec{\tau}$). In this case, I used the vector formula notation (explicit by the cross product; this could analogously be done for the force, equation (6.1), above). There is another extra step compared to the linear momentum: calculation of the **angular momentum requires a reference point** (usually the hinge, $\vec{P}$, as above, but any point could be chosen). The calculated angular momentum actually changes depending on which reference point is chosen. Returning to the example of the door, rotating on a hinge (Fig. 6.4). We are free to approximate the door by a point mass in point $\vec{D}$, to see in which way the results deviate from common experience. Angular momentum $\vec{L}$ relative to point $\vec{P}$, which we denote $\vec{L}_{\vec{P}}$, is calculated by using the cross product of the vector from point $\vec{P}$ to $\vec{D}$ ($\vec{r}_{\vec{P}\vec{D}}$) and the linear momentum $\vec{p} = m\vec{v}$:

$$\vec{L}_{\vec{P}} = \vec{r}_{\vec{P}\vec{D}} \times \vec{p} = m\vec{r}_{\vec{P}\vec{D}} \times \vec{v}$$

From cross product properties, we see that the direction of $\vec{L}$ does not change during the slamming (stays parallel to the hinge axis), and the magnitude stays constant and greater than zero if velocity is greater than zero.

In contrast, the point mass door has a different angular momentum when calculating it relative to point $\vec{D}$: the position vector $\vec{r}_{\vec{D}\vec{D}}$ becomes zero, so $\vec{L}_{\vec{D}} = \vec{0}$ at all times. This is non-intuitive: the door rotates, after all; yet remember that we are reducing it to a point mass here and $\vec{L}_{\vec{D}} = \vec{0}$ only holds for the point mass case! Other points and non-point-mass-objects require more complex calculations, except maybe in the special case where rotation happens about the center of mass ($C\vec{O}M$). In that special case, we speak of “spin angular momentum”, which is an intrinsic property of the object in a sense that we do not need to give a reference point any more because everyone knows that the $C\vec{O}M$ is the reference point for spin angular momentum.

But the general idea is that one needs to keep track of a reference point for angular momentum, and thus also for all calculations regarding angular momentum or mass moment of inertia.

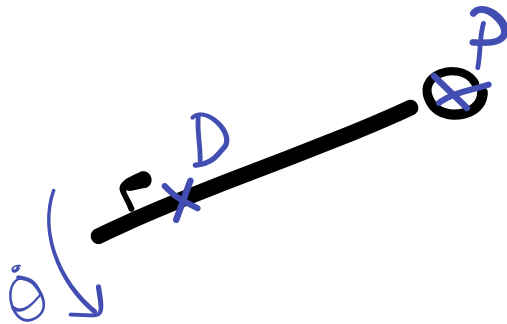


Figure 6.4: *The slamming door, isolated sketch.*

6.3.4 Screw Theory and Wrenches

The formulas for force and moment above resemble each other. In fact, a common trick to help novice physicists remember the angular formulas is to replace position $\vec{x}$ by θ (and derivatives accordingly), and mass m by mass moment of inertia I (Tipler and Mosca, 2007). The force-moment-analogy is no co-incidence. Force is a 3D vector, acting at a point. Moment is a 3D vector with a reference point. The moment is in fact just the corresponding force acting at a lever. And the lever is usually the same position vector used to calculate angular momentum; in other words, given that we stay in the same reference frame, we might just need the force and the reference point, and the moment follows.

Would it not be useful then to join force and moment?

It would. There is a framework to unify the linear and angular components of momentum, which is called **Screw Theory** (Ball, 1876; Lynch and Park, 2017).

The trick is to concatenate the vectors $\vec{F}$ and $\vec{M}$ to form a six element vector $\vec{W}$, **the Wrench** (Fig. 6.5, cf. https://en.wikipedia.org/wiki/Screw_theory). “6D” might seem unhandy, and one may ask “why”. The pure and simple answer is computational convenience (Dumas et al., 2004; Müller, 2018): the wrench saves us from double computation, by linking the two intrinsically related parts of momentum.

Any modern scripting language used to solve inverse dynamic questions can easily work on six element arrays/vectors instead of three-dimensional ones. A mathematical shorthand for writing wrenches are square brackets and a semicolon indicating that the (column vector) components are concatenated (to form a longer column vector):

$$\vec{W} = [\vec{F}; \vec{M}]$$

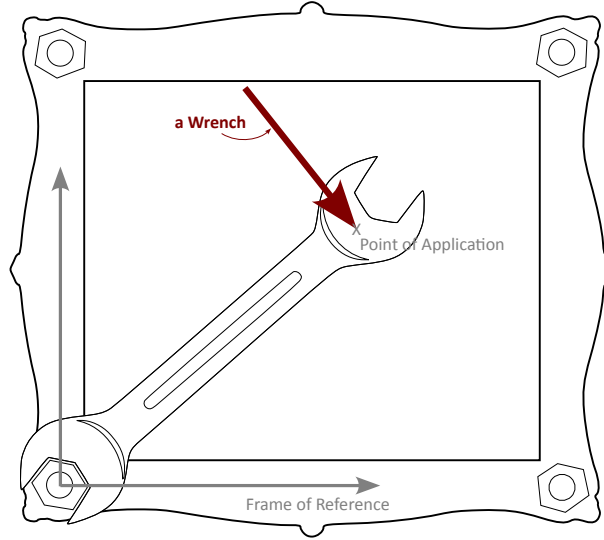


Figure 6.5: Wrench notation. A Wrench is a combined vector of force and moment, consisting of six elements, and requiring a point of application. Wrench vectors are depicted as arrows, yet in contrast to forces, their tip points towards the point of application (they can be moved around, adjusting the moment component; see text). Note that this figure is actually a 3D sketch, yet the limitations of a printed PhD thesis prohibit illustration beyond the paper plane projection.

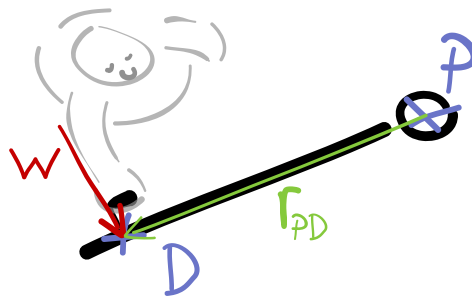


Figure 6.6: The door slamming experiment in wrench notation (see text). Point $\vec{P}$ is the origin of the reference frame, $\vec{r}_{\vec{P}\vec{D}}$ is the position vector to the attack point $\vec{D}$ of the wrench $\vec{W}$.

6.3.5 Balance of Wrenches

As with force and moment, we can equate a dynamic wrench to external wrenches in the calculation of a “balance of wrenches”. This can be achieved by, again, vertically concatenating force and moment balances.

For one, we get a “dynamic wrench” on the left hand side of the balance. Plugging in the non-zero part of equations (6.1) and (6.2) from above, we get an expression for what can be referred to as the **dynamic wrench** (defined at the $C\vec{O}M$):

$$\vec{W}_{dyn} = \left[m \frac{d^2 \vec{x}}{dt^2}; I \frac{d^2 \theta}{dt^2} \right] \quad (6.3)$$

Wrenches can be added up, for example the dynamic wrench can be easily shifted (vector $\vec{s}$) to a different reference point, which does not require extra adjustment of the mass moment of inertia:

$$\vec{W}_{dyn}^* = \vec{W}_{dyn} + [\vec{0}; \vec{s} \times \underbrace{(m \frac{d^2 \vec{x}}{dt^2})}_{\vec{F}}] \quad (6.4)$$

The $\vec{x}$ and θ in the dynamic wrench refer to the movement of the rigid body itself and are immediately derived from the object’s kinematics. The inertial properties m and I have to be measured or approximated separately. The dynamic wrench expresses the rate of change of momentum of an object.

On the other hand side of the equation, we have any **external wrenches** acting on the object. The most important ones are the gravitational wrench, friction wrench, and the tip wrench (stay tuned).

$$\vec{W}_{\vec{g}} = [m\vec{g}; 0]$$

$$\vec{W}_f \approx [0; 0]$$

$$\vec{W}_{\vec{t}ip} = [\vec{F}_{\vec{t}ip}; \vec{M}_{\vec{t}ip}]$$

We can sum up all external wrenches to get a total external wrench $\vec{W}_{ext}$:

$$\vec{W}_{ext} = \sum_{ext} \vec{W}_i = \vec{W}_{\vec{t}ip} + \vec{W}_{\vec{g}} + \vec{W}_f + \dots \quad (6.5)$$

Any other known external forces and moments should be added in wrench form, keeping in mind that correct addition requires shift to the same reference point (usually the $C\vec{O}M$).

Equations (6.3) and (6.5) constitute the **Balance of Wrenches**:

$$\vec{W}_{dyn} = \sum_{ext} \vec{W}_i \quad (6.6)$$

In other words:

The rate of change of [linear;angular] momentum of a rigid body is equal to the sum of external wrenches acting on it.

6.3.6 Reference Frames

For simplicity, I ignored the issue of reference frames in all equations above. Just like reference points, reference frames need to be well documented. Transformation from one to another reference frame is possible by matrix multiplication with a transformation matrix. It turns out that reference frames are not required for the calculations above: one can and should simply stay in the inertial reference frame.

If one, however, decides to venture into the reference frames (as the author did), be aware that fictitious forces add to the balance. Interested readers may refer to a practical guide by the author (Mielke, 2020).

6.3.7 Quaternions

On top of the force/moment complexity comes the issue of quantifying rotation in space. Angular momentum depends on angular acceleration $\frac{d^2\theta}{dt^2}$, which is the second temporal derivative of “angle”. But getting that angle in 3D is not exactly trivial, and there are several complementary notations.

When we think about rotation in space, we might think about an axis and an angle (Fig. 6.7). And that is a good point to start.

$$\text{axis: } \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \text{ angle: } \theta$$

On the swinging door, the axis would be the line through the hinge points, and the angle could be measured as the radians separating open and closed state. Ideally, one would want to combine axis and angle into a single object $\vec{q}$, and it turns out that the following way is convenient (Challis, 2020; Flashner and McNitt-Gray, 2019; Pennestri and Valentini, 2010):

$$\begin{pmatrix} \cos(\theta/2) \\ x \cdot \sin(\theta/2) \\ y \cdot \sin(\theta/2) \\ z \cdot \sin(\theta/2) \end{pmatrix} = \vec{q} = \begin{pmatrix} q_w \\ q_x \\ q_y \\ q_z \end{pmatrix} \quad (6.7)$$

There also exists a conjugate, $\vec{q}^{-1}$:

$$\vec{q}^{-1} = (q_w, -q_x, -q_y, -q_z)$$

And this object $\vec{q}$ allows for convenient spatial rotation, by almost-simple multiplication: take any 3D point's position vector $\vec{D}$ with 4D coordinates³ $(0, a, b, c)$. This could be the door handle point from our example: after rotation, it will be $\vec{D}_{rot}$, which can be calculated by:

$$\vec{D}_{rot} = \vec{q} \cdot \vec{D} \cdot \vec{q}^{-1}$$

The dot “.” is just vector multiplication, and computational tools can handle this operation quickly and efficiently. Better yet, one can stack quaternions by (non-commutatively) multiplying them:

$$\vec{q}_2 \cdot (\vec{q}_1 \cdot \vec{P} \cdot \vec{q}_1^{-1}) \cdot \vec{q}_2^{-1} = (\vec{q}_2 \cdot \vec{q}_1) \vec{P} (\vec{q}_2 \cdot \vec{q}_1)^{-1}$$

Working with time series of quaternions ($\vec{q}(t)$, i.e. the angular position of an object changing over time), there is a formula for getting the angular velocity $\frac{d\theta}{dt}$ from the time differences $d\vec{q}/dt$ which fortunately even works with numeric differentials $\frac{\Delta\vec{q}}{\Delta t}$ of real measurements (Baker, 1999).

$$\frac{d\theta}{dt} = 2 \frac{\Delta\vec{q}}{\Delta t} \cdot \vec{q}^{-1}$$

A quaternion variant for the Kabsch algorithm exists (Kabsch, 1976; Kneller, 1991; Lawrence et al., 2019). This algorithm is useful to find the rotational difference between rigid bodies, including the change of rotation of a rigid body between two measurement times or the position of a segment relative to another.

Quaternion operations are usually included as functions in common quaternion toolboxes (for example in Matlab, R, and Python). And toolboxes also provide the capability of conversion to other notations of spatial rotation (Euler- or Tait-Bryan angles, Rotation Matrices).

Just like with wrenches, quaternions cause discomfort to untrained analysts because they seemingly leave the familiar space of 3D geometry. However, if one simply accepts that maths can handle 4D or 6D problems, the reward is computational convenience. Superposition and the generation of temporal derivatives are readily available. Quaternions also avoid some of the pitfalls of the alternative methods (gimbal lock problem, order dependence). They have an intuitive geometric interpretation through the axis-angle-analogy. And if anatomically relevant axes are required for communication, they can easily be transformed to such.

This does not imply that inverse dynamic calculations are impossible with Euler Angles. Yet in most applications, quaternions are mathematically elegant and efficient.

³The position vector $\vec{D}$ is transferred to 4D space (formally necessary for multiplication with 4D quaternions) by simply appending a “zero” as the first element; the quaternion multiplication ensures that $\vec{D}_{rot}$, though also 4D, always has a zero-valued first element, being effectively 3D.

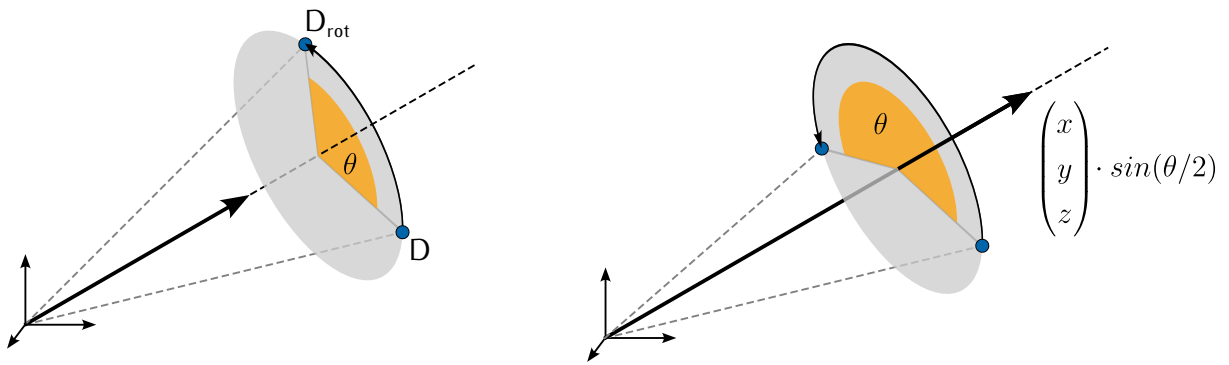


Figure 6.7: Spatial rotation with quaternions. A point $\vec{D}$ (blue, with gray dashed position vector) can be rotated by an angle θ around an axis $(x; y; z)$ to position $\vec{D}_{rot}$. The higher the rotation angle value (indicated by the orange area), the further the rotation. A quaternion (large black arrow) can encode both the axis and the angle in a 4D vector, in a manner which conveniently enables the execution of the rotation by simple vector multiplication.

6.4 Dynamics: Application

6.4.1 Generic Approach

The theoretical considerations above assemble several technical advantages which are worth summarizing.

- The Balance of Wrenches, equation (6.6), can be calculated entirely in the inertial reference frame, without repeated switch to joint coordinate systems (and return).
- Wrenches can easily be moved to other reference points via equation (6.4).
- A convenient point to calculate the balance of wrenches is the center of mass ($C\vec{O}M$); rotation about that point is an expression of spin angular momentum (an intrinsic property) and most components simplify (e.g. gravitational wrench $\vec{W}_{\vec{g}}$, dynamic wrench $\vec{W}_{dyn}$).
- Expressing rotational motion by quaternions avoids issues of other methods; favorable features are linear superposition and direct derivative calculation.

These favorable technical properties have not gone unnoticed (Dumas et al., 2004, 2007). Dumas et al. (2004) presented a numeric procedure to perform segment-wise balance of wrenches (equation 15 in Dumas et al., 2004). This enables the complete solution of inverse dynamic problems with a single, easy to apply numeric formula. Dumas *et al.* have provided and since then improved an implementation of their procedure in Matlab (Dumas, 1999). The sceptical author of this thesis has reproduced that procedure in Python, both numerically and analytically (Mielke, 2021).

6.4.2 Algorithm

With wrenches, quaternions, or the numerical formula at hand, one can compute the Balance of Wrenches for a single segment. Limb segments are usually arranged in a chain, e.g. a human leg consists of the thigh, a lower leg, a foot, and toes. However, there is an issue. The inertial wrenches (gravitational wrench, friction, ...) are known; the dynamic wrench is measured from kinematics and inertials; yet connecting wrenches are usually unknown: each segment (e.g. the thigh) receives an unknown wrench input from all adjacent segments (e.g. the pelvis and the lower leg).

There is one exception: the distal-most segment which has ground contact. Instrumented runways can contain force plates to measure contact forces and moments at the contact of the distal limb on the ground. Force plates usually generate electronic read-outs from directional, pressure sensitive piezoelectric crystals. By calibrating the plate, one can convert a voltage profile to forces, moments, and a contact point, all of which are changing over time.

Hence, the practical strategy (Dumas et al., 2004; Lynch and Park, 2017; Robertson et al., 2013) is to start at the most distal limb segment (e.g. the foot-toe or hoof complex), and measure its movement and ground contact wrench. In that situation, all wrenches in the balance are known, except the wrench at the proximal joint (human foot example: the ankle wrench). The external wrench at the ground contact point, in the balance calculation of the segment which is in ground contact, is also called the tip wrench; trivially, it is zero if the distal segment is lifted off the ground. Note that the read-out of a force plate gives a contact point, which is not the $C\vec{O}M$, so the wrench can be taken at contact point and must then be shifted. One can calculate the balance of wrenches by subtracting all external wrenches (ground contact wrench, gravitational wrench, proximal joint wrench, ...; all shifted to the $C\vec{O}M$ as the common reference point) from the dynamic wrench of the segment (measured linear and angular movement, also with respect to the $C\vec{O}M$). Exactly one of the external wrenches is unknown, the one on the proximal joint; we can solve for it (careful with reference points: the calculation comes out at the $C\vec{O}M$). The proximal joint wrench is shifted to the proximal joint position, its natural “point of attack”⁴. Inverted in sign (Newton’s n^{th} law), the joint wrench becomes the tip wrench of the proximally adjacent segment and adds to the balance of wrenches of that segment (Fig. 6.8). There, the procedure is repeated.

6.4.3 Inertial Properties

The procedure is quite trivial to implement, however the measurement uncertainty for all the involved quantities is not homogeneous. Whereas force plates and spatial coordinates can be accurately measured in a calibrated lab reference frame, the **inertial properties** of each segment require rough measurements and approximation. The relevant inertial properties are mass m , center of mass $C\vec{O}M$, and mass moment of inertia I . This justifies a closer look at them.

There are several ways to retrieve inertial properties. The classical method requires a dissection of the segment; this can be a precise anatomical disarticulation or straight-cut slicing of a frozen specimen. In the case of our slamming door, one would take the door off the hinge. The segment mass is then measured on a simple scale. Center of mass could be derived from a classical pendulum experiment (Fig. 6.9): hang an object from a thread, measure oscillation period, apply $T = 2\pi\sqrt{\frac{l}{g}}$ solved for effective length l , repeat the procedure with the object hung along a different axis. A common simplification is to express the position of the $C\vec{O}M$ as a fraction of the segment length, inaccurately assuming that it lies on the segment long axis. Anatomical landmark references are required to transform the relative $C\vec{O}M$ position into the lab reference frame. Finally, the classical method to determine mass moment of inertia is by using a trifilar pendulum (Korr and Hyer, 1962; Schedlinski and Link, 2001; Wells and DeMenthon, 1987).

⁴This series of shifts should have illustrated why the wrench is a useful construct: computational rigor can make sure that the reference points are respected.

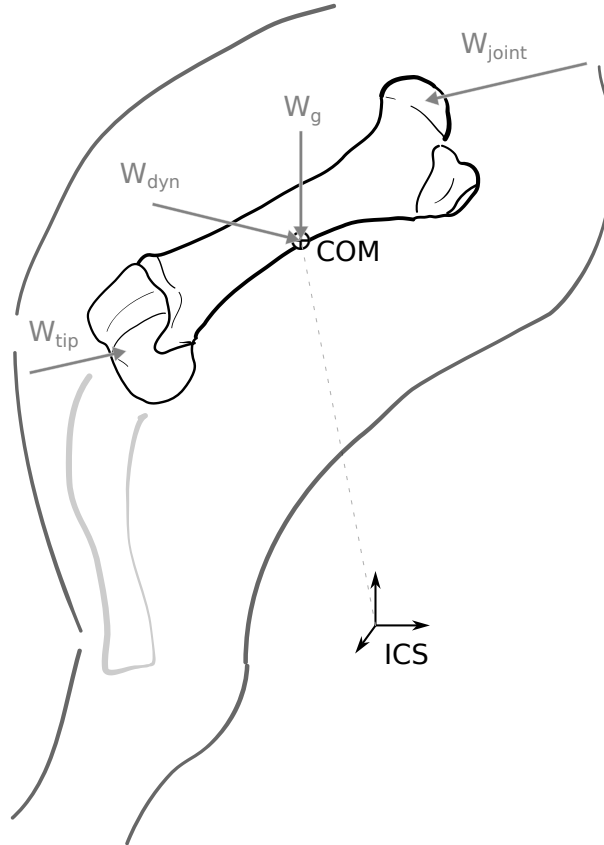


Figure 6.8: Wrench balance of a piglet femur (sketch), as it can be computed for each segment from distal to proximal. The tip wrench ($\vec{W}_{tip}$, at the distal joint) is the input from the distal adjacent segment or from ground contact. Dynamic wrench ($\vec{W}_{dyn}$, at the center of mass/ $\vec{COM}$) is calculated from kinematics. Gravitational wrench ($\vec{W}_g$, at $\vec{COM}$) is known by the mass of the segment. These wrenches allow to calculate the joint wrench ($\vec{W}_{joint}$, at the proximal joint), which is the tip wrench input for the proximal adjacent segment. All calculations can be performed in the inertial coordinate system (ICS , i.e. the inertial reference frame).

The so-retrieved measurements are filled in to the algorithm outlined above.

Measurement uncertainty of the inertial properties is not determined by the accuracy of the measurement instruments. This is where the “slamming door” analogy fails, because a door is actually rigid. The pendulum methods above yield numerically accurate results. However, the dissected limb is only a limited representation of the segment *in vivo*, for several reasons. A first issue is the choice of a disarticulation cutting surface: which mass particles are assigned to which segment? This choice affects mass, but also $C\vec{O}M$ and mass moment of inertia, and because cuts are generally further away from the $C\vec{O}M$, this effect matters.

Second, muscles (as a complex composition of fiber bundles and fibers) deform, shift and slide past each other during contraction and movement (Böl et al., 2013). Their cross section and fibre distribution can change on non-isometric movements. Gravity pulls soft tissue into a direction (Hansraj et al., 2022) which might be different *in vivo* and on the preserved configuration of the limb. Zeugopodial and autopodial segments often include multiple bones which might experience a shift in their relative position during a stride cycle. All this reduces congruence of the preserved, measured body part and the *in vivo* situation: mass is constant, yet the $C\vec{O}M$ of a segment is not stable during locomotion, and the mass moment of inertia tensor will most certainly vary during the process.

Third, there might be shrinkage and tissue loss during the preservation of the specimen. Formol, ethanol, or temperature are typical agents used as a fixative; any fixative has its effect on soft tissue (Buytaert et al., 2014; Pech et al., 1987). This issue can putatively change all inertial properties of the sample.

Fourth, 3D volumetric imaging (e.g. CT scans) could be an option to replace the measurements above, yet such scans do not solve all of the other issues, and instead come with their own set of problems and limitations (topic of the next chapter).

The implications of these conceptual issues will be discussed in detail in the next chapter. Yet note already that these four effects violate an assumption above: inertial properties are hard to determine, and the velocity product terms in the balance of forces are non-negligible because the mass distribution is not constant.

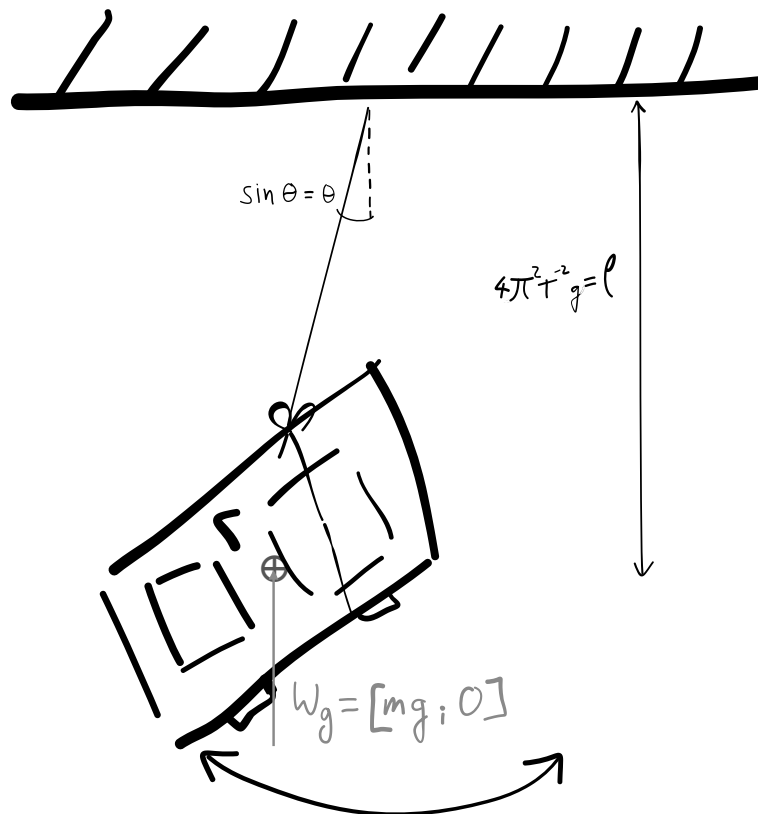


Figure 6.9: *Measuring inertial properties of the slamming door: the center of mass.*

6.5 Summary

This chapter serves as a brief overview of the basic methodology of inverse dynamics. This method exploits the fundamental physical phenomenon of conservation of momentum to model and infer the reasons for movement of systems of more or less rigid bodies. I outlined the general algorithm, and documented one specific procedure applicable to inverse dynamic measurements of quadruped, terrestrial locomotion.

In particular, the Wrench-Quaternion method (Dumas et al., 2004) has proven useful for joining kinematics and force measurements for dynamic calculation. The reason is computational convenience. Wrenches are the conceptual link of forces and moments, and they allow convenient computer implementations. Quaternions have favorable properties to calculate spatial rotation in the context of rigid body motion (superposition, singularity avoidance, transformability). This establishes the conceptual framework to perform the relevant calculations of the conventional algorithm, which involves starting at the distal segments where external forces are known (e.g. ground contact) and processing the calculation towards the torso of the animal.

A crucial simplification of that procedure is that we treat body segments of animals as rigid bodies, which they are not. By *post mortem* fixation and dissection, or by digital volumetric segmentation, segments can be separated and their inertial properties measured. The accuracy of these methods will be the topic of the next chapter.

Chapter 7

Approximating Inertial Properties of Biological Specimens using CT Data, or Not

7.1 Abstract

An essential part of the inverse dynamic workflow is the measurement of inertial properties, in particular mass, center of mass, and mass moment of inertia. In the past, these have often been approximated by assigning average density values to limb segments. Recent studies have explored the potentially more accurate retrieval of inertials from computed tomography (CT) data, which would work by associating voxel gray values of the CT image stack with densities. However, CT images actually depict attenuation, which is a different attribute of matter and only roughly correlates with density. Though this was known since the early days of CT, the temptation of easily retrieving inertials has recently led researchers to re-attempt linear approximation methods.

In this study, I document my own attempts in retrieving density values from CT images. Though a conversion is possible by using a regression, I demonstrate that the resulting mass and inertia are highly erratic. By performing a thorough sensitivity analysis, and by simulating common CT artifact, I can quantitatively evaluate what effect uncertainties in the regression have for the outcome measures.

The results are alarming: in the test case, mass is overestimated by more than a fifth, and error propagation indicates that this has an even worse effect on moment of inertia, which directly translates to errors in inertial dynamics. However, there is no reason to expect alternative, conventional methods to be more accurate. Thus, though CT images are not ideal for the purpose, they might be the best option available in most cases. In conclusion, I emphasize the value of a sensitivity analysis, and point at promising research progress of others which might lead to improving CT reconstructions for the purpose of estimating material characteristics.

7.2 Introduction

7.2.1 Computed Tomography and Density

Since the advent of x-ray imaging, people are intrigued by the ability to see the inner structure of objects and living creatures (such as the famous hand of Röntgen’s wife, “Über eine neue Art von Strahlen”, 1895). This desire even increased by the development of computed tomography (CT; Beckmann, 2006; Hounsfield, 1973), a set of techniques which enable the reconstruction and visualization of three-dimensional structural images. The transmission images obtained via high energy electromagnetic radiation often serve to answer qualitative questions (e.g. whether a bone is fractured). Quantitative questions are obvious with regard to the shape of a scanned structure (e.g. the length of a fracture, the shape of a bone, trabecular and cortical micro-/architecture). However, researchers have been struggling with the quantitative extraction of material properties from CT data. The material property of primary interest is physical density; from a given density distribution, other relevant inertial properties can be calculated.

Researchers suggested early on to relate the gray value of CT images to density (Mull, 1984; Phelps et al., 1975), yet it was immediately noted that the relation is nothing more than a correlation which only holds under specific circumstances. Even these pioneer works acknowledge that there must be discrepancies between real and x-ray-derived densities, which are often associated with (i) the polychromatic character (source spectrum) of the used radiation, (ii) chemical composition (absorption spectrum), and (iii) scan artifacts.

These issues demand a detailed explanation, below. Notwithstanding this list of known problems, people have repeatedly attempted to extract density and inertial properties from CT scans (Du Plessis et al., 2013; Durston et al., 2022; Phillips and Lannutti, 1997, as well as the present study). The purpose of this chapter is to explore whether or not (or under which circumstances) CT gray values can be used to estimate density distributions, and thereby mass and other derived inertial properties. I will start by giving a brief intro to crucial aspects of CT scanning technology, before reviewing similar attempts by others, then introducing a simple experimental setup to measure dynamics of an excised piglet femur, and with that re-attempting the extraction of inertial properties.

7.2.2 Emission and Absorption

Visible light and x-ray radiation are, to a large degree, analogous. The reason for this is that both are electromagnetic radiation; they differ in wavelength, thereby energy, and thereby in the way they interact with matter. To illustrate what is happening with x-ray light during a CT measurement, I will use the analogy of visible light.

Most plants have evolved to be green; their cells hold organelles with the green pigment chlorophyll, which absorbs some non-green (e.g. red) wavelengths of light very well (Fig.

7.1, green curve). Light are photons, “elementary particles which are the quantum of the electromagnetic field” (Wikipedia: photon). Filters let through some of them, while prohibiting others from passing. Think of party lights or colored window sheets: a light source is placed behind one or more filter sheets which block most colors and let through certain others. In case of a green filter sheet, relatively more green photons pass the filter, the filtered spectrum will appear green to our visual perceptive system. In plants, the ensemble of non-absorbed photons also appears green to us, which is why leaves generally appear green to our view. But leaves only appear green if there is a green content in the incident light! Green leaves will look red under red light (and “darker”, i.e. lower number of reflected photons than for an equal input intensity of green light). If you want to reduce the chance of being seen while looking for something in a dark forest, better use a red flashlight.

To characterize light which is emitted, reflected, or absorbed, one can plot a spectrum (Fig. 7.1) ¹. Red LEDs are an alternative to red filters: they immediately produce a spectrum which is biased towards low frequencies. One can assess that red LEDs have a different **source spectrum or emission spectrum** than white LEDs or “vintage” (tungsten) light bulbs. If the light from a white source is filtered (e.g. by a red sheet of thin, translucent plastic), the spectrum is altered. A red LED viewed through a green-pass filter (e.g. a leaf) might appear to be “off”, when really power is “on”, because the majority of photons are absorbed by the filter.

In fact, light hitting *any* object will change its spectral composition, depending on the material’s **absorption, transmission, and reflection spectrum**. This is also a filtering process: the outgoing light will depend on the incoming spectrum and the material properties of the object. Plant leaves will almost entirely absorb a narrow band of incoming red light. Interactions of photons and matter crucially depend on the energy/frequency/wavelength of the photons, and the energetic/vibrational/resonant molecular properties of the matter. A (differential) spectrum is a way to depict a wavelength dependence.

Crucially, the interaction of x-rays (i.e. high-energy, low wavelength photons) with matter differs from that of visible light and matter. Hence, the analogy between visible light and x-rays only partly holds. X-rays (Berger et al., 2018a; Buzug, 2008) are also photons, just at a different wavelength (range 10 *pm* to 10 *nm*). An x-ray detector is a complex sensor device, similar in function to digital camera sensors, which will detect any light which is not absorbed or scattered by the sample or air on its way through the scanner chamber. X-ray sources are “targets”, i.e. anodes made of certain metals (Tungsten, Molybdenum, ...), shot at with an electron beam which is evoked by applying voltage (e.g. 60 *kV*) to the warmed-up anode and cathode. When hit by electrons, the target will emit x-ray photons of a specific composition of energies (spectrum, Fig. 7.2). The source emission is generally **polychromatic**, i.e. consisting of multiple colors/energies (just as the spectra in Figs. 7.1 and 7.2). Most household CT scanners (in contrast to synchrotrons) have a polychromatic

¹Patient readers of this thesis will see a connection to Fourier methods (Ch. 2).

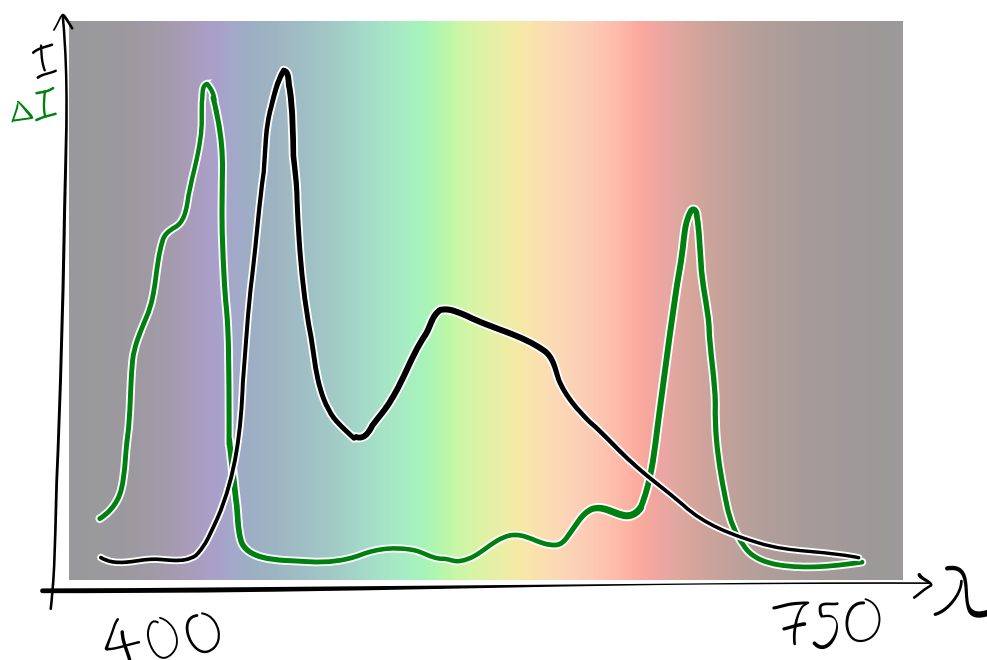


Figure 7.1: Spectra. The horizontal axis shows all wavelengths λ (nm) of relevance (in this case the range of visible light, for illustration). Related to wavelength and therefore viable axis label alternatives are frequency (units: Hz) and photon energy (units: eV). The vertical axis shows photon occurrence probability, or normalized intensity I , or the intensity difference ΔI for the absorption spectrum. Black curve: emission spectrum of a white LED (Tanabe et al., 2005). Green curve: the approximate absorption spectrum of chlorophyll A (Zscheile, 1934). The chlorophyll will hardly get excited by the LED, yet the plant will look bright and green.

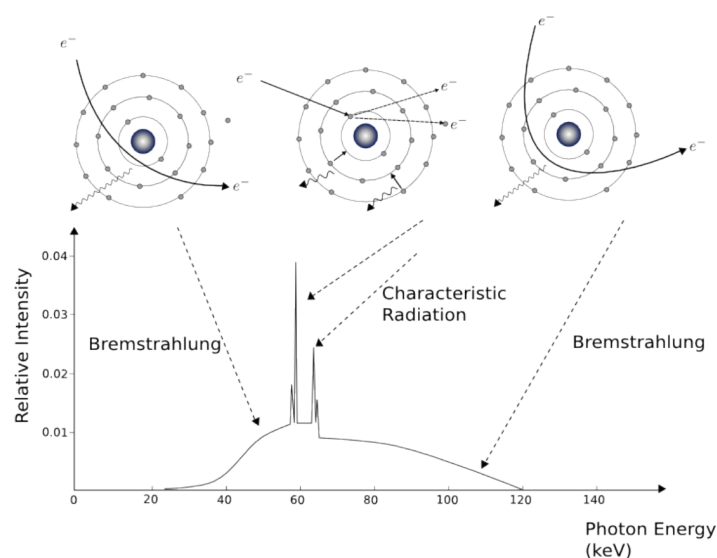


Figure 7.2: X-ray emission spectrum. An x-ray source will emit photons with a variety of energies (“colors”). Bremsstrahlung will cover a wide range of photon energies, whereas discrete peaks are caused by specific emission processes in the target material. (taken from Berger et al., 2018b, creative commons license).

source. On the other hand, x-ray detectors usually produce monochrome images, but are not monochromatic! Much like a monochrome digital camera photo cell measures light intensity of any wavelength within the visual range (restricted to *visual* by another filter); in contrast, a truly monochromatic detector would detect only a specific wavelength. They integrate intensities over a wide range of wavelengths (in a specific way that could be measured as the sensitivity spectrum of the scintillator²). There is ongoing development on the frontier of “spectral CT” (Liu et al., 2023), yet spectral and spatial resolution are currently still below par (whereas price is not).

Once set on their path from within the x-ray source, x-ray photons interact stochastically with matter they encounter on their trajectory (photoelectric effect, Compton scattering). “**Attenuation**” is the term to describe that not all photons reach the detector, i.e. the detected x-ray intensity is lower than the incident intensity. The probability of either of the possible interactions depends on (i) the wavelength (photon energy), (ii) the elementary composition of the material (absorption spectrum, K-edges), and (iii) the trajectory of the photon (thickness of the material, angle of incidence).

The varying degrees of attenuation, when measured from multiple incident angles for 3D view, is what enables the extraction of structural information (Lambert-Beer’s Law). If children place their hands inside a conical light beam, an animal-shaped shadow will be projected onto the wall. Their hands attenuate the light. Photos of the hand’s attenuation pattern from all possible directions (i.e. rotating the hand) can in fact be used to reconstruct the 3D shape of the hand to falsify the hypothesis that an actual animal was causing the projection.

And attenuation is precisely the property which is thought to correlate with physical density: the higher the density, the higher the attenuation. Or so it seems. Yet think of a case in which the specific emission of the source does not match the absorption peaks of the material - as in the example of a white LED not exciting chlorophyll of a green leaf (Fig. 7.1). Or the opposite case, a substance with an absorption spectrum which is mostly congruent to the emission spectrum of the source. Examples of substances problematic for x-ray are water and formol, because they absorb a broad range of photon energies within the x-ray range.

To summarize: both in the visual and x-ray range of electromagnetic radiation, emission and absorption are determined by stochastic interactions of photons and matter. Spectra summarize ensemble properties of a given light source or material; differential spectra measure relative absorption. The filtering properties of matter can be used to acquire images and reconstruct 3D structure, even in the absence of precise spectral information.

²The scintillator is the device which enables detection of x-ray, by converting x-ray light into visible photons which are then detected by a sensor array.

7.2.3 Scan Artifacts

X-ray images do not always look as one would want them to look. The unfavorable image features are commonly called “artifacts” (Triche et al., 2019). In the opinion of the provocative author, there is actually no such thing as CT scanning artifacts. The term “artifact” linguistically implies³ that there is an “error” or flaw in the measurement, yet instead it can be ascertained that correctly obtained x-ray images are highly accurate. Any tomographic reconstruction just shows exactly what is measured, convoluted with ideally negligible reconstruction algorithm characteristics.

Acknowledged, some aspects of the measurement might be unfavorable to the observer, because they deviate from the image which that observer expects, based on their personal experience of the real world. For example, a “beam hardening artifact” occurs if absorption in the superficial layers of an object alter the spectrum of the beam on its trajectory, which will affect the virtual representation of the deeper regions (Van Gompel et al., 2011). Low energy photons have lower penetration depth than high energy photons, because they are more likely to interact with matter. This causes a gradual change of the spectrum, which will shift towards higher average energy along the ray’s trajectory through thick material. An observer will know that a cylinder is homogeneous, and reject the image which shows a radial density gradient. However, that gradient is in fact a normal manifestation of the actual physical process (the stochastic interaction of electromagnetic radiation and matter). Curiously, beam hardening can be minimized by pre-hardening the beam with the use of metal filter plates (Triche et al., 2019).

Similarly, ring artifacts stem from sensitivity variations on the image detector, which are technically inevitable (due to constraints of the physical detection process), but can be rectified reliably (Sijbers and Postnov, 2004). Partial volume effects are caused by finite scan resolution and voxel volume. Streak artifacts are caused by limited dynamic range and “photon starvation” (i.e. beam hardening, again). All these could be considered properties of the scan, rather than interpreting them as an annoyance.

Another group of so-called artifacts might stem from the choice and limitations of the reconstruction algorithm. There are iterative/algebraic and analytical reconstruction methods (Andersen and Kak, 1984; Feldkamp et al., 1984; Geyer et al., 2015; Gilbert, 1972; Hansen et al., 2021), all of which have their specific limitations. These algorithms are constantly refined and improved, and specific algorithm variants can already overcome specific scan limitations (e.g. Frenkel et al., 2022; Six et al., 2019). In the future, the question of reconstruction artifacts will rather be one of algorithm choice. Today, the most common reconstruction algorithm, used in a vast majority of the CT service facilities with

³Artifact: “A product of artificial character (as in a scientific test) due usually to extraneous (such as human) agency”; Merriam-Webster online dictionary, ; <https://www.merriam-webster.com/dictionary/artifact>. “A false or inaccurate result or effect that is produced accidentally”; The Britannica Dictionary; <https://www.britannica.com/dictionary/artifact>. Accessed February 6th, 2024.

cone-beam setups, is “filtered backprojection” (FBP, or Feldkamp/Davis/Kress = FDK algorithm; Feldkamp et al., 1984).

To summarize: scan artifacts, if one wants to use that label, are perfectly normal. Some might be partially circumvented, others show intrinsic features of the technical and computational tomography procedure. Artifacts are interpreted as “something is not as it is supposed to be”, despite attenuation-based images being close to technical perfection. Artifacts are in conflict with our psychological image of the real world: metal beads do not look like “stars” (radial light rays) to us in the real world, so the image is rejected. I suspect that the major reason people still perceive artifacts as problematic is that we actually think of matter and physical objects as a distribution of density, i.e. a mass distribution, whereas x-ray scanning really yields a distribution of x-ray attenuation.

On the other hand, if assuming that density cannot be recovered from attenuation images, the only appropriate way to measure an exact mass distribution of an object would be to physically slice it into millions of little voxel cubes and weigh each of them. Compared to that option, computational tomography is certainly a time-saving alternative.

Yet it must be kept in mind that x-ray images do not quantify density, they quantify attenuation, most often lumped over a spectrum.

7.2.4 Density Approximations: Two Case Studies

The fact that two physical properties (attenuation and density) are fundamentally different things does not imply that researchers cannot use one to measure the other. Scientists have repeatedly suggested and attempted to convert CT scan gray values to approximate density, which I will discuss on two examples (Du Plessis et al., 2013; Durston et al., 2022).

Du Plessis et al. (2013) acknowledge general difficulties of accuracy and repeatability in extracting density from CT data.

They then average gray values of putatively homogeneous blocks of polymer plastic material, apply linear regression which, as they point out themselves, does not appropriately cover two of the measured data points. The latter problem is attributed to differences in chemical composition (without discussing the known composition of the objects). The authors then argue that chemical composition can be assessed by performing a Dual-Energy CT scan (DECT), i.e. scanning at two different tube voltages and taking the ratio of gray values. Note that this mode of execution of DECT is not really “dual” energy: the spectrum of energies elicited in a 60 *kV* scan is included in the 230 *kV* scan; the photons below 60 *kV* might even contribute the majority of light in the high energy scan. A better differential would have been achieved by incorporating metal filters in the 230 *kV* scan (just as in beam hardening filter, see above). Nevertheless, comparing the calibration line and the DECT ratio results, the putative outliers do not stand out more than other regression elements (especially the PTFE sample puts the original regression in doubt: it is perfectly fit by the

calibration regression, is a modest outlier on the DECT ratio, despite a special chemical composition containing fluorine).

For a toy pig of unknown plastic material, Du Plessis et al. (2013) retrieve higher than actual mass estimates; they identify chlorine and calcium content as the cause. On another unknown sample which is assessed to be “similar enough” in chemical composition, they retrieve an accurate prediction. The authors do not quantify or report measurement uncertainty which manifests in gray value distributions, sub-optimal regression goodness of fit and potentially high regression sensitivity (PTFE), or other sources of variability (Macaulay et al., 2017). Finally, they destroy a toy pig for the study, which cannot be excused.

In a nutshell, Du Plessis et al. (2013) attempt density prediction in a particular use case (homogeneous, chemically identical objects) and suggest a DECT ratio to assess chemical composition.

Durston et al. (2022) attempt a huge leap from there and measure inertial properties of frozen cadaver parts, both conventionally and computationally. Emphasis: they use whole birds, and their considerable amount of work must be appreciated!

As the authors above, Durston et al. acknowledge the critical assumption of a linear relation between attenuation and density, and consequentially use a simple linear relation as a conversion from CT scan in Hounsfield Units to actual density. They supplement the calibration regressions, which show systematic errors at close look (the regression line lies tilted compared to the majority of relevant calibration objects; it looks biased by the “air” sample point yet that is a valid and important reference point; the authors do not discuss this). Still, that the linear fit works at all is surprising, given that this study uses tissue phantoms provided for medical CT, all of which will putatively have a slightly different chemical composition (linearity assumption violated).

To validate their results from CT density estimates, the authors apply two approaches. The first is a comparison of virtual and physical dissection, with regard to segment mass measurements. The second is a trifilar pendulum “ground truth” for one axis of the mass moment of inertia. There is also a validation of the pendulum method, by applying it to manufactured nylon blocks, yet quantitative error estimates remain vague.

Overall, results of the Durston et al. (2022) study remain superficial considering the author’s valuable efforts on this project. They juxtapose pie charts of segment masses to verify mass distribution, which is far too inaccurate. They compare the virtually and physically derived dorso-ventral axis of the moment of inertia, and present what must be a clear mismatch, considering the lack of meaningful error margins. They discuss the influence of partial volume effects on density (which, agreeably, could be a problem with feathers), however that is irrelevant for mass and moment of inertia because volume and density both change if a voxel contains air and tissue. They present “moment of inertia distributions” in various ways, i.e. the contribution of each voxel to a $C\vec{O}M$ -centered I for an entire articulated skeletal system, which is of little practical relevance (which they confirm themselves, by

comparing extended and retracted wing configuration). It is a good reproducibility control that their data confirms previous findings of segment masses of some bird species. The article is published in the “methods and techniques” section of the Journal of Experimental Biology, therefore fully focuses on the CT method, and does not discuss any further application of the generated data.

This critique of these studies seems harsh, and I apologize for pointing a finger. Many other studies have lumped the whole scan volume into a single density material, whereas the presented examples explore potential improvements. I selected them because they are related to my own work, and to point out a unifying feature and a reason for the highlighted flaws in theses and other studies: they are *output driven*, and fall short of discussing the mechanistics of CT imaging. As shown in these examples, practitioners often simply assume that structural CT data represents physical density, instead of failing to falsify this claim. The results are studies which yield approximate density distributions, yet fail to quantify the inaccuracy and uncertainty of their quantitative data.

The author of this thesis is no exception. As the authors of the studies reviewed here, I lacked insight and was driven by good hope when starting the work on the present project. In consequence, this study is centered around the hypothesis that CT gray values can not be converted to physical density, hopefully highlighting pitfalls to avoid and limitations to be aware of for future researchers.

7.2.5 Reductionist Approach

This study follows up on the idea of using a calibrated CT scan (Du Plessis et al., 2013), and adds a biologically relevant test object as well as an evaluation of the error margins (Arroyave-Tobon et al., 2022; Hughes and Hase, 2010; Myers et al., 2015), which is crucial for practical use. Calibration is attempted with cheap, leftover plastic pieces of known material and density, as well as commercial bone mineral density calibration phantoms. I chose a dissected porcine femur as the object of interest of which I seek to estimate the density distribution, and thereby total mass, center of mass, and mass moment of inertia (Fig. 7.3). I also attempted to validate the method on dissected animal specimens (as Durston et al., 2022, but less sophisticated).

The hypothesis that CT images are a valid approximation of physical density is already falsified by the theoretical considerations above. The following relevant questions remain:

- How far off the true value is the calculated mass moment of inertia, i.e. what is the measurement error?
- How (much) exactly do known artifacts contribute to the measurement error?
- Are the cheap plastic parts appropriate calibration options for organic tissue?

It should be pointed out that preliminary (naive) results of this study were first presented at the conference of the Society of Experimental Biology (SEB) in 2021, one year prior to the Durston et al. (2022) study.



Figure 7.3: *Experimental Setup.* A piglet femur was excised and marked with metal bead markers. Together with various calibration objects, the sample was fixed (“mounted”) in a PET bottle to fit into the cylindrical scan volume of the FleXCT scanner of the University of Antwerp. The image on the left shows the sample mount, magnified on the inset. On the right is one CT projection of the sample, with the femur with markers clearly visible on top.

7.3 Materials and Methods

7.3.1 The Flying Femur

As discussed in the previous chapter, the goal of inverse dynamics is to calculate the segment-wise balance of wrenches, thereby elucidating which forces and moments each joint has to handle in a motor task. The unit of calculation is therefore a segment. In a reductionist approach, the purpose of this experiment is to perform all procedures and calculations on one extracted rigid part of a segment.

For this purpose, a femur was excised from a piglet specimen which had been used previously in XROMM⁴ experiments. The animal was euthanized after successful completion of the experiment. Procedures have been approved *a priori* by the Ethical Committee of Animal Experimentation, University of Antwerp, Belgium (approval number 2017–25). The femur was extracted by carefully dis-articulating the parts of the right hindlimb of the piglet, isolating the femur, and removing surrounding soft tissue.

The extracted bone underwent the full XROMM and inverse dynamic modelling procedure (Fig. 7.4). Metal bead markers, 0.5 mm standard soldering balls made out of a lead-tin alloy ($Sn_{63}Pb_{37}$), were glued to the extracted bone to simulate the typical XROMM necessity of marked bones and facilitate landmark tracking on videos. Biplanar x-ray video recordings were performed on the University of Antwerp’s 3D²YMOX system (Nguyen et al., 2021; Sanctorum et al., 2020), with two experimental settings. In one setting, the femur hung on a long, thin nylon string pendulum (appearing on camera to be “flying”). In a second setting, the bone was rotated to mimic in vivo motion by a motor to which it was attached with a radio-translucent, custom-made plastic clamp.

After the experiment, the femur was subjected to a calibrated scan at a local micro-CT facility (FleXCT, imec-VisionLab, University of Antwerp, Belgium). Plastic parts for calibration were donated by the mechanical workshop of the university, covering a range of physical density from polypropylene (PP, $916 \frac{kg}{m^3}$) to Polytetrafluoroethylene (PTFE, $2210 \frac{kg}{m^3}$).

In addition to the plastic debris, two professional calibration phantoms (Bruker, USA) were included in the scan. Those are hydroxyapatite cylinders, immersed in water over night, with a diameter of about half a centimeter and a length of a centimeter. They calibrate bone *mineral* density⁵ of $0.25 \frac{kg}{m^3}$ and $0.75 \frac{kg}{m^3}$, and they have a *physical* density of $1254 \frac{kg}{m^3}$ and $1485 \frac{kg}{m^3}$. Their physical density was measured by dividing the weight (measured with a fine scale) and the volume (from CT scan), after the author had found out that those are different quantities. Air and tap water volumes in the scan have known densities and

⁴X-ray Reconstruction Of Moving Morphology, Brainerd et al., 2010

⁵Bone Mineral Density (BMD) quantifies the amount of inorganic content within a volume of bone tissue. Though sharing similar units with physical density, BMD measures just a fraction of the physical bone density.

complete the calibration series.

Expected outcomes of the “flying femur” experiment are twofold. Firstly, this provides a test case as a reference for actual XROMM and inverse dynamic calculations, just with simpler calculations, but including an “order of magnitude” estimate of the moments and forces required to move the bone (not reported herein). Secondly, the validity of the “calibrated CT” method for extracting inertial properties is to be evaluated.

7.3.2 Piglet Data

Originally, the femur was part of an animal, and thus also part of a bigger whole set of experiments. That project involved newborn piglets reared temporarily at the veterinary facilities of the University of Antwerp, and subjected to XROMM recording sessions. Experiments were approved by the Ethical Committee of Animal Experimentation, University of Antwerp, Belgium (approval number 2017–25) and performed in June and July 2019.

The part of these experiment relevant for this chapter are the CT scans. Those were not calibrated with the full array of plastic debris; they contained only air, water, and the bone mineral density phantoms. One of the piglet scans was chosen for the validation of the density acquisition method by comparing known segment- and total weight to the outcome of the CT method. For this purpose, the animal was dissected both physically and virtually (segmentation, Fig. 7.5) and the results compared. Segments were weighted individually on a precision scale. Virtual segments were processed with the density regression procedure to compute segment-wise mass estimates.

7.3.3 Scan Parameters

The femur scan was performed on a customized Tescan Unitom XL (FlexCT, University of Antwerp), with appropriate settings (Tab. 7.1). The 1.0 *mm* aluminum filter plate and the relatively high voltage were chosen to reduce beam hardening and metal bead artifacts. Three “averages” indicate that every projection is the average of three scan images from the same angle, which is a common trick to reduce pixel noise. Scan geometry was set to give sufficient resolution (small voxel size) at reasonable scan duration. The reconstruction value range was chosen to cover the entire histogram, excluding the metal beads, which gives best dynamic range on the plastic and organic tissue.

The piglet scans for the second experiment were performed at a different facility (Royal Belgian Institute of Natural Sciences, Brussels, Belgium) and with slightly different settings (110 *kV*, 500 *mA*, 62 μm resolution, 5 averages). Piglets were scanned in frozen state and in bags of two individuals at a time, to save time and cost. Air, water, and the bone mineral density phantoms were available for calibration in these scans. Resolution was lower, but exposure settings were comparable to the ones on the femur experiment.



Figure 7.4: *The "flying femur" experiment:* an excised piglet femur used in XROMM experiments as an allegory of reductionism. Left: the femur oscillating on a pendulum. Right: installation with a motor capable of moving the femur in a realistic way.

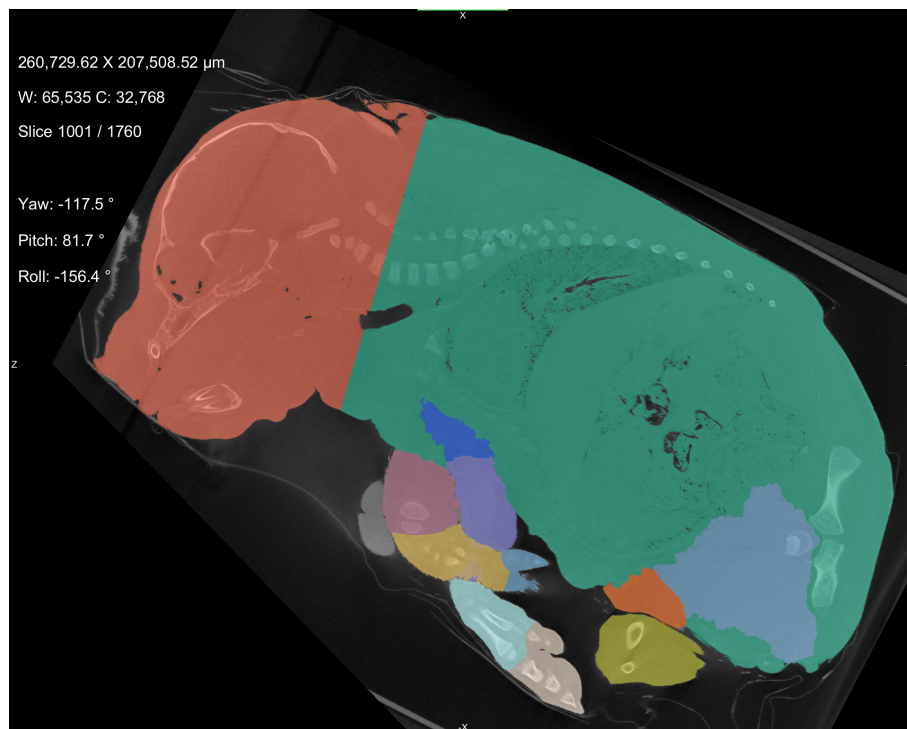


Figure 7.5: *Virtual dissection (segmentation) of a piglet CT scan. Example virtual slice.*

7.3.4 Inertial Properties

All code used on this project, including Python implementations of the mathematical formulas below, can be found online (https://git.sr.ht/~falk/flying_femur and http://mielke-bio.info/falk/posts/23.ct_density).

CT Segmentation

Of course, if e.g. the mass of a bone is to be calculated, one intends to sum up the mass of only the bone, excluding the mass of the surrounding air or support material. A preliminary step to calculate inertial properties from CT scans is the segmentation of the scan. Segmentation is the separation of the “relevant” and “irrelevant” sub-volumes within the image stack, in this case the bone and the surrounding air or background. More generally, limb segments which are treated as a unit have to be labeled in dedicated software (e.g. 3D Slicer), typically with a kind of “color brush” or “magic wand” tool; thresholding of gray values, boolean operations, algebraic operations (“filling” etc.), and other tricks may simplify the segmentation procedure in certain situations. The outcome is a 3D boolean mask which can be used to distinguish the voxels in the scan which are associated with the bone of interest.

This step might seem trivial, time consuming, yet necessary, and indeed it is all of that.

Mass

The mass of a volume element m_i of a rigid body is calculated as the density ρ_i of that element, multiplied with its volume V_i (Fig. 7.6). Summing up all volume elements will give the total mass of the object:

$$\sum_i m_i = \sum_{i \in V} \rho_i V_i \quad (7.1)$$

The crucial part here is to get ρ_i , the density per voxel, which is attempted via a regression.

Center of Mass

The center of mass is the mass-weighted average position vector of all volume elements of a rigid body (Fig. 7.6). For each volume element i , mass m_i is a scalar, the position vector r_i is a three-by-one vector, and their product is summed up for all voxels of the segmented bone. The result is normalized by dividing the total mass. The outcome is the three-by-one position vector of the center of mass ($C\vec{O}M$, shorthand for $\vec{r}_{C\vec{O}M}$).

$$C\vec{O}M = \frac{1}{\sum_i m_i} \sum_{i \in V} m_i \vec{r}_i \quad (7.2)$$

Table 7.1: *CT scan settings* for the calibrated femur scan.

voltage	150 <i>kV</i>
power	55 <i>W</i>
current	≈ 360 <i>mA</i>
filter	1.0 <i>mm Al</i>
detector field	1920×1896 <i>px</i>
pixel size	$(0.15 \text{ mm})^2$
exposure time	50 <i>ms</i>
averages	3
projections	2879
source-detector	800 <i>mm</i>
source-object	267 <i>mm</i>
binning	<i>none</i>
voxel size	50 μm
scan duration	10 <i>min</i>
reco value range	$[-0.2, 1.0]$

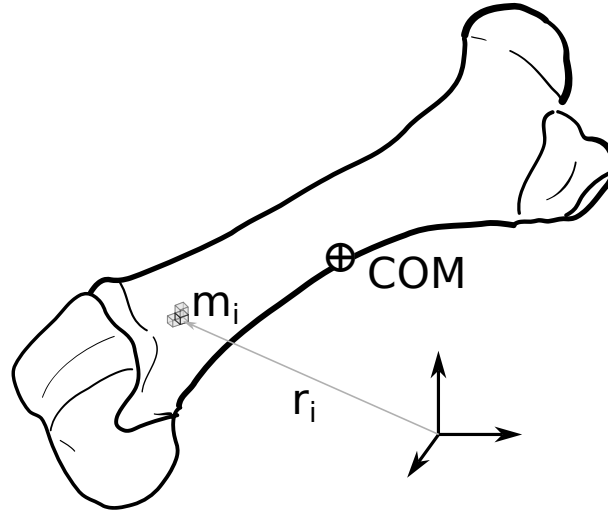


Figure 7.6: *Calculation of inertial properties of a limb segment (femur) from CT scans.* *CT scans are 3D images, consisting of voxels (cubes) each of which is associated with a gray value. If these could be associated with physical density, it would be possible to calculate the mass, center of mass (COM) and mass moment of inertia. In the calculation procedure, voxels are treated as little mass elements m_i , which are positioned at a given vector position r_i from an arbitrary CT origin.*

Mass Moment of Inertia

As stated before, mass moment of inertia is a tensor (speak: 3×3 matrix) which measures the resistance of an object to angular acceleration. It can be calculated for any rigid body (see Fig. 7.6) via an integral formula over all volume elements (Wikipedia/Moment of Inertia, 2023):

$$I = \sum_{i \in V} m_i \left(\|\vec{r}_i\|^2 E_3 - \vec{r}_i \otimes \vec{r}_i \right) \quad (7.3)$$

... with the rigid body's volume V split up into voxels i which have mass m_i and position vector $\vec{r}_i$; E_3 is the 3×3 identity matrix and $\otimes$ the outer product.

Initially, this is always calculated with respect to the $C\vec{O}M$ as reference point, the algorithm usually includes subtracting the $C\vec{O}M$ from the position vectors. The inertia tensor is a symmetric matrix, and can be brought to a principal form by an Eigenvalue procedure, in which the off-diagonal ("products of inertia") become zero. The reference point of I can be changed by the generalized parallel axis / Steiner's theorem (Lynch and Park, 2017, p. 245).

$$I_p = I_0 + m \cdot \left(\vec{s}^T \vec{s} E_3 - \vec{s} \vec{s}^T \right) \quad (7.4)$$

... with $\vec{s}$ being the shift vector connecting the original and target points.

It is difficult to get an intuition about Mass Moment of Inertia, but classroom demos can illustrate how this measure depends on the geometric distribution of mass of an object (Dalvi, 2017; Lewin, 1999a,c).

7.3.5 Density Regression

The critical step in the procedures above is the relation of gray values γ and physical densities ρ . This is the search for an (idealized) function $\gamma = f(\rho)$, which can predict the CT gray value for any given density. The inverse, $\rho = f^{-1}(\gamma)$, can then be used to assign densities to gray values from the scan. Several regression functions were attempted, based on a guessed relation of the observed gray values of the calibration objects.

The regression was performed in Python, namely using the `scipy.optimize.minimize` function to minimize the Euclidean difference between observed and fit values with the Nelder-Mead algorithm, tolerance set to 10^{-16} (Gao and Han, 2012). Convergence was supported by setting meaningful parameter start values close to the expected outcome (Tab. 7.2).

Table 7.2: Density Regression Functions and regression start values.

	function	start values
linear	$a + b \cdot x$	$a = 0.175, b = 10^{-4}$
exponential	$a + b \cdot e^{c \cdot x}$	$a = 0.12, b = 0.05, c = 0.0015$

Ideally, all measured points would lie on either of these functions. Yet that turns out not to be the case in the actual data. Subsets of the calibration objects were selected for the regression to be plausible: one group were air and the plastic parts, which approximate a linear relation; the other group were air, water, and the two bone mineral density phantoms, which seem to follow an exponential relation; the sample labeled “PVC” was plausible for neither group and excluded.

7.3.6 Scan Artifacts

To estimate the effect of beam hardening artifacts, one of the scanned calibration objects was selected by segmentation and virtually modified. I selected PTFE, because it has the highest attenuation in the data set and is therefore most prone to suffering from beam hardening. The PTFE cylinder has almost ideal cylindrical shape, and was well aligned with the scan rotation axis. The object was segmented and cropped out of the whole scan so that the long axis of the cylinder aligns with the center of the cropped volume.

Beam hardening manifests in a radial attenuation gradient in objects: superficial layers which are reached first by x-ray beams seem to attenuate more (in absolute terms) than internal volume elements, because they receive a higher input radiation; the inner elements are partially shielded by the outer elements. The attenuation gradient produces a cup-like profile in the gray value images, which is why beam hardening of cylinders is called “cupping”. To quantify the amount of cupping, the cylinder was transformed to cylinder coordinates (`scikit image/transform/warp polar`, https://scikit-image.org/docs/stable/api/skimage.transform.html#skimage.transform.warp_polar), and then flattened by averaging along its vertical (long) axis.

$$\mathcal{I} = a + b \cdot e^{k \cdot r} \quad (7.5)$$

By using an exponential regression (7.5) to the exponential part of that cylinder’s intensity profile $\mathcal{I}$, one can extract a parameter k which quantifies the amount of cupping (k as in the Dutch word “kopje”). With that k known, one can rectify the gray values of the scan by applying a correction factor which would push up the exponential line to a flat constant (obviously not applicable for irregularly shaped objects).

Similarly, one can multiply the intensity values of each voxel in the original volume with an exponential of its distance to the center point. Depending on the chosen value for k in that exponential, one can virtually manipulate the amount of beam hardening (Fig. 7.7).

The focus of this virtual experiment is Mass Moment of Inertia. One might argue that, though the value of the tensor I depends on the mass, its principal axis orientation only depends on the mass distribution. Given an expected mismatch of the calculated total mass and the actual measured total mass of the test object, an appealing strategy might be to scale the voxel-wise masses so that total masses match. I call this strategy “scaling”, and it

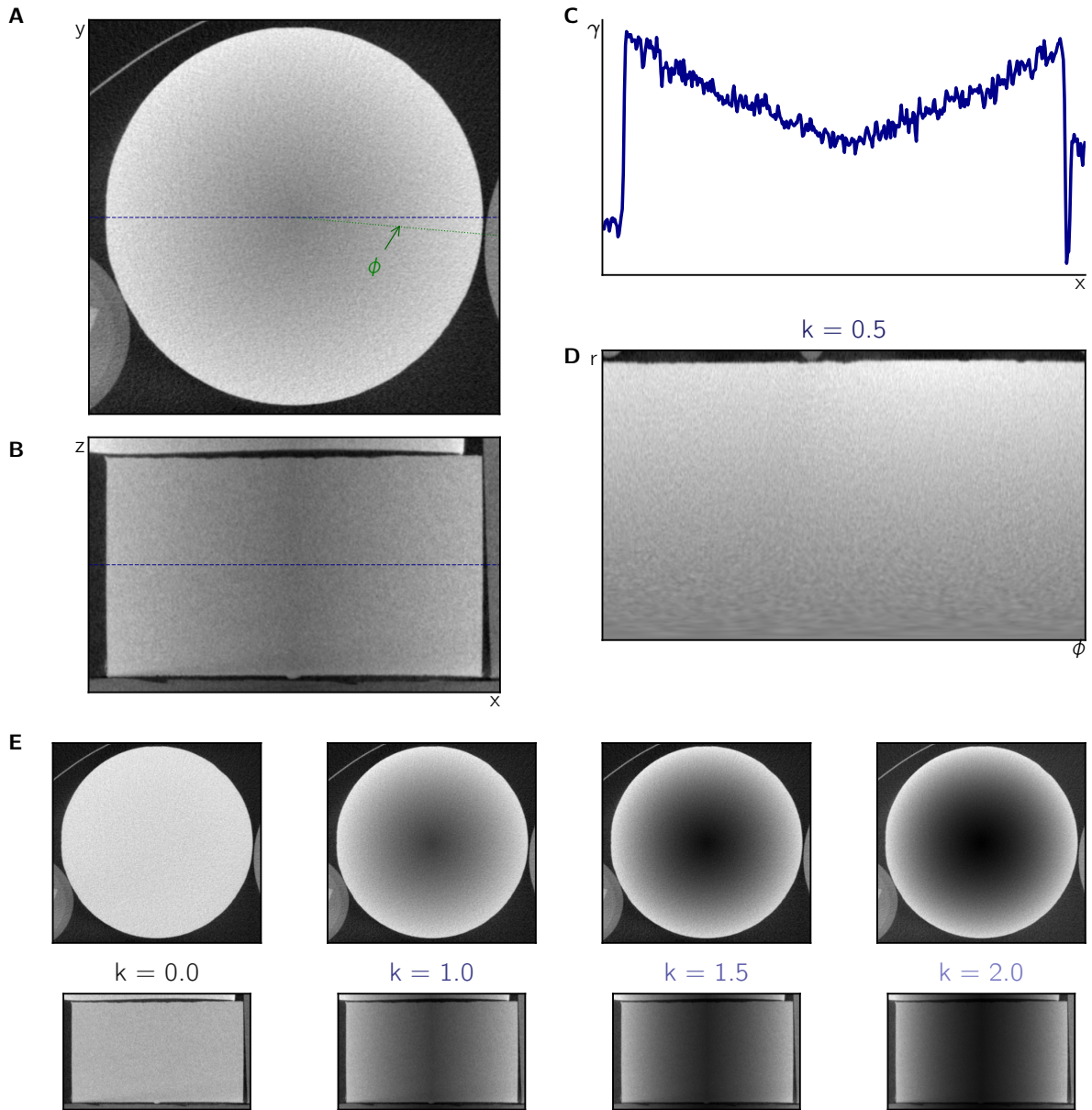


Figure 7.7: Beam hardening artifact simulation. The top view (A) and side view (B) of a PTFE cylinder with beam hardening strength k virtually set to $k = 0.5$. (C) Line profile of scan gray values along the blue line indicated in panel A shows the typical “cupping”. (D) cylinder coordinate transformation of the slice marked by a blue line in panel B gives the gray values along the radial lines r at angles ϕ indicated in panel A. (E) Top and side view for different simulated beam hardening strengths k .

is optionally compared to the “unscaled” gray values of the simulated PTFE cylinders.

7.3.7 Error Propagation

What magnitude do miscalculations of the moment of inertia have, and how much of a problem is that for calculations of joint moment? To evaluate this, one has to consider error propagation (Arroyave-Tobon et al., 2022; Hughes and Hase, 2010; Myers et al., 2015) with the balance equations (7.6). The variable I only enters the balance equations through the dynamic wrench, $\vec{M} = I \frac{d^2\theta}{dt^2}$. This function is linear in I , so error propagation is quite simple⁶ (Normann, 2016):

$$\Delta \vec{M} = \sqrt{\left(\frac{d\vec{M}}{dI} \Delta I\right)^2 + \underbrace{\left(\frac{d\vec{M}}{d\left(\frac{d^2\theta}{dt^2}\right)} \Delta \left(\frac{d^2\theta}{dt^2}\right)\right)^2}_{=0}} = \Delta I \frac{d^2\theta}{dt^2} \quad (7.6)$$

Sensitivity is linear at this level: any percentage of measurement inaccuracy in I will directly propagate to the joint moment calculation.

Yet the situation is more involved: calculating I via equation (7.3) is itself subject to errors, by inaccuracy in the mass itself (regression), and by uncertainty in the $C\vec{O}M$ position and therefore the position vectors $\vec{r}_i$ of each mass element.

$$\Delta I = \Delta I|_{m_i} + \Delta I|_{\vec{r}_i} = \sqrt{\sum_i \left[\left(\frac{dI}{dm_i} \Delta m_i\right)^2 + \left(\frac{dI}{d\vec{r}_i} \Delta \vec{r}_i\right)^2 \right]}$$

The direct dependence on mass is visible in equation (7.3), uncertainty of the mass elements contribute linearly but are added up; errors in the mass of further away mass elements contribute more to the total error of the moment of inertia.

$$\Delta I|_{m_i} = \sqrt{\sum_i ((\|\vec{r}_i\|^2 E_3 - \vec{r}_i \otimes \vec{r}_i))^2 \cdot \Delta m_i}$$

The partial derivative with respect to $\vec{r}_i$ must be added, because I is taken relative to the $C\vec{O}M$ (hence, $\Delta \vec{r}_i = \Delta C\vec{O}M$). Given that the choice of reference point for moment of inertia can be adjusted by using generalized Steiner, the error in the $C\vec{O}M$ propagates as a Steiner component, equation (7.4), with $s = \Delta C\vec{O}M$. Yet $\Delta C\vec{O}M$ also depends non-trivially on the relative mass of volume elements, equation (7.2).

$$\Delta C\vec{O}M = \sqrt{\frac{1}{\sum_i m_i} \sum_i (\vec{r}_i \cdot \Delta m_i)^2 + \left(\frac{\sum_i \vec{r}_i \cdot m_i}{\Delta \sum_i m_i}\right)^2}$$

⁶It should not be assumed that the error in measuring $\frac{d^2\theta}{dt^2}$ is zero, but I will here focus only on the I component.

With this last equation I have outlined how the target parameter, I , ultimately depends logically on mass elements m_i and their position in different ways.

In practical terms, the expected error can be conveniently computed by iteratively changing the variables of interest by a percent, to infer how “a percent change” in a parameter effects the result of the calculation. This is a numerical “**sensitivity analysis**”. From the analytical considerations above, it can be seen that the variables to change in the numeric computation are the regression parameters ($\pm 5\%$ of their value), the $\vec{COM}$ position ($\pm 5\%$ of femur length, 0.25 mm in each direction), and the total mass ($\pm 5\%$, which are 0.26 g).

7.3.8 Error Estimation

Calculating the effect of an error is only one part of sensitivity analysis. The second, equally relevant part is an estimate of how uncertain parameter estimates actually are. A great way to measure this are probabilistic statistics (Ch. 4). Besides yielding outcomes for the intercept and slope, they also allow to estimate the uncertainty of these parameters by sampling a model residual.

To estimate the uncertainty in the density regression, I repeated the procedure in a probabilistic modeling framework (PyMC, Version 5.6, Salvatier et al., 2016). The linear regression function was almost unaltered:

$$\gamma = a + b \cdot \rho + \varepsilon_j$$

In this framework, the residual ε is calculated, measuring how much residual difference remains after optimization of the regression parameters. In a linear regression, ε corresponds to an uncertainty in intercept (it affects all calibration objects alike). Beyond the standard model residual, I introduced ε_j , with index j referring to the materials used for calibration. Each material thus receives its own residual, all of which are drawn from a higher level normal ε (hierarchical model design). This helps to account for systematic increase of the gray value uncertainty with increasing density (“widening” of the distribution), as it was also observed in Du Plessis et al. (2013).

Hence, this regression serves three purposes:

- confirm results of the conventional least squares regression
- provide a measure for the intercept uncertainty (ε)
- identify potential slope uncertainties (ε_j)

All parameters were approximated by Normal distributions, except for the ε_j (`HalfNormal`, i.e. bound to be > 0). Results from the conventional regression and wide enough standard deviations were used as priors to initialize the sampling. The posterior distributions were

calculated from the last $4 \cdot 2^{12}$ samples (after sufficient tuning of the NUTS sampler, see Ch. 4). Gelman-Rubin statistics and trace plots indicated convergence and well-behaved sampling.

One important distinction from least squares regression is that probabilistic regression does not require averaging of the observed gray values. Hence, the regression took a large number of voxels into account, capturing the variability in gray values of the homogeneous polymer materials. A subset of 10000 randomly selected voxels per material were added to the data array. Subsetting is necessary because the number of observations adds computational load for the MCMC sampler.

7.4 Results

7.4.1 Density Calibration Regression

To establish a transform function which can convert CT gray values γ to physical density ρ and back, I compared different regression functions which fit the observed gray values of different polymer- and other materials of known densities (Fig. 7.8).

The random plastic parts obtained from the workshop have known densities, and most of them show a linear relation when plotting γ against ρ . The exception is PVC, which was therefore excluded from the regression (whether it was truly PVC, or a mislabeled different polymer, cannot be assessed retrospectively). Bone mineral density (BMD) phantoms and water do not fall on that line. Their gray values imply an exponential relation of γ and ρ . Both regressions were performed by minimizing least square difference of observations and the target function. Both the linear and exponential function can be inverted. The femur of unknown density distribution lies right between the two regression lines, but spans a large range of gray values and densities because of its heterogeneous composition (marrow, cavities, muscle tissue, calcified bone, metal markers and glue).

With the regression outcome, it is technically possible to convert any observed or simulated voxel gray value to putative density values, which applies for example to the femur.

7.4.2 Issue 0: Mass Mismatch

To confirm the outcome of the regressions above, I compared the actual mass of the femur (which is 5.271 g) to the outcome of the digital approximation.

As seen in the gray value distributions (Fig. 7.8), the majority of femur voxels have a gray value which falls right in between the two regression lines. This propagates to the calculated outcomes. With the linear regression, the femur mass is calculated to be 6.38 g (+21%). In contrast, the exponential regression yields a lower than actual mass (4.96 g, -6%). Though it appears that the correct regression would lie somewhere between the exponential and linear attempt, there is no data, nor theoretical justification supporting an intermediate hypothesis. Neither is there any reason to think that all biological tissues generally fall in between the presented cases: for other samples, composition might be different.

I conclude that, at least for a femur, the regression method is not reliable.

To exclude that problems arose from the special composition of this isolated bone, I applied a similar regression method to more complex data.

In a second experiment, a piglet body was CT scanned, virtually segmented, and analogously physically dissected to see whether the total calculated masses of the virtual segments match their actual counterparts (Tab. 7.3). The corresponding scan was performed at a different facility and long before the density calibration experiments, and at the time the only available calibration points were air, water, and two density calibration phantoms.

Although an exponential regression was used (under-estimation in case of the femur), this second density regression leads to a over-estimation of the mass of most of the segments, and the specimen total (average +12%).

The contradicting results highlight how erratic the digital density approximation is.

7.4.3 Issue 1: Beam Hardening

To evaluate what effect beam hardening has on the calculated mass moment of inertia, I simulated beam hardening on a well described PTFE cylinder within the scan volume.

The only reliable recovery of the true mass moment of inertia (Fig. 7.9B, horizontal lines) is found to be if there is exactly zero cupping. This generally holds for homogeneous objects: moment of inertia is determined by the geometric distribution of mass elements around the $C\vec{O}M$; if (and only if) the relative mass of all mass elements is well approximated by the density regression (which is only and exclusively the case in homogeneous objects), and only if cupping is negligible, then one can calculate appropriate I values. However, if the object is non-homogeneous (as a femur, or a cylinder with considerable cupping), then the relative density of voxels are not appropriately covered, and I will be incorrect. If there are **any beam hardening or streak artifacts**, which is non-trivial to remove from FBP reconstructed scans, **the moment of inertia will be incorrect**. Mass scaling, which was suggested as a strategy to recover total mass from density distribution, will produce erroneous moments of inertia if artifacts are present. Attention should also be put on the units of the calculated mass moments of inertia: it is in the order of 10^{-7} kgm^2 for this cylinder. This is small, justifying extra attention to numeric accuracy of computational methods as well as sources of physical inaccuracy (e.g. $C\vec{O}M$ miscalculation).

I conclude that the chance of getting an accurate measurement for mass moment of inertia from conventional CT scans is low under the influence of common error sources.

7.4.4 Issue 2: Metal Bead Artifacts and Chemical Composition

Another typical kind of artifacts in CT scans appear around metal beads (Fig. 7.10).

Metals are chemical elements with over-proportionally higher x-ray attenuation than carbohydrate compounds. They absorb a large fraction of the incoming radiation, demanding the use of higher energy for sufficient penetration to avoid shadows arising from missing information in the projections; yet higher energy can be detrimental to contrast on softer tissue. The shadows appear as “streaks” or “starbursts” on CT scans, they complicate segmentation, and are unwanted in XROMM visualizations (Brainerd et al., 2010). As with other artifacts, metal bead artifacts have their source in limitations of the tomographic method (limited dynamic range, rotationally symmetric scan geometry), and might be corrected by appropriate counter-measures. Beam hardening reduction by using metal filter plates can improve dynamic range and contrast (at the cost of exposure). The

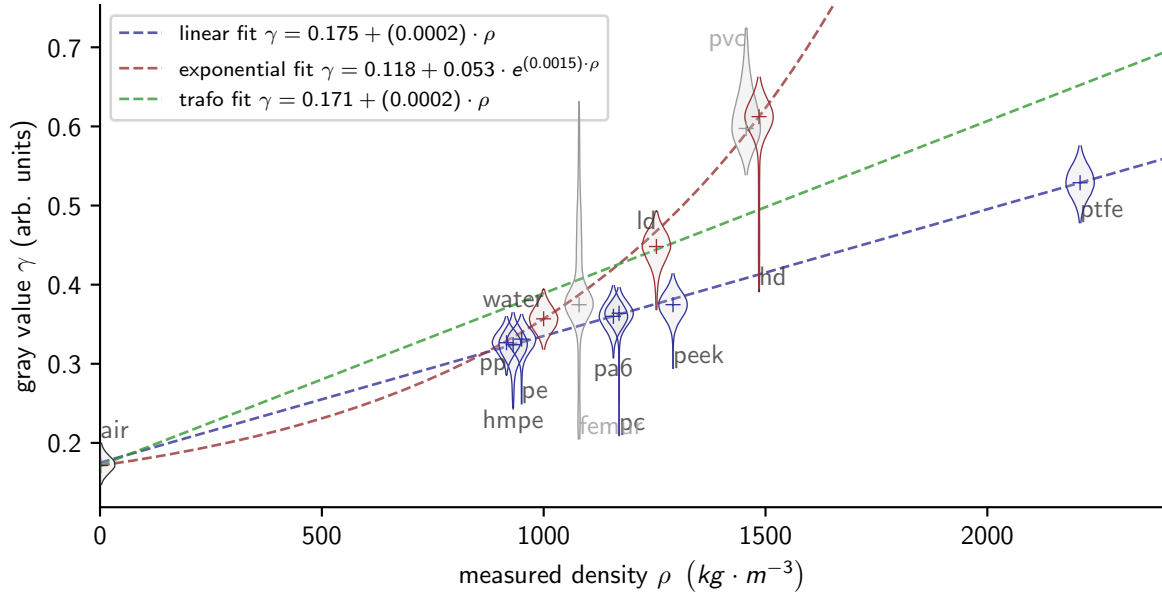


Figure 7.8: Density calibration of a CT scan. The polymer materials show a linear relation between density ρ and CT gray value γ (exception: PVC, excluded from regression), whereas the positions of BMD phantoms (ld: low density; hd: high density) implies an exponential relation. Violins illustrate the distribution of gray values, which appears due to stochastic attenuation, yet each object in the scan is associated with exactly one true (average) density.

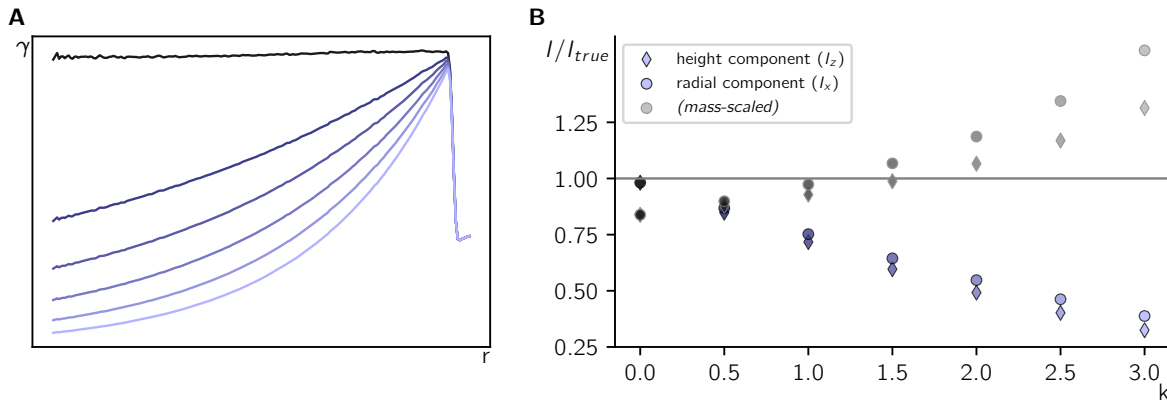


Figure 7.9: Effect of beam hardening on calculated mass moment of inertia of a PTFE cylinder. (A) Repeat of radial profiles (as in Fig. 7.7C) for different beam hardening strengths. (B) Calculated mass moments of inertia, unscaled (blue) and scaled (gray); scaling was optional to make the calculated mass of the cylinder match the actual mass (i.e. normalizing mass, see text). In this cylinder, I_z is always larger than $I_x = I_y$. Calculated values were normalized by the theoretical inertia of the PTFE cylinder (I/I_{true} ; with $I_z = 4.6 \times 10^{-7}$ and $I_x = 3.3 \times 10^{-7}$), hence a correct calculation would yield values of 1.0. The more beam hardening k , the lower the (unscaled) mass moment of inertia. Noteworthy, at $k = 0$ the inertia values match the expected values. However, this is only true for the unscaled curves; when scaling, the correct inertia is reached at an arbitrary, non-zero k , and the calculated inertia increases with beam hardening.

Table 7.3: Segment mass verification on an actual piglet specimen. On the lighter segments, measurement accuracy and rounding are limiting. Note that the total mass at euthanasia was measured 1.125 kg, a slight loss is expected due to condensation and preservation.

segment masses	CT segmentation [kg]	dissection [kg]	difference
head	0.247	0.230	+7%
torso+trunk	0.666	0.580	+15%
LF humerus	0.033	0.027	+22%
LF radio-ulna	0.016	0.014	+14%
LF metacarpal	0.009	0.006	+50%
LF hoof	0.006	0.005	+20%
LH femur	0.052	0.050	+4%
LH tibia-fibula	0.027	0.020	+35%
LH metatarsal	0.012	0.010	+20%
LH hoof	0.004	0.004	+0%
RF humerus	0.031	0.031	+0%
RF radio-ulna	0.019	0.015	+27%
RF metacarpal	0.008	0.006	+33%
RF hoof	0.005	0.005	+0%
RH femur	0.050	0.053	-6%
RH tibia-fibula	0.023	0.021	+10%
RH metatarsal	0.013	0.011	+18%
RH hoof	0.005	0.003	+67%
TOTAL	1.226	1.091	+12%

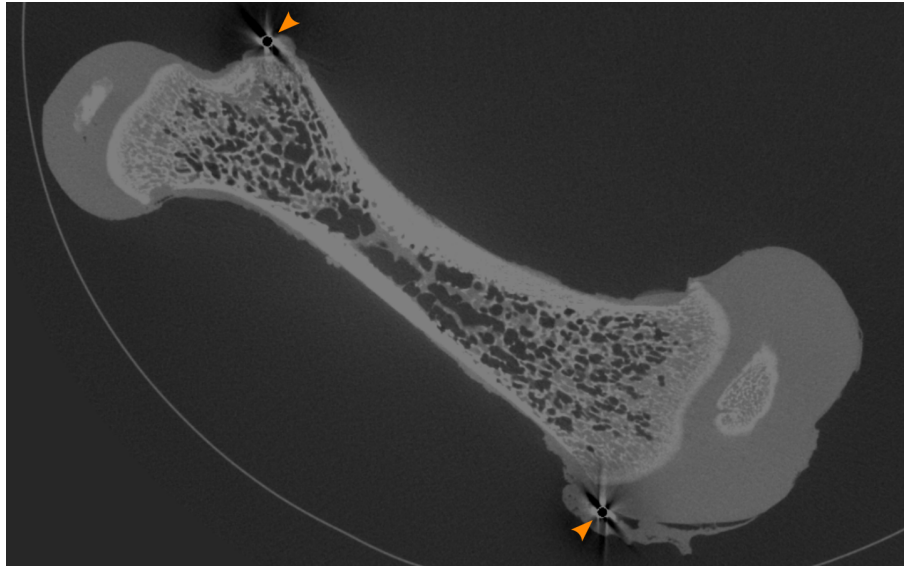


Figure 7.10: Metal bead artifacts, marked with orange arrows. Despite the use of an aluminium filter to harden the x-ray beam, metal beads produce conspicuous, star-shaped streaks on the scan reconstructions. The reason that the center of the metal beads appears black is a clipping setting during image preprocessing (intentional, do ensure they do not contribute to the total mass).

use of a helical scan geometry in a cone-beam CT should reduce the length and contrast of streaks, because projections from above or below the bead elevation level will project through the tissue immediately behind the bead.

Though it is pronounced on metal beads due to their spatial confinement, the problem is not restricted to metal. Streak artifacts also appear between multiple regions of dense bone or strongly absorbent solute (i.e. water), especially if two such regions align along the beam trajectory in the projection geometry.

Bead artifacts are a kind of “beam hardening”, taken to the extreme: the spectrum of the photon ensemble changes along the projection trajectory, and much more so if x-ray interactions with matter are more numerous. They are more numerous in some materials than others (irrespective of density, e.g. formol): chemical composition is key.

7.4.5 Sensitivity Analysis

To put all the aforementioned issues into perspective, I calculated how a change in one parameter affects the outcome of the calculation of inertial properties (Tab. 7.4).

Moderate changes in the density calibration regression directly influence the mass, and thereby the moment of inertia; even the $\vec{COM}$ is mildly affected. The effect on I is in the order of 7 – 8% (determined as Euclidean difference in the principal inertia values). Adjusting the mass has a similar effect as a change in regression parameters, and a 5% mass increase causes a ΔI of more than 8%. Shifting the $\vec{COM}$ alone does affect inertial parameters depending on the shift direction, which is expected (Steiner theorem). The effect of the notable shift (5% of the femur length) is a change in I of about 4%. This is an artificial adjustment, changing the $\vec{COM}$ without a change in mass or mass distribution, but it does help to estimate errors in an inverse dynamics workflow, where a $\vec{COM}$ misplacement will also affect a joint wrench according to the Steiner theorem. Mass changes seem to affect all inertia axes alike, whereas the error due to $\vec{COM}$ shift affects the axes differently. Finally, the exponential regression leads to a moderate shift in $\vec{COM}$ (0.25 mm), but an underestimation of mass (–29% compared to linear regression; –6% compared to actual mass) and moment of inertia (–32% compared to linear regression). These simulations confirm that moment of inertia crucially depends on the mass and position assigned to each voxel in the CT data set; the relative error in the outcome will exceed the input inaccuracy.

To also get an estimate how variable the actual gray value observations were, a probabilistic linear regression was applied (Fig. 7.11).

The intercept was sampled at 0.1734 ± 10^{-4} , and the slope was $1.6052 \times 10^{-4} \pm 10^{-7}$, both consistent with the least squares regression. The result for the model residual ε is 0.024 ± 0.008 , which is 14% of the intercept. Note that reference to the intercept here is in line with using a percent adjustment of the regression parameters above; a more realistic estimate

Table 7.4: Sensitivity analysis of the mass moment of inertia calculations of the femur scan. For reference, the actual femur weighed 5.271 g, had a length of 46 mm, volume of $4.637 \times 10^{-6} \text{ m}^3$ and an average density of $\approx 1140 \text{ kgm}^{-3}$. Regression a and b are fit parameters (see text).

parameter	input change	$\sum_i m_i$ [g]	$\Delta C\vec{O}M$ [px]	I_{xx} [$\times 10^{-6} \text{ kgm}^2$]	I_{yy} [$\times 10^{-6} \text{ kgm}^2$]	I_{zz} [$\times 10^{-6} \text{ kgm}^2$]
(reference)	none	6.38	0	0.52	1.36	1.67
regression a	+5%	6.12	<1	0.50	1.31	1.60
regression b	+5%	6.07	0	0.49	1.30	1.59
mass $\sum_i m_i$	+0.26 g	6.70	0	0.55	1.43	1.75
$C\vec{O}M_x$	+2.5 mm	6.38	50	0.52	1.40	1.71
$C\vec{O}M_y$	+2.5 mm	6.38	50	0.56	1.36	1.71
$C\vec{O}M_z$	+2.5 mm	6.38	50	0.56	1.40	1.67
regression type	expon.	4.96	5	0.42	1.09	1.34

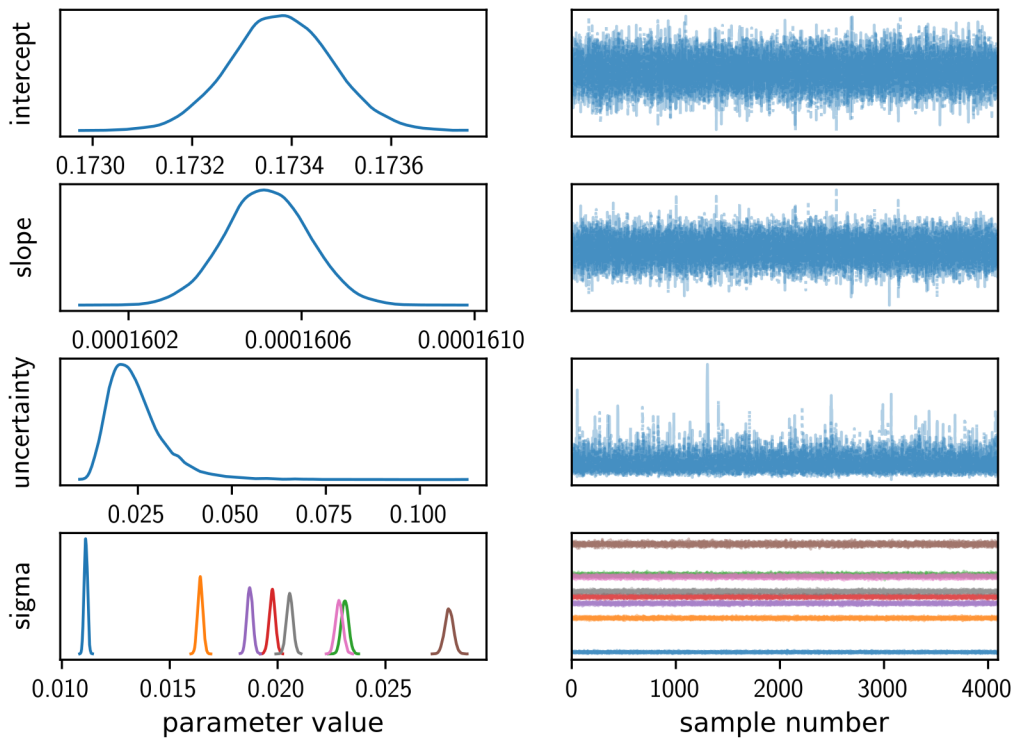


Figure 7.11: Probabilistic density regression result. “Traceplot” showing the posterior distributions of model parameters (left) and sampling traces (right). The “uncertainty” (ε) quantifies the hierarchical model residual, i.e. the total uncertainty remaining after convergence. The colored distributions for “sigma” are the material-wise uncertainties in the gray values (ε_i).

for “percent mass change” would be the average gray value of the femur ($0.39 \rightarrow 6.3\%$); that value is also in the range of the mass adjustment simulated above. Furthermore, the material-wise residuals ε_j turn out to increase with increasing density, yet that overall trend is not systematic for each material. The specific uncertainty for air is lowest of all materials with 0.011, whereas the specific uncertainty for PTFE is 0.021. The highest uncertainty is measured for HMPE (0.023) and polycarbonate (0.027), potentially indicating that these objects were less homogeneous than for example PTFE. Overall, one can attribute roughly half of the model residual to uncertainty in the slope.

These results indicate that the 5% error margins used above are rather conservative estimates: the actual uncertainty is higher, and affects both the slope and the intercept.

7.5 Discussion

7.5.1 Density and Attenuation

In this chapter, I challenge the common misconception that structural information from CT images can be easily converted to a density distribution. Humans think of objects in terms of their mass because that is what we are confronted with in our everyday experience of the physical world. And the structures which we recognize in CT visualizations resemble that experience. However, the raw images quantify attenuation, which is a fundamentally different phenomenon.

And even the voxel-wise attenuation cannot be accurately extracted, for the following reasons. X-ray radiation used in commercial CT scanners is often polychromatic. Absorption is wavelength-dependent in a non-trivial way. The whole process contains stochastic aspects (e.g. scattering). Scientists tend to introduce the constraint that chemical composition must be homogeneous within samples, yet that constraint is neither realistic, nor practical. Finally, scan “artifacts” are often described as problematic, but as argued above they might just be the manifestation of the inappropriate modeling of a normal physical process by our limited arsenal of reconstruction methods.

7.5.2 Miscalculated Mass

The thorough measurements reported above confirm that the outcome of a CT scan cannot be directly related to physical density. I followed the strategy of others, using a kind of calibration regression to convert CT values to density (Du Plessis et al., 2013; Durston et al., 2022; Fath et al., 2023; Mull, 1984; Phillips and Lannutti, 1997). The presented measurements show severe differences between the actual density and the calculated values, with errors of more than 20% of the actual mass.

Neither plastic parts, nor calibration phantoms provided an accurate reference for the density of the test object (a piglet femur). The bone mineral density phantoms arguably serve a different purpose (calibrating *mineral* density), but resemble the compact parts of the bone in their chemical composition. Yet it is not only the chemical composition of the scanned object, but also the spatial arrangement of its components, because the trajectory of the x-ray beam affects attenuation (“self-filtering”). It is insufficient to just assume “comparable chemical composition”. A calibration object should also have the same shape and spatial arrangement as the object of interest. The ideal object to calibrate an unknown femur would thus be an identical femur of known mass distribution (which is actually attempted in clinical tomography, cf. Lennie et al., 2021).

The reason, as illustrated in the introduction, is that photons of different wavelengths have specific absorption probabilities. The spectrum of the beam along its trajectory changes at each instance in the tissue. The most obvious example of this is beam hardening, and

I showed by a simple virtual simulation that beam hardening leads to a systematic under-estimation of mass and moment of inertia (Fig. 7.9). In a more extreme case of this, radio-opaque metal beads will absorb almost all passing photons (“photon starvation”), leading to problems with conventional reconstruction algorithms which appear as starburst artifacts.

Recent developments attempt to improve the reconstruction algorithm (Batenburg and Sijbers, 2011; Frenkel et al., 2022; Six et al., 2019; Yang et al., 2021), but their application requires prior information and extra work. Others might eagerly await “spectral CT” (Liu et al., 2023), yet technical and economic challenges are still limiting the capability and availability of such machines. Dual-/Multi-Energy CT might be another route to overcome the limitations of monochrome detector screens, and they might combine well with iterative, model-based reconstruction techniques.

7.5.3 Non-Intuitive Inertia

It is perfectly normal that physical measurements are subject to measurement inaccuracy. However, the error range must be quantified and reported.

Above, I demonstrated an analytical assessment of which error contributions should be expected, as well as a numerical sensitivity analysis. I also quantified the actual regression uncertainty with probabilistic statistical modeling. The results enable an evaluation of the relative importance of different parameters for the accuracy of the outcome. For example, changes in the mass or in the regression parameter values directly influence the moment of inertia, at a rate higher than the input error. A $\vec{COM}$ miscalculation might cause errors of lower magnitude, but those are not necessarily isotropic.

Yet by far the greatest effect was due to a change in the regression type, switching from linear to exponential. Neither of these is accurate, yet in the absence of a model which covers the nature of attenuation (which would rather have to adjust the reconstruction algorithm than the regression), there is no other reasonable option. The exponential fit seemed plausible from the BMD phantoms, which are designed to resemble ossified tissue. The majority of observed femur gray values were not far off the exponential regression line. Nevertheless, the calculated moment of inertia was about 30% off the reference. Note that this error is the sum of both an over-estimation in the linear reference regression, and an under-estimation of I in the exponential attempt. Also, the high percentage difference is explained by the fact that moment of inertia values are relatively small in baseline magnitude. A shortcoming of the present study is that I am unable to determine a “ground truth” inertia for the test object, lacking sufficiently accurate physical measurement apparatus and expertise.

Nevertheless, the presented findings have practical application. Any regression should be as accurate as possible (optimal exposure and dynamic range, noise reduction e.g. by more averaging, beam hardening filter, more/diverse calibration objects). A probabilistic regression is favorable, since it yields error margins for the regression parameters. With

those, a sensitivity analysis can and should be performed. Results I obtain are technically accurate, confirming that cheap plastic parts are sufficient and might be preferred in most situations over expensive calibration phantoms.

I also evaluated that scaling voxel-wise density, so that the mass matches the known total mass, is not an option: this modification can amplify issues from beam hardening (Fig. 7.9) or from marker streaks and thus make the outcome more unpredictable. The rationale behind scaling was that, for moment of inertia, the distribution of densities is relevant. Scaling masses is identical to adjusting the intercept parameter in the linear regression (Tabs. 7.2, 7.4). An alternative, which I did not attempt, would be to adjust the regression slope so to match the mass checksum. However, both these adjustments (regression intercept/mass scaling, regression slope) have conceptual flaws. Firstly, the resulting modified regression would not correctly fit any of the calibration objects, and the question arises which of the measurements is “more trustworthy” or “more true”. Second, slope adjustment over-values the “air” observation: “air” has a negligible density, yet might have a non-negligible attenuation (which is visible from its gray value distribution; x-ray flux strongly depends on source-detector-distance in non-vacuum). This (energy-dependent) air issue might actually explain the aforementioned systematic errors in the supplemented regression plots of Durston et al. (2022). Third, again, there is no mechanistic reason to assume a linear relation; the polymer plastics of identical chemical composition are a rare corner case, most other series of objects could be non-linear.

7.5.4 Inertia Components

To my knowledge, this study is the first to perform a detailed sensitivity analysis of the CT density calibration. Sensitivity is high, so are the measured error margins, which implies that natural inaccuracies are potentially fatal for the derived measurements.

Yet whether or not an error margin is “fatal” also depends on the field of application. The purpose of the inertial estimation is to calculate joint moments in an inverse dynamic workflow. Joint moments of a limb in ground contact are dominated by the ground reaction force, i.e. by the gravitational load of the animal and by the dynamics of the heavy torso and head (which constitute 75% of the mass of the animal, Tab. 7.3). For grounded limbs, the inertial properties of the individual limb segments themselves are negligible (order of magnitude of limb segment mass *versus* total mass), although, depending on the behavior under study, accelerations might be high and caution might be appropriate. When calculating the inertia of limbs in the swing phase, the inertial contributions of more distal segments dominate the joint moment calculations. In any of those cases, measurement inaccuracy should receive a close inspection.

There is another aspect to the miscalculation of inertials: the “rigid body” assumption. Segments *in vivo* are not rigid. Anisotropic muscle contractions and internal segmental mass

movements (e.g. radius-ulna rotation) will shift the $C\vec{O}M$ and might even rotate the inertial axes. This is practically hard to capture for most dynamical behaviors (though it might be modeled with dynamic CAD models), but the magnitude of this problem depends on the study question.

Another factor is the preservation of the specimen (frozen in this case; formol, alcohol, staining, ...) which has an influence on the chemical composition and inertial properties (not assessed herein). Finally, in vivo movements are often restricted to one or few axes. If a femur swings forward during walking, it will mostly rotate around the medio-lateral axis. In that case, the moment of inertia around that axis is relevant. The segment might slightly rotate around its long axis, but long axis rotation will hardly affect total momentum due to a low I on that axis. Thus, in this case, the (uncertainty of the) medio-lateral inertial component demands special attention.

7.5.5 Lack of Alternatives

The femur CT scan used in this experiment was recorded with all necessary diligence, under ideal conditions, with good image quality and resolution. For example, a beam hardening filter was used, sample mounting was flawless, contrast and dynamic range were optimized. Still, I see a distribution for gray values which hinders determining the femur mass at an accuracy better than 10% (see probabilistic regression).

The proximal conclusion would be to reject the density estimation from CT data. However, such an attitude demands the evaluation of alternatives.

The mass can be measured accurately with sufficiently fine scales, yet dissection of segments to determine segment-wise mass literally introduces fuzzy edges. Center of mass can be estimated by pendulum methods, yet in addition to the dissection issues, the problem of accurately relating the $C\vec{O}M$ to an anatomical reference frame arises. The conventional instrument to measure moment of inertia is the trifilar pendulum (Korr and Hyer, 1962; Schedlinski and Link, 2001; Wells and DeMenthon, 1987). As Durston et al. (2022) have demonstrated, this method is far from trivial, requires a well-prepared setup, and still suffers from limitations (e.g. that all axes have to be measured separately, and then transformed to an anatomical reference frame). Pendulum methods usually rely on negligibility of the instruments' masses and on the small angle approximation. Hence, the accuracy of these methods remains hard to estimate.

A slightly more involved approach is the utilization of model-based inertials (Allen et al., 2009; Hutchinson et al., 2007). This approach could be termed the “indirect CT-based” method: volumetric images are segmented into volumes of homogeneous tissue, and tissue types can be assigned their approximate physical density. Full CAD models might be assembled, and to a certain extent these might even incorporate dynamic mass redistribution during locomotion. The method requires good soft tissue contrast of the

scans, e.g. by applying staining protocols. Since tissue is not homogeneous, but continuous transitions can occur (e.g. muscle-tendon), a validation and error estimate of this approach would be required. The main challenges of this approach are the workload of segmentation, and again transferability to *in vivo* measurements

Finally, an alternatives to CT scans for the present purpose might be *better* CT scans. Projection images are monochrome, not monochromatic, and the step towards “spectral CT” might improve density estimation by incorporating models of chemical composition. This will doubtlessly require technical and computational innovation such as new algorithms, most likely algebraic ones which incorporate a measured source spectrum, which iteratively adjust a model volume to match simulated and real CT projections (which would then be polychromatic). All those are promising developments that would certainly impact the accuracy of inverse dynamic models derived from XROMM experiments.

Neither of these options currently yields quick, transferable, generalizable results, which might seem like CT gray value regression should be the method of choice, despite considerable error. However, conceptual and practical issues from CT gray value conversion uncovered in this study demand caution and careful evaluation of the limited circumstances in which CT scans may yield usable inertial approximations. Doubts about measurement accuracy must count as an obstacle in the ethical evaluation of planned experiments. It must be stated that there currently is no validated method to reliable and efficiently measure dynamic inertials.

7.6 Summary

At the end of this chapter, it should be clear that using CT images to calculate inertial properties has a number of conceptual and computational pitfalls which can put serious doubt on the results. Tomographic images depict attenuation, not physical density, and though these are occasionally correlated, they remain fundamentally different properties of matter. Above, I discussed potential flaws in the CT-based approximation of mass, center of mass, and mass moment of inertia, as well as corner cases in which the calculation might co-incidentally yield appropriate results. In all other cases, the calculation will certainly yield wrong results. The magnitude of the error is non-negligible, since some axes of inertia in balance equations have very small base magnitude and are therefore susceptible to faults. However, with all that objection put to paper, it must be stated that there is currently no reliable, efficient method to measure inertial properties of biological systems. There is no “density tomography”, CT-based regression is error-prone, and common alternative methods potentially suffer from equally large error margins for other reasons (e.g. dissection inaccuracy, non-static configuration during movements, preservation issues).

Using attenuation values to calculate densities might be about as accurate as setting $\pi \equiv 3$ in physics classes (a totally valid thing to do, if at the same time setting $g \equiv 10$). Typical pitfalls (beam hardening, metal beads, symmetries) are known and can be avoided, keeping track of error magnitudes is essential (by performing a sensitivity analysis, involving e.g. Gaussian error propagation and probabilistic regression). Under this condition, the calculated inertials might be a just sufficiently accurate approximation of the true values for some experimental questions; yet accuracy must be validated and cannot be taken for granted.

7.7 Acknowledgements

A significant part of this project was supported by the following fabulous people. A lot of the personal insights about x-ray theory herein have been made possible by my time at the FleXCT (2022-2023). Thanks to Jan Sijbers and Jan De Beenhouwer for this opportunity, and special thanks to the colleagues at the VisionLab, especially Joaquim Sanctorum, for the great discussions! Jonathan Brecko (“naturalscience” museum Brussels) and Björn De Samber (FleXCT) were the operators performing the CT scans. Thank you to Chris Van Ginneken for the most accurate piglet dissection possible (which I was allowed to assist). Jan Scholliers designed and printed the most ingenious radio-translucent clamp-on-a-motor construction possible, thank you also for the great support on all the parts of this project! Maja Mielke, Joaquim Sanctorum, Peter Aerts and Chris Van Ginneken provided helpful comments on an earlier version of this manuscript.

Part IV

Epilogue

Chapter 8

General Discussion

8.1 First Principles

The three parts of this thesis each had a slightly different focus, spanning the range of methodologies which is characteristic for biomechanical research (kinematics, dynamics, statistics). They follow a common scheme: each of the parts took a basic principle of physics research and applied it to the study of biomechanics.

Part one was concerned with kinematic data analysis, and the generation of meaningful measures from videos of moving animals. The basic principle therein is coordinate transformation (Tipler and Mosca, 2007): physicists are well aware that computational problems simplify if one converts the data to an appropriate coordinate space. For locomotor kinematics, I reproduced and applied a suggestion of others (Bernstein, 1927a; Webb and Sparrow, 2007) which relies on the repetitive, periodic character of many locomotor behaviors (Ch. 2). By applying a Fourier Series Decomposition, it is possible to transform joint angle profiles to the frequency domain, where their mathematical structure is much simplified (affine components are accessible; time series are represented by an array of harmonic constituents, instead of a raw signal with an undefined number of noisy samples). The Fourier Series method is a flexible tool which opens up a data set to multivariate statistics (Ch. 3). It is also a prerequisite of the subsequent part, which could hardly be adapted to untransformed kinematic measures.

Part two focused on statistical modeling of kinematic measurements. It built on the application of another physical principle: namely that most measured variables do not take exact values, but rather follow probability distributions (LibreTexts, 2022b). This is especially true in the biological sciences, where variability in a trait, in reproductive success, and in the correlation of these two are the key prerequisites of our central working model (Darwin, 1859). Statistical models which incorporate variability, though rarely applied in that context to date, are useful in biological applications (De Groote and Falisse, 2021; Røraas et al., 2019). I provided a brief tutorial to outline why and how probabilistic models work in practice (e.g. the MCMC sampling methodology, Ch. 4). In that framework, I applied probabilistic, predictive modeling to a comprehensive data set of piglet kinematics (Ch. 5). Prediction in conventional statistics would be deterministic, i.e. yielding averages, at best with the vague guidance of standard deviations, which only superficially capture the stochastic, variable nature of the phenomenon. I demonstrated that probabilistic prediction can be used to generate relevant insights into locomotor maturation of piglets.

Finally, the third part of the thesis was devoted to inverse dynamics. The goal of inverse dynamics are estimates of joint forces and moments, which can be calculated from measured or estimated external forces and moments by treating limb segments as rigid bodies, and by (stepwise) calculation of their balance of wrenches (Ch. 6). Rigid bodies must be associated with inertial properties (mass, center of mass, mass moment of inertia), which are crucial

prerequisites for the balance equations. As I then demonstrated, the assumption that inertial properties can be accurately derived from x-ray computed tomography cannot stand up to closer inspection (Ch. 7). They can be derived, but not accurately, yet they might provide the “best educated guess” available. This finding was enabled by a third basic principle commonly applied in physics (which is actually a first principle): the quantification of a measurements inaccuracy (Hughes and Hase, 2010).

There are certainly more examples of basic principles in physics which would deserve application in biology in general (e.g. quantum cognition, Aerts and Aerts, 1995; Busemeyer and Wang, 2015), or in biomechanics in particular (e.g. wrenches and quaternions, Dumas et al., 2004).

8.2 Predictive Power

Yet are these adaptations from the physical sciences inevitable? Why/when should conventional methods be questioned and compared to the yet uncommon methods suggested in this thesis?

The two decisive analysis characteristics are the occurrence of variability, and predictive modeling.

Variability is omnipresent in biology, and I find it surprising how many studies use statistics which model point estimates for actually variable measurements.

Yet the ultimate quality criterion of statistical models (and classifiers, and neural networks/“AI”) is how well they can be used to predict experimental circumstances which deviate from those they were trained on, or in other words how they transfer to novel circumstances. This refers to the distinction of descriptive, explanatory, and predictive modeling (Shmueli, 2010; Shmueli and Koppius, 2011). In biomechanics, explanatory modeling is rarely applied, because the cause for animal movement are known to be physical first principles (conservation of momentum, Ch. 6). Virtually all quantitative analyses in locomotor science are descriptive: they “summarize or represent the data structure in a compact manner”, and “focus is at the measurable level rather than at the construct level” (Shmueli, 2010). For example, spatiotemporal gait characteristics (e.g. dimensionless speed) are often measured and compared between study groups. There are numerous recent examples of studies in which spatiotemporal gait variables are the basis and (own) goal of comparison (e.g. Cheu et al., 2022; Druelle et al., 2021; Ekhtator et al., 2023; Jones and Hasiotis, 2023; McHenry and Hedrick, 2023; Plocek and Dunham, 2023; Young et al., 2023b). However, it is rarely modeled how and why a study group reaches a certain speed (explanatory statistics), and the predictive power of studies on spatiotemporal variables is low (e.g. we could not predict the speed of an individual with a stiff joint from measuring normal individuals).

Some examples should illustrate how predictive modeling could complement existing analyses.

One exemplary study of kinematics was performed by Hutchinson et al. (2006). The authors tracked landmarks on an enormous number of strides of different species of elephants¹. Their measurements are extensive (in terms of the number of derived quantities), yet descriptive. The authors discuss whether more or less subtle discontinuities in the spatiotemporal measures actually reflect gait changes (Alexander, 1989), but emphasize the overall smooth speed-dependence of spatiotemporal parameters. They put some effort into a convincing argumentation that a subset of their measurements were close to what can be considered the maximum speed elephants can reach. And they document the restricted repertoire of elephant gaits: these animals seem unable to trot or gallop, reaching the highest measured locomotor speeds with their default gait.

Yet, if joint kinematics are available, these questions could be addressed and expanded in the frequency domain: how does the ensemble of Fourier coefficients change with speed (see Ch. 3; amplitudes or mean of specific joint angles will certainly change), and could we predict what happens at higher speeds (limits to effective ROM)? Could we even predict which joint loadings are critical if we set up a virtual elephant model to gallop? Are African and Asian elephants really identical in their intralimb coordination, or do they achieve the same spatiotemporals with altered joint profile patterns (analogous to Ch. 5)?

A particularly well studied species are humans, yet inference about our ancestors is tricky (Cazenave and Kivell, 2023; Polk, 2004; Stamos and Alemseged, 2023). Available techniques rely on morphological information and inverse kinematics, i.e. the inference of kinematics from output patterns such as fossilized footprints, and dynamic constraints such as stability (Hatala et al., 2016; Nicolas et al., 2006, 2007; Pronost et al., 2006). Interesting debates are ongoing with regard to the bipedality and arboreality of particular fossil specimens. Note that the methodological advances presented in this thesis could be used for quantitative predictions of hominid (or any other animals’) locomotion. Trained with locomotor data of e.g. extant humans and chimpanzees (phylogenetic bracket), a probabilistic model which relates anatomical, morphological, and developmental traits to joint angle profiles could be tuned to predict stride cycles of intermediate forms (Ch. 4). The emphasis is on “stride cycles”: this refers to predicting complete joint angle profiles, from which all corresponding joint positions can be recovered (information retention of the Fourier Series, Ch. 3). These predictions could be turned into dynamic simulations and virtual animations of the complete behavior, and matched to trace fossils (as in Nyakatura et al., 2019, but with less manual work). Adding qualified estimates of inertial properties (Ch. 7) to the probabilistically predicted strides would enable forward dynamic modeling, and an estimation of the plausibility of certain hypothesized gaits (e.g. by comparing

¹For reference: this was at a time when high-end consumer GPUs housed about 256MB of memory; long before the era of DeepLabCut.

energy efficiency). The possibility of computing independent parameter spaces for posture, range of motion, and coordination (Fig. 8.1) enables a more fine-grained exploration of possible ancestral locomotor modes.

These two examples illustrate how the methodological advances presented herein could extend our knowledge about terrestrial locomotion, with datasets which already exist. There are multiple potential goals of predictive approaches (Shmueli, 2010), such as hypothesis generation, measurement comparison, improvement of explanatory models, and assessing predictive accuracy of a model. The latter goal was targeted herein in a comparative study of newborn piglets.

8.3 Transferability

Explanatory or even predictive considerations require returning to the construct level, and for legged terrestrial locomotion this would be the level of the joints, e.g. measuring the (joint) angles between segments. Yet there is a standing problem with that construct: joint angles are constrained by morphology, which has been associated with issues of transferability (Gatesy and Pollard, 2011). This complicates comparison of joint angles between morphologically disparate groups (e.g. between species, or birth weight categories). For example, Gatesy and Pollard (2011) assess that motion transfer between a human and flamingo fails², causing “undesirable motion artifacts” due to morphological discrepancy^{3,4}. They also speculate that underlying patterns of coordination could be invariant to morphological disparity, yet they expose a lack of methods to directly assess coordination. As shown herein, FCAS enables the isolation of affine components of joint angle profiles, and, by partly removing them, the extraction of measures of coordination. I have demonstrated how to generate relative joint angle profiles, which are direct measures of coordination. Take a hypothetical lineage which evolves towards elongation of a single segment (non-isometric, i.e. the other segments would not elongate similarly; Fig. 8.1): to retain the ability to walk despite altered morphology, there are three options. These animals must either (i) change their posture, e.g. by walking with a more extended limb. In other cases, e.g. an elephant-like, straight limb situation as the “null” posture, they must (ii) adjust their effective range of motion to retain locomotor capacity. Then, (iii) an adjustment of the timing of joint angle changes, i.e. coordination, would follow. This example re-emphasizes the importance of being able to separate affine components (i, ii) and to process the remaining variables (iii) alone. The path for future researchers would be

²Cf. <https://www.youtube.com/watch?v=1Wh8d9b2Wrw>

³Another popular example for “motion transfer” is the transfer of measured running kinematic data of chicken, *Gallus domesticus*, to animate a *Tyrannosaurus rex*, cf. <https://www.amnh.org/exhibitions/dinosaurs-ancient-fossils/theropod-biomechanics/walk-dont-run>.

⁴These considerations must be flanked by discussions on dynamic similarity (Aerts et al., 2023b).

to select a suitable phylogenetic branch or example (maybe Tragulidae or Giraffidae, or more broadly Ruminantia), apply FCAS, and measure how a multivariate block of morphological variables correlates to affine and non-affine components of the joint angle data (a suitable method might be Two-Block Partial Least Squares, Rohlf and Corti, 2000). FCAS is the door-opener: at first instance, it can help to identify the mismatching aspects which prohibit motion transfer (Gatesy and Pollard, 2011). It allows for the separation of different logical components of kinematics (especially coordination; Fig. 8.1), potentially making some of them transferable. Motion transfer is in fact the extrapolation of kinematic aspects of one species or comparative category to another, i.e. a predictive problem. The application of a predictive model of NBW kinematics to LBW in this thesis is an implicit form of motion transfer: which age/size would we assign to an NBW which walked like the LBW we observe? FCAS thus enables predictive modeling at the construct level of locomotion.

The Fourier-method is not always applicable, but often. Oscillation, i.e. that a limb configuration changes repetitively over time, is a common scheme in all kinds of locomotor patterns: be it a baboon hindlimb flexing-extending-flexing-extending on a bipedal walk to a water hole (Druelle et al., 2021), or be it an ungulate head-nodding in rhythm with its locomotion (Loscher et al., 2016). There are multiple reasons for the abundance of oscillation in vertebrate locomotion. The bone-joint system of vertebrates is well approximated by rigid bodies, which can rotate, but hardly translate with respect to each other. There are other structural constraints, e.g. elastic elements such as the nuchal ligament which has a crucial role in head nodding. There is neural organization: brain circuits are known to exhibit oscillations as well (Gupta and Chen, 2016), which might project downstream. Furthermore, the evolution-like optimization of behavioral tasks: there is little use to gain efficiency in a one-time control unit, yet repeated modules of a task are worth adjustment; this situation might have favored recurrent locomotor modes (but note that those are not omnipresent: think of cephalopod benthic crawling). In the many cases which show repetitive or recurrent behavior, Fourier analysis is the natural first choice for transformation.

8.4 Methodological Advances

To summarize: the Fourier method presented herein complements existing analysis strategies for locomotor kinematics by transforming the data to the frequency domain. This enables the isolation of affine components, which quantify the dynamic posture, effective range of motion, and the coordination of segments. Though these components could in principle be extracted from raw profiles, the transformation yields them directly. The advantages are the following.

- The transformation changes the **data structure**. Fourier coefficients are more accessible to multivariate statistics, such as PCA and statistical modeling (Ch. 3 and 5).
- **Information retention:** FSD conserves information and is reversible, which means that any model interpolations or predictions can be converted back to kinematic data, which enable dynamic simulations (e.g. to test for stability).
- FCAS enables **motion transfer within** a sequence of joints: some or all affine components can be transferred to adjacent joints (Ch. 3), leaving the residual, relative movement, i.e. coordination.
- The method is **complementary** with existing analysis strategies: for example, one can revert the “coordination” residual of an FCAS to temporal profiles and perform statistical parametric mapping (appendix 9.3).
- **Motion transfer among** study groups: because of the separation of affines and the altered data structure, one can study and test evolutionary trajectories (again, via simulation).
- Fourier Series and its inversion are trivial to apply, they are **deterministic transformations**, and code is readily available or easy to obtain in any common programming language (Ch. 2.7.1).

These new possibilities come with a major limitation: Fourier Series decomposition demands temporal periodicity. Joint angle profiles must be cyclic, i.e. end at the angle where they started. Another limitation is that rather complex profiles might theoretically require a rather high number of coefficients, as do situations with multiple joints and 3D angles (which were not exhaustively explored in this work).

On the other hand, the nature of locomotion is that there are oscillating limb elements which often move in cyclical patterns. There are non-invasive tricks to achieve cyclicity, some of which I discussed herein (e.g. end-start matching, 2.4.4).

This connects the first two parts of this thesis to the third one. It would have been desirable to link Fourier coefficients of piglet kinematics to dynamic simulations of their limb segments. For example, an immediate question is which segment movements are responsible for which components of joint moments. Because of the inconclusive attempts to extract inertial properties from CT scans, and because of the reported inaccuracies in markerless motion tracking, I failed to establish such links between kinematics and dynamics (Ch. 7). I anticipate that frequency domain considerations hold some unexplored potential in this direction.

Finally, there is some benefit in using probabilistic frameworks when modeling locomotion. The motor output of biological systems is intrinsically variable, and an animal which is

stable on one stride can tumble on the next one. Working with likelihoods and probabilities of kinematic patterns seems to better reflect the nature of the process.

8.5 How and Why Piglets Move

The primary goal of this thesis was to falsify the hypothesis that there are differences in locomotion of low- and normal birth weight piglets (LBW and NBW, Ch. 1). There are two main aspects in which the study groups could differ: in kinematics (*how* animal limb segments move) and dynamics (*why* animal limb segments move as observed). And there are putatively confounding effects: locomotor maturation (i.e. age), size and weight of the animals (i.e. physical appearance, cf. Aerts et al., 2023b), and the variability of locomotor behavior.

Previous studies from our group have compared kinematics on the level of spatiotemporal characteristics of newborn piglet gait (Vanden Hole et al., 2018a). They discussed delayed locomotor maturation as a proximal explanation of the observed difference. I herein extended their analysis and confirmed maturational delay by training probabilistic models on the FCAS data of a large number of NBW piglet strides, applied to LBW observations (Ch. 5). The FCAS data entering the model contains practically all kinematic information which could be used to distinguish the study groups. The predictive models accurately incorporate the variability of the phenomenon. The models predict size and weight of the LBW piglets as if they were “normal”, i.e. there are no apparent kinematic differences which would lead the model to infer non-normal subject characteristics. This finding must be taken with the discussed grain of salt: conceptually, indifference is hard to prove with the predictive modeling strategy chosen herein, and of course more data would be desirable to settle the case. The only differences found are related to the age of LBW piglets. This indicates that, if models would not consider age or maturation as a model factor, low- and normal birth weight piglet locomotion would be indistinguishable. The “age” model provided evidence that maturation of LBW locomotion halts approximately four hours *post partum* (developmental delay). We suspect competition within litters and limited metabolic reserves as the reason for developmental delay (Vanden Hole et al., 2019).

Note that these observations are restricted to walking gait at voluntary speed; we cannot exclude more obvious LBW/NBW differences emerging in more challenging motor tasks. However, the apparent developmental delay in even this basic motor skill indicates the high sensitivity of the method. Also, the relevance of any hypothetical, “more challenging” stair, treadmill, or wind tunnel experiment for animals whose major priority is to find and access their mother’s teats is questionable. Overall, the models provide at least some evidence against the existence of fundamental differences in *how* LBW and NBW piglets move: their coordination seems innate.

In fact, this lack of differences in the collective output of the animals neuromotor system

is surprising. The FCAS data quantifies joint angle profiles - which are indifferent to size. Yet piglets of low- and normal birth weight operate under different body mass constraints: NBW are about twice as heavy as LBW. To produce the same kinematics, one would think that they must tune their motors differently, like using different gears. These would be differences in *why* we observe the (indistinguishable) movements. Different weight implies different physical constraints, size does matter (Aerts et al., 2023b).

Yet we encountered conceptual and practical issues in addressing this second big question (i.e. *why* they move as observed). The first issue is a practical one (Ch. 7): inertial parameters, which others have extracted from calibrated CT scans, cannot be determined at sufficient accuracy. Then, again, there is the issue of variability, as evident from the study of kinematics (Ch. 5): differences between NBW and LBW (when normalized for size) seem to be lower in magnitude than the effect of maturation, and lower than normal variability of the behavior. This was actually confirmed by analysis of 3D kinematics (which we did, unpublished): we originally attempted the full experimental procedure for XROMM. Of course those recordings are variable, but of the 392 recordings we obtained in several recording sessions in summer 2019, not a single one showed locomotion which strictly fulfilled the “steady state” criterion and gave single limb forces (hitting the force plate correctly was likely, but not guaranteed, and happened in 189 recordings, 85 of them walking gait). Not a single stride could be considered “representative”, because young piglets in our setup literally always tumbled, jumped, sidestepped, stopped shortly after, or interfered with each other, both LBW and NBW. The XROMM and inverse dynamics procedure is a low through-put method, and it was not possible to track landmarks on a sufficient number of strides to assess LBW differences in joint moment orchestration on the background of motor variability.

These issues notwithstanding, I succeeded in porting the XROMM workflow (Brainerd et al., 2010) to an alternative software environment (cf. https://git.sr.ht/~falk/foss_xromm, Fig. 8.2). This workflow includes the complete calculation of inverse dynamics, i.e. joint forces and moments can be visualized. The software conventionally used for XROMM caused issues: due to its large dependency tree and lack of documentation, XMALab (Brown University, USA, <https://xromm.org>) was dysfunctional on the Linux operating system for about half a year during the course of this project; Maya AutoDesk is non-free software. Except for ORS Dragonfly (Object Research Systems, Canada, <https://www.theobjects.com>), which was used for CT segmentation but could be replaced by 3D Slicer (Slicer Community, <https://www.slicer.org>), all the tools in the adapted workflow are free and open source. The example animation (<https://ody.sh/fZe30pgHGN>) can be viewed in Blender (The Blender Foundation, <https://blender.org>) from any perspective, and replayed at self-chosen speed. At the current (incomplete) stage of development, we observe segment position glitches when markers enter or leave the focal volume of the biplanar x-ray. Also, joint forces and moments do not seem to withstand a visual quality check (their direction and magnitude

are sometimes questionable, which falls within the error margin of inertial properties). This might be an interesting, publishable observation, or a “bug”; ultimately an error in the computational pipeline cannot be excluded. Worth noting is that others are currently working in similar directions to open up software alternatives to the conventional XROMM workflow (Falkingham, 2023). At any rate, the primary issues remain the workload of the low through-put XROMM procedure, in contrast to the high through-put demand of appropriate statistics, given the variability of locomotion, coupled with the inaccuracy in determination of inertial segment properties.

The few (three, non-steady) strides which I analyzed and visualized with XROMM techniques with considerable efforts are not representative and cannot provide falsifying evidence with regard to birth-weight dependent differences in piglet locomotor dynamics.

This of course does not imply that there can be no studies of few, representative behavioral observations. It must be emphasized once more that this major limitation of the present study is tied to two things: the fact that subjects are newborns, and the research question posed herein. The high variability in locomotor behavior is a specific feature of the “newborn piglet” model; at later ages, locomotor behavior stabilizes and becomes more recurrent. Our research question was whether one can falsify diagnostic differences in LBW and NBW piglets, and the answer must account for stochastic variation to exclude co-incidental findings. In many other situations, there is either no reason to assume that measurement uncertainty and variability in a behavior are considerable, or they would not prohibit general mechanistic conclusions (e.g. Aerts, 1998; Astley and Roberts, 2012; Montuelle and Williams, 2015). Single or rare observations of a behavior can be instructive in many regards (e.g. Dawson et al., 2011; Druelle et al., 2020; Scheidt et al., 2022; Schwarz et al., 2020). Thus, with most other study settings and subjects, mechanistic evaluations of single strides or actions can generate valuable insights into the workings of the vertebrate musculoskeletal system.

To summarize: this project provided evidence that there is little difference in *how* LBW and NBW piglets move. Though differences in *why* segments move in a particular way are unlikely in the light of this first finding, methodological limitations prohibited definitive conclusions on dynamic differences in the study groups.

8.6 Piglets, Baboons, and Humans

An important premise of this project was that piglets are a valid model species for humans, in terms of locomotion. This premise demands constant revision.

Piglets have long been suggested as a model species for human newborns (Book and Bustad, 1974; Cooper, 1975; Mayerl et al., 2023b), though there might be alternatives (Mellor and Cockburn, 1986). This suggestion is perfectly justified in the fact that it can be ethically problematic to study human newborns, especially on individuals burdened with an impeded

or altered development trajectory. Piglets are social, sedentary, and share similarity in their appearance (hairless, rotund habitus). Of course, one has to carefully discuss differences and similarities with the model species, and evaluate alternatives.

One aspect which sets piglets apart from other mammalian model organisms is the relatively high frequency of occurrence of IUGR, Intra-Uterine Growth Retardation (Van Ginneken et al., 2022; Widdowson, 1971; Wu et al., 2006). This diagnosis is different from low birth weight (Wootton et al., 1983), i.e. the lowest quantile of weights in a litter. Then there is also preterm birth: preterm individuals can have a weight which is lower than normal, but appropriate for gestational age. Preterms can have low weight for gestational age, or even IUGR, thus double or triple the issues. Finally, there is low vitality (Vanden Hole et al., 2018a), which is a momentary performance measure (locomotion/respiration scoring), likely linked to the other three state variables. The situation is complex, classification non-trivial, and there is persistent debate on which indicators matter in which situation.

Nevertheless, researchers study newborn piglets, tracing back epidemiological or congenital issues to histological or structural anomalies. These can be neurological studies: piglets are somewhat similar to humans in the way their cerebral cortex develops (e.g. Lind et al., 2007), they are available for gene editing (Lind et al., 2007), and a lot of research has focused on neural development in piglets and implications for human infants (Conrad et al., 2012; Dickerson et al., 1971; Fanous et al., 2020; Leyshon et al., 2016; Mudd and Dilger, 2017; Radlowski et al., 2014). The piglet model has enabled progress with regard to the research of respiratory problems (Spengler et al., 2019; Williams and Galvis, 1974), metabolic issues (Mellor and Cockburn, 1986; Mota-Rojas et al., 2011; Vanden Hole et al., 2019), mastication and gut (un-)health (Che et al., 2010; Cilieborg et al., 2011; D’Inca et al., 2011; Mayerl et al., 2023a; Sangild et al., 2006; Vanden Hole et al., 2021).

Maybe most importantly for this thesis, muscle and bone histology and function have been studied (Aerts et al., 2023b; Alvarenga et al., 2013; Andersen et al., 2016; Magrini et al., 2023; Vanden Hole et al., 2018b). Results vary, indicating differences in locomotor performance and muscle mass, but failing to associate them with differences in fibre composition or force generating capacity.

Yet, common to all the organ systems under investigation, it must be stated that analogies of human infants and piglets are usually related to homogeneities on the cellular- or tissue level. Whereas developmental anomalies in gut, lung, and brain tissue can be related to dysfunctions in neonates of both species, characteristic anomalies of muscle tissue are lacking. Therefore, though it seems tempting to turn to piglets in light of all the prior research, it might be questionable to choose them as a model *for musculoskeletal aspects*, or even for behavioral aspects which rely on that system.

Other characteristics of the piglet prohibit transfer of findings to our own species, with regard to locomotion. They are rather precocial (Wischner et al., 2009; Young et al., 2023a), and locomotion of newborns matures at a baffling pace (Vanden Hole et al., 2018a, 2017). They

are unguligrade, quadrupedal, strictly non-arboreal, and it must be assumed that during the approximately 94 million years of evolution that separate us (Kumar et al., 2017) their neural and musculoskeletal anatomy and morphology adapted to different locomotor constraints. To conclude, I must re-emphasize that piglets are an appropriate and inevitable model species with regard to anomalies on various organ systems. However, there is little reason to expect relevant conclusions about human locomotor development from piglet research: our locomotor systems are fundamentally different. Additionally, I quantified a lack of differences in locomotion between low- and normal birth weight piglets (Ch. 5): it must be assumed that motor control in precocial animals is innate and therefore locomotor capacity indistinguishable. I cannot exclude that the most severe IUGR or preterm conditions can elicit deficits in kinematics, tracing to impeded neural or musculoskeletal development. Also, metabolic difference are likely, limiting endurance and vitality of LBWs, which I did not quantify. But even if such differences occur, they cannot be easily extrapolated to humans. This opens the discussion about alternative model species. Non-human primates, e.g. baboons, are sufficiently similar to humans in anatomy and habitus, and their usefulness in medical and evolutionary research is widely accepted (Aerts et al., 2023a; Boulinguez-Ambroise et al., 2021; Druelle et al., 2021; Liang et al., 2023; Nardone et al., 2017). Their developmental trajectory is less steep than that of piglets (Ch. 4), but still on a manageable timescale for longitudinal studies (Druelle et al., 2017). Ethical concerns in using primates might weigh more strongly (for reasons unknown). Those objections can be partially mitigated for the kind of locomotor research demonstrated herein: kinematic studies from calibrated videography are non-invasive. The inaccuracies of inertial measurements (Ch. 7) justify doubt on the common practice of euthanizing individuals after experiments; one could as well use naturally deceased specimens, model data, or literature data to transfer the measured inertials to new observations (which would also enable longitudinal inverse dynamic studies). All this of course demands ethical assessment of, amongst more, proportionate radiation dose (especially on young individuals), appropriate experimental circumstances (e.g. habituation requirement), and general welfare (Young and Shapiro, 2018). The main limiting factor for such studies is the availability of non-human primates and their offspring, even more of IUGR-like conditions, which, as is often argued, is one of the benefits of using domestic piglets as a model. Yet in the light of the present thesis, it seems more worthwhile to strive for appropriate conditions and experimental circumstances to study species which are more closely related to humans.

8.7 Conclusion

Independent of the choice of model species, this thesis presented a number of advances in developmental or general studies of locomotor behavior.

Conventional studies of kinematics mainly rely on irreversibly derived spatiotemporal variables.

Raw kinematic traces are often qualitatively compared, yet inaccessible for multivariate statistical analysis. Fourier Analysis methods are known and rarely used, yet a subtle completion made them more accessible for kinematic analysis: equation (3.6), which determines the main phase of a signal. Together with the delay theorem, a cyclic joint angle profile can be “rolled” around the stride cycle for precise temporal alignment. The affine components (mean, amplitude, phase) of the joint angle profiles can be isolated, allowing a Procrustes-like alignment and more relevant quantitative comparison of temporal patterns in joint angle movement.

I have demonstrated the potential of this set of mathematical tricks on a broad data set of ungulates (comparative/descriptive) and on a longitudinal, developmental data set from piglet kinematics (comparative/predictive).

Transforming kinematic data to the frequency domain and extracting affine components opens up countless possibilities for data analysis.

Variability in locomotor output might be considered a nuisance, yet especially in neonates, variability and the progressive reduction of it are arguably a feature of the system. Variability enables learning and locomotor maturation; even seemingly innate neuromotor control has to pass the “reality test”. Being able to quantify and compare this feature is a huge advantage of probabilistic models. Predictive modeling, in particular, can be implemented to compare groups (as demonstrated), refine model design, support hypothesis building; models can take an evolutionary, comparative, longitudinal or cross-sectional, or exploratory flavor, or many more. The modeling toolbox chosen herein has proven to be sufficiently adaptable to handle even complex model constructs, which relate subject characteristics (including morphometrics), spatiotemporal variables, and raw kinematics in FCAS form. The achievement of the author with regard to probabilistic modeling is mainly educational: these methods are popular in other fields, yet I daresay that few biomechanists are fully aware of their potential. Consequently, all code produced for this thesis is publicly available for others to adapt and replicate.

With all these data types and modeling capabilities at hand, future researchers can fine tune their models to answer highly specific questions about study subjects and their locomotor development.

In contrast to kinematic data analysis and modeling, my contribution to inverse dynamic modeling is less than what I hoped to achieve.

As discussed above, this is due to the discrepancy in data throughput of the inverse dynamic workflow, and variability of the phenomenon of interest. Also, initial technical misconceptions about CT scans (Ch. 7) have temporarily misled my efforts; there were visible problems in the outcome of my custom-made inverse dynamics workflow. In reaction, the “flying femur” experiment as a reductive approach enabled the understanding of all details of the workflow. This included a sensitivity analysis, which reveals surprisingly high error margins for the inertial properties, which might be prohibitive for inverse dynamic calculations. I take this as a clear reminder that for anything we measure, we should put all necessary efforts into understanding the mechanistic fundamentals of our measurement, and we must keep track of measurement uncertainty.

Both the reductive approach and the sensitivity analysis are advisable strategies hopefully adopted by future biomechanists.

It is my sincere hope that other researchers take these analysis tools, reproduce, adapt, and correct them where needed, and improve our understanding of how and why various animals move in different phases and circumstances of their existence. I hope this goal explains some extensive, tutorial-like chapters within this thesis, and it hopefully excuses the colloquial tone of some paragraphs which might not have been to the intended entertainment or taste of every recipient.

Thank you for your interest in my work!

Part V

Appendix

9.1 Bibliography

- Abell, C. E., Johnson, A. K., Karriker, L. A., Rothschild, M. F., Hoff, S. J., Sun, G., Fitzgerald, R., and Stalder, K. J. (2014). Using classification trees to detect induced sow lameness with a transient model. *Animal*, 8(6):1000–1009. doi: 10.1017/S1751731114000871 .
- Adams, D. C. (2014). A generalized K statistic for estimating phylogenetic signal from shape and other high-dimensional multivariate data. *Systematic Biology*, 63(5):685–697. doi: 10.1093/sysbio/syu030 .
- Aerts, D. and Aerts, S. (1995). Applications of quantum statistics in psychological studies of decision processes. *Foundations of Science*, 1:85–97. doi: 10.1007/BF00208726 .
- Aerts, P. (1998). Vertical jumping in *Galago senegalensis*: the quest for an obligate mechanical power amplifier. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 353(1375):1607–1620. doi: 10.1098/rstb.1998.0313 .
- Aerts, P., Goyens, J., Berillon, G., D’Août, K., and Druelle, F. (2023a). From quadrupedal to bipedal walking ‘on the fly’: the mechanics of dynamical mode transition in primates. *Journal of Experimental Biology*, 226(2):jeb244792. doi: 10.1242/jeb.244792 .
- Aerts, P., Mielke, F., Vanden Hole, C., Van Gorp, M. J., and Van Ginneken, C. (2023b). Early development of locomotion in the term piglet model: does size matter? *Integrative and Comparative Biology*, page icad054. doi: 10.1093/icb/icad054 .
- Aerts, P., Van Damme, R., Van Elsacker, L., and Duchêne, V. (2000). Spatio-temporal gait characteristics of the hind-limb cycles during voluntary bipedal and quadrupedal walking in bonobos (*Pan paniscus*). *American Journal of Physical Anthropology*, 111(4):503–517. doi: 10.1002/(SICI)1096-8644(200004)111:4<503::AID-AJPA6>3.0.CO;2-J .
- Alexander, R. and Jayes, A. (1980). Fourier analysis of forces exerted in walking and running. *Journal of Biomechanics*, 13(4):383–390. doi: 10.1016/0021-9290(80)90019-6 .
- Alexander, R. M. (1984). Elastic energy stores in running vertebrates. *Integrative and Comparative Biology*, 24(1):85–94. doi: 10.1093/icb/24.1.85 .
- Alexander, R. M. (1989). Optimization and gaits in the locomotion of vertebrates. *Physiological Reviews*, 69(4):1199–1227. doi: 10.1152/physrev.1989.69.4.1199 .
- Alexander, R. M. and Jayes, A. S. (1983). A dynamic similarity hypothesis for the gaits of quadrupedal mammals. *Journal of Zoology*, 201(1):135–152. doi: 10.1111/j.1469-7998.1983.tb04266.x .
- Allen, V., Paxton, H., and Hutchinson, J. R. (2009). Variation in center of mass estimates for extant sauropsids and its importance for reconstructing inertial properties of extinct archosaurs. *The Anatomical Record*, 292(9):1442–1461. doi: 10.1002/ar.20973 .
- Alvarenga, A., Chiarini-Garcia, H., Cardeal, P., Moreira, L., Foxcroft, G., Fontes, D., and Almeida, F. (2013). Intra-uterine growth retardation affects birthweight and postnatal development in pigs, impairing muscle accretion, duodenal mucosa morphology and carcass traits. *Reproduction, Fertility and Development*, 25(2):387–395. doi: 10.1071/RD12021 .

-
- Amdi, C., Krogh, U., Flummer, C., Oksbjerg, N., Hansen, C. F., and Theil, P. K. (2013). Intrauterine growth restricted piglets defined by their head shape ingest insufficient amounts of colostrum. *Journal of Animal Science*, 91(12):5605–5613. doi: 10.2527/jas.2013-6824 .
- Andersen, A. D., Sangild, P. T., Munch, S. L., van der Beek, E. M., Renes, I. B., Ginneken, C. v., Greisen, G. O., and Thymann, T. (2016). Delayed growth, motor function and learning in preterm pigs during early postnatal life. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 310(6):R481–R492. doi: 10.1152/ajpregu.00349.2015 .
- Andersen, A. H. and Kak, A. C. (1984). Simultaneous Algebraic Reconstruction Technique (SART): A Superior Implementation of the Art Algorithm. *Ultrasonic Imaging*, 6(1):81–94. doi: 10.1177/016173468400600107 .
- Animal Diversity Web (2019). [diverse animal profiles]. The University of Michigan, Museum of Zoology, <http://animaldiversity.org>. Accessed 2019-04-05.
- Arbter, K., Snyder, W. E., Burkhardt, H., and Hirzinger, G. (1990). Application of affine-invariant Fourier descriptors to recognition of 3-D objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):640–647. doi: 10.1109/34.56206 .
- Arnold, S. J. (1983). Morphology, performance and fitness. *Integrative and Comparative Biology*, 23(2):347–361. doi: 10.1093/icb/23.2.347 .
- Arroyave-Tobon, S., Drapin, J., Kaniewski, A., Linares, J.-M., and Moretto, P. (2022). Kinematic modeling at the ant scale: propagation of model parameter uncertainties. *Frontiers in Bioengineering and Biotechnology*, 10:767914. doi: 10.3389/fbioe.2022.767914 .
- Astley, H. C. and Roberts, T. J. (2012). Evidence for a vertebrate catapult: elastic energy storage in the plantaris tendon during frog jumping. *Biology Letters*, 8(3):386–389. doi: 10.1098/rsbl.2011.0982 .
- Audigié, F., Pourcelot, P., Degueurce, C., Denoix, J. M., and Geiger, D. (1999). Kinematics of the equine back: flexion-extension movements in sound trotting horses. *Equine Veterinary Journal*, 31(S30):210–213. doi: 10.1111/j.2042-3306.1999.tb05219.x .
- Austin, G. P., Garrett, G. E., and Bohannon, R. W. (1999). Kinematic analysis of obstacle clearance during locomotion. *Gait & Posture*, 10(2):109–120. doi: 10.1016/S0966-6362(99)00022-3 .
- Austin, P. C. and Steyerberg, E. W. (2015). The number of subjects per variable required in linear regression analyses. *Journal of Clinical Epidemiology*, 68(6):627–636. doi: 10.1016/j.jclinepi.2014.12.014 .
- Ayuso, M., Irwin, R., Walsh, C., Van Cruchten, S., and Van Ginneken, C. (2021). Low birth weight female piglets show altered intestinal development, gene expression, and epigenetic changes at key developmental loci. *The FASEB Journal*, 35(4):e21522. doi: 10.1096/fj.202002587R .

-
- Baker, M. J. (1999). Euclidean Space: Maths/Quaternions. EuclideanSpace.com
url: <https://www.euclideanspace.com/maths/algebra/realNormedAlgebra/quaternions>. accessed 2023-06-15, <https://www.euclideanspace.com/physics/kinematics/angularvelocity/QuaternionDifferentiation2.pdf>.
- Ball, R. S. (1876). The theory of screws: A study in the dynamics of a rigid body. *Mathematische Annalen*, 9(4):541–553. doi: 10.1007/BF01442479 .
- Barrett, R., Vonk Noordegraaf, M., and Morrison, S. (2008). Gender differences in the variability of lower extremity kinematics during treadmill locomotion. *Journal of Motor Behavior*, 40(1):62–70. doi: 10.3200/JMBR.40.1.62-70 .
- Basu, C., Wilson, A. M., and Hutchinson, J. R. (2019). The locomotor kinematics and ground reaction forces of walking giraffes. *Journal of Experimental Biology*, 222(2):jeb159277. doi: 10.1242/jeb.159277 .
- Batenburg, K. J. and Sijbers, J. (2011). DART: A Practical Reconstruction Algorithm for Discrete Tomography. *IEEE Transactions on Image Processing*, 20(9):2542–2553. doi: 10.1109/TIP.2011.2131661 .
- Baxter, E., Jarvis, S., D’Eath, R., Ross, D., Robson, S., Farish, M., Nevison, I., Lawrence, A., and Edwards, S. (2008). Investigating the behavioural and physiological indicators of neonatal survival in pigs. *Theriogenology*, 69(6):773–783. doi: 10.1016/j.theriogenology.2007.12.007 .
- Baxter, E. M., Jarvis, S., Palarea-Albaladejo, J., and Edwards, S. A. (2012). The Weaker Sex? The Propensity for Male-Biased Piglet Mortality. *PLOS ONE*, 7(1):1–9. doi: 10.1371/journal.pone.0030318 .
- Beattie, V. E., O’Connell, N. E., Kilpatrick, D. J., and Moss, B. W. (2000). Influence of environmental enrichment on welfare-related behavioural and physiological parameters in growing pigs. *Animal Science*, 70(3):443–450. doi: 10.1017/S1357729800051791 .
- Beckmann, E. C. (2006). CT scanning the early days. *The British Journal of Radiology*, 79(937):5–8. doi: 10.1259/bjr/29444122 .
- Benasson, I., Wagnac, E., Diotalevi, L., Moore, D., Mac-Thiong, J.-M., and Petit, Y. (2020). Gait analysis of a post induced traumatic spinal cord injury porcine model. In *2020 42nd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, pages 3803–3806.
- Berger, M., Yang, Q., and Maier, A. (2018a). X-ray Imaging. In *Medical Imaging Systems: An Introductory Guide [Internet]*. Springer. doi: 10.1007/978-3-319-96520-8_7 .
- Berger, M., Yang, Q., and Maier, A. (2018b). X-ray Imaging. Figure 7.8 “X-ray spectrum of a tungsten tube”. *Medical Imaging Systems: An Introductory Guide [Internet]*. doi: 10.1007/978-3-319-96520-8_7 . Available from: <https://www.ncbi.nlm.nih.gov/books/NBK546155/figure/ch7.fig8>.
- Bernstein, N. A. (1927a). Analyse aperiodischer trigonometrischer Reihen. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, 7(6):476–485. doi: 10.1002/ZAMM.19270070606 .
-

-
- Bernstein, N. A. (1927b). Kymozyklographion, ein neuer Apparat für Bewegungsstudium. *Pflüger's Archiv für die gesamte Physiologie des Menschen und der Tiere*, 217(1):782–792. doi: 10.1007/BF01723725 .
- Bernstein, N. A. (1934). The techniques of the study of movements. In Whiting, H. T. A., editor, *Human Motor Actions: Bernstein Reassessed (1984)*, volume 17 of *Advances in Psychology*, chapter 1, pages 1–26. North-Holland. doi: 10.1016/S0166-4115(08)61367-9 .
- Bernstein, N. A. (1935). The Problem of the Interrelation of Co-Ordination and Localization. In Whiting, H., editor, *Human Motor Actions: Bernstein Reassessed (1984)*, volume 17 of *Advances in Psychology*, chapter 2, pages 77–119. North-Holland. doi: 10.1016/S0166-4115(08)61370-9 .
- Biancardi, C. M. and Minetti, A. E. (2012). Biomechanical determinants of transverse and rotary gallop in cursorial mammals. *Journal of Experimental Biology*, 215(23):4144. doi: 10.1242/jeb.073031 .
- Böl, M., Leichsenring, K., Weichert, C., Sturmat, M., Schenk, P., Blickhan, R., and Siebert, T. (2013). Three-dimensional surface geometries of the rabbit soleus muscle during contraction: input for biomechanical modelling and its validation. *Biomechanics and modeling in mechanobiology*, 12:1205–1220. doi: 10.1007/s10237-013-0476-1 .
- Bongaardt, R. and Meijer, O. G. (2000). Bernstein's Theory of Movement Behavior: Historical Development and Contemporary Relevance. *Journal of Motor Behavior*, 32(1):57–71. doi: 10.1080/00222890009601360 .
- Book, S. A. and Bustad, L. K. (1974). The Fetal and Neonatal Pig in Biomedical Research. *Journal of Animal Science*, 38(5):997–1002. doi: 10.2527/jas1974.385997x .
- Bookstein, F. L. (1991). *Morphometric tools for landmark data: geometry and biology*. Cambridge University Press.
- Borghese, N. A., Bianchi, L., and Lacquaniti, F. (1996). Kinematic determinants of human locomotion. *The Journal of Physiology*, 494(3):863–879. doi: 10.1113/jphysiol.1996.sp021539 .
- Boulinguez-Ambroise, G., Herrel, A., Berillon, G., Young, J. W., Cornette, R., Meguerditchian, A., Cazeau, C., Bellaiche, L., and Pouydebat, E. (2021). Increased performance in juvenile baboons is consistent with ontogenetic changes in morphology. *American Journal of Physical Anthropology*, 175(3):546–558. doi: 10.1002/ajpa.24235 .
- Bracewell, R. N. (2000). *The Fourier Transform And Its Applications*. McGraw-Hill series in electrical and computer engineering. Circuits and systems. McGraw Hill, 3 edition.
- Brainerd, E. L., Baier, D. B., Gatesy, S. M., Hedrick, T. L., Metzger, K. A., Gilbert, S. L., and Crisco, J. J. (2010). X-ray reconstruction of moving morphology (XROMM): precision, accuracy and applications in comparative biomechanics research. *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*, 313A(5):262–279. doi: 10.1002/jez.589 .
- Braune, W. and Fischer, O. (1895-1904). *Der Gang des Menschen*. B.G. Teubner, Leipzig.
-

-
- Brett, M., Penny, W., and Kiebel, S. (2003). Introduction to random field theory. *Human brain function*, 2:867–879. <https://www.fil.ion.ucl.ac.uk/spm/doc/books/hbf2/pdfs/Ch14.pdf>.
- Brown, R. A., Lauzon, M. L., and Frayne, R. (2010). A General Description of Linear Time-Frequency Transforms and Formulation of a Fast, Invertible Transform That Samples the Continuous S-Transform Spectrum Nonredundantly. *IEEE Transactions on Signal Processing*, 58(1):281–290. doi: 10.1109/TSP.2009.2028972 .
- Broyden, C. G. (1970). The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations. *IMA Journal of Applied Mathematics*, 6(1):76–90. doi: 10.1093/imamat/6.1.76 .
- Brugman, H. and Russel, A. (2004). Annotating multimedia/multi-modal resources with ELAN. In *Proceedings of LREC 2004, Fourth International Conference on Language Resources and Evaluation*. Max Planck Institute for Psycholinguistics, The Language Archive, Nijmegen, The Netherlands. <https://tla.mpi.nl/tools/tla-tools/elan>.
- Bruton, M. R., O’Dwyer, N., and Adams, R. (2013). Sex differences in the kinematics and neuromuscular control of landing: Biological, environmental and sociocultural factors. *Journal of Electromyography and Kinesiology*, 23(4):747–758. doi: 10.1016/j.jelekin.2013.04.012 .
- Busemeyer, J. R. and Wang, Z. (2015). What Is Quantum Cognition, and How Is It Applied to Psychology? *Current Directions in Psychological Science*, 24(3):163–169. doi: 10.1177/0963721414568663 .
- Buytaert, J., Goyens, J., De Greef, D., Aerts, P., and Dirckx, J. (2014). Volume Shrinkage of Bone, Brain and Muscle Tissue in Sample Preparation for Micro-CT and Light Sheet Fluorescence Microscopy (LSFM). *Microscopy and Microanalysis*, 20(4):1208–1217. doi: 10.1017/S1431927614001329 .
- Buzug, T. M. (2008). *Computed tomography: From photon statistics to modern cone-beam CT*. Springer.
- Caro, T. M., Graham, C. M., Stoner, C. J., and Vargas, J. K. (2004). Adaptive significance of antipredator behaviour in artiodactyls. *Animal Behaviour*, 67(2):205–228. doi: 10.1016/j.anbehav.2002.12.007 .
- Cazenave, M. and Kivell, T. L. (2023). Challenges and perspectives on functional interpretations of australopith postcrania and the reconstruction of hominin locomotion. *Journal of Human Evolution*, 175:103304. doi: 10.1016/j.jhevol.2022.103304 .
- Challis, J. H. (2020). Quaternions as a solution to determining the angular kinematics of human movement. *BMC Biomedical Engineering*, 2(1):5. doi: 10.1186/s42490-020-00039-z .
- Chamberlin, T. C. (1890). The Method of Multiple Working Hypotheses. *Science*, ns-15(366):92–96. doi: 10.1126/science.ns-15.366.92 .
- Che, L., Thymann, T., Bering, S. B., Huërou-Luron, L., D’inca, R., Zhang, K., Sangild, P. T., et al. (2010). IUGR does not predispose to necrotizing enterocolitis or compromise

-
- postnatal intestinal adaptation in preterm pigs. *Pediatric research*, 67(1):54–59. doi: 10.1203/PDR.0b013e3181c1b15e .
- Cheu, A. Y., Reed, S. A., Mann, S. D., and Bergmann, P. J. (2022). Performance and Kinematic Differences Between Terrestrial and Aquatic Running in *Anolis Sagrei*. *Integrative and Comparative Biology*, 62(4):840–851. doi: 10.1093/icb/icac031 .
- Christiansen, P. (2002). Locomotion in terrestrial mammals: the influence of body mass, limb length and bone proportions on speed. *Zoological Journal of the Linnean Society*, 136(4):685–714. doi: 10.1046/j.1096-3642.2002.00041.x .
- Cilieborg, M. S., Boye, M., Mølbak, L., Thymann, T., and Sangild, P. T. (2011). Preterm birth and necrotizing enterocolitis alter gut colonization in pigs. *Pediatric research*, 69(1):10–16. doi: 10.1203/PDR.0b013e3181ff2a89 .
- Conrad, M. S., Dilger, R. N., and Johnson, R. W. (2012). Brain Growth of the Domestic Pig (*Sus scrofa*) from 2 to 24 Weeks of Age: A Longitudinal MRI Study. *Developmental Neuroscience*, 34(4):291–298. doi: 10.1159/000339311 .
- Cooper, J. E. (1975). The use of the pig as an animal model to study problems associated with low birthweight. *Laboratory Animals*, 9(4):329–336. doi: 10.1258/002367775780957188 .
- Corcoran, A. J., Schirmacher, M. R., Black, E., and Hedrick, T. L. (2021). ThruTracker: Open-Source Software for 2-D and 3-D Animal Video Tracking. *bioRxiv*. doi: 10.1101/2021.05.12.443854 .
- Crall, J. D., Gravish, N., Mountcastle, A. M., and Combes, S. A. (2015). BEEtag: A Low-Cost, Image-Based Tracking System for the Study of Animal Behavior and Locomotion. *PLOS ONE*, 10(9):1–13. doi: 10.1371/journal.pone.0136487 .
- Cronin, N. J. (2021). Using deep neural networks for kinematic analysis: Challenges and opportunities. *Journal of Biomechanics*, 123:110460. doi: 10.1016/j.jbiomech.2021.110460 .
- Daley, M. A. (2018). Understanding the Agility of Running Birds: Sensorimotor and Mechanical Factors in Avian Bipedal Locomotion. *Integrative and Comparative Biology*, 58(5):884–893. doi: 10.1093/icb/icy058 .
- Dallal, G. E. (2003). Why P=0.05? Tufts University
url: https://www.webpages.uidaho.edu/~brian/why_significance_is_five_percent.pdf. accessed 2023-07-25.
- Dalvi, S. (2017). Moment of inertia by Walter Lewin. via youtube.com
url: <https://www.youtube.com/watch?v=lvfzdibrUFA>. accessed 2023-07-11.
- D’Août, K., Aerts, P., De Clercq, D., De Meester, K., and Van Elsacker, L. (2002). Segment and joint angles of hind limb during bipedal and quadrupedal walking of the bonobo (*Pan paniscus*). *American Journal of Physical Anthropology*, 119(1):37–51. doi: 10.1002/ajpa.10112 .
- Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection*. Murray, London.
-

-
- Dawson, M. M., Metzger, K. A., Baier, D. B., and Brainerd, E. L. (2011). Kinematics of the quadrate bone during feeding in mallard ducks. *Journal of Experimental Biology*, 214(12):2036–2046. doi: 10.1242/jeb.047159 .
- Day, L. M. and Jayne, B. C. (2007). Interspecific scaling of the morphology and posture of the limbs during the locomotion of cats (Felidae). *Journal of Experimental Biology*, 210(4):642. doi: 10.1242/jeb.02703 .
- De Groote, F. and Falisse, A. (2021). Perspective on musculoskeletal modelling and predictive simulations of human movement to assess the neuromechanics of gait. *Proceedings of the Royal Society B: Biological Sciences*, 288(1946):20202432. doi: 10.1098/rspb.2020.2432 .
- D’Eath, R. (2012). Repeated locomotion scoring of a sow herd to measure lameness: consistency over time, the effect of sow characteristics and inter-observer reliability. *Animal Welfare*, 21(2):219–231. doi: 10.7120/09627286.21.2.219 .
- Demuth, O. E., Herbst, E., Polet, D. T., Wiseman, A. L. A., and Hutchinson, J. R. (2023). Modern three-dimensional digital methods for studying locomotor biomechanics in tetrapods. *Journal of Experimental Biology*, 226(Suppl1):jeb245132. doi: 10.1242/jeb.245132 .
- Denayer, T., Stöhr, T., and Van Roy, M. (2014). Animal models in translational medicine: Validation and prediction. *New Horizons in Translational Medicine*, 2(1):5–11. doi: 10.1016/j.nhtm.2014.08.001 .
- DiBerardino III, L. A., Polk, J. D., Rosengren, K. S., Spencer-Smith, J. B., and Hsiao-Wecksler, E. T. (2010). Quantifying complexity and variability in phase portraits of gait. *Clinical Biomechanics*, 25(6):552–556. doi: 10.1016/j.clinbiomech.2010.03.007 .
- Dickerson, J., Merat, A., Widdowson, E., et al. (1971). Intra-uterine growth retardation in the pig. 3. The chemical structure of the brain. *Biology of the Neonate*, 19(4/6):354–362. doi: 10.1159/000240429 .
- Downey, A. (2013). Are your data normal? Hint: no. Probably Overthinking It (blog) url: <https://allendowney.blogspot.com/2013/08/are-my-data-normal.html>. accessed 2023-07-25.
- Druelle, F., Aerts, P., D’Août, K., Moulin, V., and Berillon, G. (2017). Segmental morphometrics of the olive baboon (*papio anubis*): a longitudinal study from birth to adulthood. *Journal of Anatomy*, 230(6):805–819. doi: 10.1111/joa.12602 .
- Druelle, F., Aerts, P., Ngawolo, J. C. B., and Narat, V. (2020). Impressive arboreal gap-crossing behaviors in wild bonobos, *Pan paniscus*. *International Journal of Primatology*, 41:129–140. doi: 10.1007/s10764-020-00140-z .
- Druelle, F., Goyens, J., Vasilopoulou-Kampitsi, M., and Aerts, P. (2019). Compliant legs enable lizards to maintain high running speeds on complex terrains. *Journal of Experimental Biology*, 222(6):jeb195511. doi: 10.1242/jeb.195511 .
- Druelle, F., Özçelebi, J., Marchal, F., and Berillon, G. (2021). Development of bipedal walking in olive baboons, *Papio anubis*: A kinematic analysis. *American Journal of Biological Anthropology*. doi: 10.1002/ajpa.24454 .
-

-
- Dryden, I. L. and Mardia, K. V. (2016). Statistical shape analysis with applications in R. In *Wiley Series In Probability And Statistics*. John Wiley and Sons Ltd, 2nd edition.
- Du Plessis, A., Meincken, M., and Seifert, T. (2013). Quantitative determination of density and mass of polymeric materials using microfocus computed tomography. *Journal of Nondestructive Evaluation*, 32:413–417. doi: 10.1007/s10921-013-0195-7 .
- Dumas, R. (1999). 3d kinematics and inverse dynamics. MATLAB Central File Exchange url: <https://www.mathworks.com/matlabcentral/fileexchange/58021-3d-kinematics-and-inverse-dynamics>. accessed 2023-06-15.
- Dumas, R., Aissaoui, R., and de Guise, J. (2004). A 3d generic inverse dynamic method using wrench notation and quaternion algebra. *Computer Methods in Biomechanics and Biomedical Engineering*, 7(3):159–166. doi: 10.1080/10255840410001727805 .
- Dumas, R., Nicol, E., and Chèze, L. (2007). Influence of the 3D Inverse Dynamic Method on the Joint Forces and Moments During Gait. *Journal of Biomechanical Engineering*, 129(5):786–790. doi: 10.1115/1.2768114 .
- Dunn, T. W., Marshall, J. D., Severson, K. S., Aldarondo, D. E., Hildebrand, D. G., Chettih, S. N., Wang, W. L., Gellis, A. J., Carlson, D. E., Aronov, D., Freiwald, W. A., Wang, F., and Ölveczky, B. P. (2021). Geometric deep learning enables 3D kinematic profiling across species and environments. *Nature methods*, 18(5):564–573. doi: 10.1038/s41592-021-01106-6 .
- Durston, N. E., Mahadik, Y., and Windsor, S. P. (2022). Quantifying avian inertial properties using calibrated computed tomography. *Journal of Experimental Biology*, 225(1). doi: 10.1242/jeb.242280 .
- D’Inca, R., Gras-Le Guen, C., Che, L., Sangild, P. T., and Le Huërou-Luron, I. (2011). Intrauterine growth restriction delays feeding-induced gut adaptation in term newborn pigs. *Neonatology*, 99(3):208–216. doi: 10.1159/000314919 .
- Edwards, S. and Baxter, E. (2015). Chapter 11. Piglet mortality: causes and prevention. In *The gestating and lactating sow*, pages 253–278.
- Ekhtator, C., Varshney, A., Young, M. W., Tanis, D., Granatosky, M. C., Diaz, R. E., and Molnar, J. L. (2023). Locomotor characteristics of the ground-walking chameleon *Brookesia superciliaris*. *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology*, 339(6):602–614. doi: 10.1002/jez.2703 .
- Falkingham, P. (2023). XROMM Blender Tools. github repository url: https://github.com/pfalkingham/XROMM_BlenderTools. accessed 2023-08-11.
- Fanous, M., Caputo, M. P., Lee, Y. J., Rund, L. A., Best-Popescu, C., Kandel, M. E., Johnson, R. W., Das, T., Kuchan, M. J., and Popescu, G. (2020). Quantifying myelin content in brain tissue using color spatial light interference microscopy (cslim). *PLOS ONE*, 15(11):1–13. doi: 10.1371/journal.pone.0241084 .
- Fath, M. A., Nguyen, S. V., Donahue, J., McMenamin, S. K., and Tytell, E. D. (2023). Static Stability and Swim Bladder Volume in the Bluegill Sunfish (*Lepomis macrochirus*). *Integrative Organismal Biology*, 5(1):obad005. doi: 10.1093/iob/obad005 .

-
- Feldkamp, L. A., Davis, L. C., and Kress, J. W. (1984). Practical cone-beam algorithm. *Journal of the Optical Society of America A*, 1(6):612–619. doi: 10.1364/JOSAA.1.000612 .
- Feldpausch, J. A., Jourquin, J., Bergstrom, J. R., Bargaen, J. L., Bokenkroger, C. D., Davis, D. L., Gonzalez, J. M., Nelssen, J. L., Puls, C. L., Trout, W. E., and Ritter, M. J. (2019). Birth weight threshold for identifying piglets at risk for preweaning mortality. *Translational Animal Science*, 3(2):633–640. doi: 10.1093/tas/txz076 .
- Figueiredo, J., Santos, C. P., and Moreno, J. C. (2018). Automatic recognition of gait patterns in human motor disorders using machine learning: A review. *Medical Engineering & Physics*, 53:1–12. doi: 10.1016/j.medengphy.2017.12.006 .
- Fischer, M. S. and Blickhan, R. (2006). The tri-segmented limbs of therian mammals: kinematics, dynamics, and self-stabilization—a review. *Journal of Experimental Zoology Part A: Comparative Experimental Biology*, 305A(11):935–952. doi: 10.1002/jez.a.333 .
- Fischer, M. S., Schilling, N., Schmidt, M., Haarhaus, D., and Witte, H. (2002). Basic limb kinematics of small therian mammals. *Journal of Experimental Biology*, 205(9):1315. doi: 10.1242/jeb.205.9.1315 .
- Fischer, O. and Braune, C. W. (1895-1904). *Der Gang des Menschen*. BG Teubner.
- Flandin, G. and Friston, K. J. (2008). Statistical parametric mapping (SPM). Scholarpedia url: http://www.scholarpedia.org/article/Statistical_parametric_mapping.
- Flashner, H. and McNitt-Gray, J. L. (2019). 3D Kinematics: Using Quaternions for Modeling Orientation and Rotations in Biomechanics. In Pallis, J. M., McNitt-Gray, J. L., and Hung, G. K., editors, *Biomechanical Principles and Applications in Sports*, pages 155–233. Springer International Publishing, Cham. doi: 10.1007/978-3-030-13467-9_7 .
- Fletcher, R. (1970). A new approach to variable metric algorithms. *The Computer Journal*, 13(3):317–322. doi: 10.1093/comjnl/13.3.317 .
- Fourier, J. (1822). *Theorie analytique de la chaleur, par M. Fourier*. Chez Firmin Didot, père et fils.
- Frantz, L., Meijaard, E., Gongora, J., Haile, J., Groenen, M. A. M., and Larson, G. (2015). The evolution of Suidae. *Annual Review of Animal Biosciences*, 4(1):61–85. doi: 10.1146/annurev-animal-021815-111155 .
- Frenkel, D., Six, N., De Beenhouwer, J., and Sijbers, J. (2022). Tabu-DART: a dynamic update strategy for efficient discrete algebraic reconstruction. *The Visual Computer*, pages 1–13. doi: 10.1007/s00371-022-02616-w .
- Frick, R. W. (1996). The appropriate use of null hypothesis testing. *Psychological Methods*, 1(4):379. doi: 10.1037/1082-989X.1.4.379 .
- Friston, K. J. (2003). *Statistical Parametric Mapping*, pages 237–250. Springer US, Boston, MA.
- Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., and Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2(4):189–210. doi: 10.1002/hbm.460020402 .
-

-
- Fukuchi, C. A., Fukuchi, R. K., and Duarte, M. (2019). Effects of walking speed on gait biomechanics in healthy participants: a systematic review and meta-analysis. *Systematic reviews*, 8(1):1–11. doi: 10.1186/s13643-019-1063-z .
- Full, R. J. and Koditschek, D. E. (1999). Templates and anchors: neuromechanical hypotheses of legged locomotion on land. *Journal of Experimental Biology*, 202(23):3325–3332. doi: 10.1242/jeb.202.23.3325 .
- Fuller, P. O., Higham, T. E., and Clark, A. J. (2011). Posture, speed, and habitat structure: three-dimensional hindlimb kinematics of two species of padless geckos. *Zoology*, 114(2):104–112. doi: 10.1016/j.zool.2010.11.003 .
- Funk, S. M., Verma, S. K., Larson, G., Prasad, K., Singh, L., Narayan, G., and Fa, J. E. (2007). The pygmy hog is a unique genus: 19th century taxonomists got it right first time round. *Molecular Phylogenetics and Evolution*, 45(2):427–436. doi: 10.1016/j.ympev.2007.08.007 .
- Gallistel, C. R. (1982). *The organization of action: A new synthesis*. Psychology Press.
- Ganley, K. J. and Powers, C. M. (2005). Gait kinematics and kinetics of 7-year-old children: a comparison to adults using age-specific anthropometric data. *Gait & Posture*, 21(2):141–145. doi: 10.1016/j.gaitpost.2004.01.007 .
- Gao, F. and Han, L. (2012). Implementing the Nelder-Mead simplex algorithm with adaptive parameters. *Computational Optimization and Applications*, 51(1):259–277. doi: 10.1007/s10589-010-9329-3 .
- Gatesy, S. M. and Pollard, N. S. (2011). Apples, oranges, and angles: Comparative kinematic analysis of disparate limbs. *Journal of Theoretical Biology*, 282(1):7–13. doi: 10.1016/j.jtbi.2011.05.009 .
- Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., and Rubin, D. (2020). *Bayesian Data Analysis*. freely available online, 3rd edition. <http://www.stat.columbia.edu/~gelman/book>.
- Geyer, H., Seyfarth, A., and Blickhan, R. (2006). Compliant leg behaviour explains basic dynamics of walking and running. *Proceedings of the Royal Society B: Biological Sciences*, 273(1603):2861–2867. doi: 10.1098/rspb.2006.3637 .
- Geyer, L. L., Schoepf, U. J., Meinel, F. G., Nance, J. W., Bastarrika, G., Leipsic, J. A., Paul, N. S., Rengo, M., Laghi, A., and De Cecco, C. N. (2015). State of the art: Iterative ct reconstruction techniques. *Radiology*, 276(2):339–357. doi: 10.1148/radiol.2015132766 .
- Gilbert, P. (1972). Iterative methods for the three-dimensional reconstruction of an object from projections. *Journal of Theoretical Biology*, 36(1):105–117. doi: 10.1016/0022-5193(72)90180-4 .
- Girard, I. (2001). Field cost of activity in the kit fox, *Vulpes macrotis*. *Physiological and Biochemical Zoology*, 74(2):191–202. doi: 10.1086/319668 .
- Goldfrab, D. (1970). A Family of Variable Metric Methods Derived by Variations Means. *Mathematics of Computation*, 24:23–26. doi: 10.1090/S0025-5718-1970-0258249-6 .

-
- Gongora, J., Cuddahee, R. E., Nascimento, F. F. d., Palgrave, C. J., Lowden, S., Ho, S. Y. W., Simond, D., Damayanti, C. S., White, D. J., Tay, W. T., Randi, E., Klingel, H., Rodrigues-Zarate, C. J., Allen, K., Moran, C., and Larson, G. (2011). Rethinking the evolution of extant sub-Saharan African suids (Suidae, Artiodactyla). *Zoologica Scripta*, 40(4):327–335. doi: 10.1111/j.1463-6409.2011.00480.x .
- Goswami, A. (1998). A new gait parameterization technique by means of cyclogram moments: Application to human slope walking. *Gait & Posture*, 8(1):15–36. doi: 10.1016/S0966-6362(98)00014-9 .
- Gower, J. C. (1975). Generalized Procrustes analysis. *Psychometrika*, 40(1):33–51. doi: 10.1007/BF02291478 .
- Grandinson, K., Lund, M. S., Rydhmer, L., and Strandberg, E. (2002). Genetic Parameters for the Piglet Mortality Traits Crushing, Stillbirth and Total Mortality, and their Relation to Birth Weight. *Acta Agriculturae Scandinavica, Section A - Animal Science*, 52(4):167–173. doi: 10.1080/090647002762381041 .
- Grasso, R., Zago, M., and Lacquaniti, F. (2000). Interactions Between Posture and Locomotion: Motor Patterns in Humans Walking With Bent Posture Versus Erect Posture. *Journal of Neurophysiology*, 83(1):288–300. doi: 10.1152/jn.2000.83.1.288 .
- Gray, R. M. and Goodman, J. W. (1995). *Fourier Transforms: An Introduction for Engineers*. The Springer International Series in Engineering and Computer Science 322. Springer US, 1 edition.
- Guesgen, M. and Bench, C. (2017). What can kinematics tell us about the affective states of animals? *Animal Welfare*, 26:383–397. doi: 10.7120/09627286.26.4.383 .
- Gupta, D. S. and Chen, L. (2016). Brain oscillations in perception, timing and action. *Current Opinion in Behavioral Sciences*, 8:161–166. doi: 10.1016/j.cobeha.2016.02.021 .
- Hales, J., Moustsen, V. A., Nielsen, M. B. F., and Hansen, C. F. (2013). Individual physical characteristics of neonatal piglets affect preweaning survival of piglets born in a noncrated system. *Journal of Animal Science*, 91(10):4991–5003. doi: 10.2527/jas.2012-5740 .
- Hallemans, A. and Aerts, P. (2009). Effects of visual deprivation on intra-limb coordination during walking in children and adults. *Experimental Brain Research*, 198(1):95. doi: 10.1007/s00221-009-1937-8 .
- Hansen, P. C., Jørgensen, J., and Lionheart, W. R. (2021). *Computed tomography: algorithms, insight, and just enough theory*. SIAM.
- Hansraj, K. K., Hansraj, J. A., Hansraj, M. D. G., Rožič, U., and Choudhary, D. (2022). Belly Fat Forces on the Spine: Finite Element Analysis of Belly Fat Forces by Waist Circumference. *Surgical technology international*, 40:405–410.
- Harrison, D. (2021). Chaos. LibreTexts: Modern Physics, Supplemental Modules url: <https://phys.libretexts.org/@go/page/4113>. accessed 2023-07-17.

-
- Hassanin, A., Delsuc, F., Ropiquet, A., Hammer, C., Jansen van Vuuren, B., Matthee, C., Ruiz-Garcia, M., Catzeffis, F., Areskoug, V., Nguyen, T. T., and Couloux, A. (2012). Pattern and timing of diversification of Cetartiodactyla (Mammalia, Laurasiatheria), as revealed by a comprehensive analysis of mitochondrial genomes. *Comptes Rendus Biologies*, 335(1):32–50. doi: 10.1016/j.crv.2011.11.002 .
- Hatala, K. G., Demes, B., and Richmond, B. G. (2016). Laetoli footprints reveal bipedal gait biomechanics different from those of modern humans and chimpanzees. *Proceedings of the Royal Society B: Biological Sciences*, 283(1836):20160235. doi: 10.1098/rspb.2016.0235 .
- Hedrick, T. L. (2008). Software techniques for two- and three-dimensional kinematic measurements of biological and biomimetic systems. *Bioinspiration & Biomimetics*, 3(3):034001. doi: 10.1088/1748-3182/3/3/034001 .
- Heideman, M., Johnson, D., and Burrus, C. (1984). Gauss and the history of the fast fourier transform. *IEEE ASSP Magazine*, 1(4):14–21. doi: 10.1109/MASSP.1984.1162257 .
- Hermanson, J. W. and MacFadden, B. J. (1992). Evolutionary and functional morphology of the shoulder region and stay-apparatus in fossil and extant horses (Equidae). *Journal of Vertebrate Paleontology*, 12(3):377–386. doi: 10.1080/02724634.1992.10011466 .
- Hicheur, H., Terekhov, A. V., and Berthoz, A. (2006). Intersegmental coordination during human locomotion: Does planar covariation of elevation angles reflect central constraints? *Journal of Neurophysiology*, 96(3):1406–1419. doi: 10.1152/jn.00289.2006 .
- Hildebrand, M. (1989). The Quadrupedal Gaits of Vertebrates. *BioScience*, 39(11):766–775. doi: 10.2307/1311182 .
- Hof, A. L. (1996). Scaling gait data to body size. *Gait & Posture*, 4(3):222–223. doi: 10.1016/0966-6362(95)01057-2 .
- Holmes, P., Full, R., Koditschek, D., and Guckenheimer, J. (2006). The dynamics of legged locomotion: Models, analyses, and challenges. *SIAM Review*, 48(2):207–304. doi: 10.1137/s0036144504445133 .
- Hounsfield, G. N. (1973). Computerized transverse axial scanning (tomography): Part 1. description of system. *The British Journal of Radiology*, 46(552):1016–1022. doi: 10.1259/0007-1285-46-552-1016 .
- Hoyt, D. F. and Taylor, C. R. (1981). Gait and the energetics of locomotion in horses. *Nature*, 292(5820):239–240. doi: 10.1038/292239a0 .
- Hsiao-Wecksler, E. T., Polk, J. D., Rosengren, K. S., Sosnoff, J. J., and Hong, S. (2010). A review of new analytic techniques for quantifying symmetry in locomotion. *Symmetry*, 2(2):1135–1155. doi: 10.3390/sym2021135 .
- Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., and Liu, H. H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1971):903–995. doi: 10.1098/rspa.1998.0193 .

-
- Hubel, T. Y. and Usherwood, J. R. (2015). Children and adults minimise activated muscle volume by selecting gait parameters that balance gross mechanical power and work demands. *Journal of Experimental Biology*, 218(18):2830. doi: 10.1242/jeb.122135 .
- Huffman, B. (2019). Ultimate ungulate guide to the world’s hoofed mammals. www.ultimateungulate.com. Accessed 2019-04-05.
- Hughes, I. and Hase, T. (2010). *Measurements and their uncertainties: a practical guide to modern error analysis*. OUP Oxford.
- Hutchinson, J. R., Ng-Thow-Hing, V., and Anderson, F. C. (2007). A 3D interactive method for estimating body segmental parameters in animals: Application to the turning and running performance of *Tyrannosaurus rex*. *Journal of Theoretical Biology*, 246(4):660–680. doi: 10.1016/j.jtbi.2007.01.023 .
- Hutchinson, J. R., Schwerda, D., Famini, D. J., Dale, R. H. I., Fischer, M. S., and Kram, R. (2006). The locomotor kinematics of Asian and African elephants: changes with speed and size. *Journal of Experimental Biology*, 209(19):3812–3827. doi: 10.1242/jeb.02443 .
- Irschick, D. and Jayne, B. (1999). Comparative three-dimensional kinematics of the hindlimb for high-speed bipedal and quadrupedal locomotion of lizards. *Journal of Experimental Biology*, 202(9):1047–1065. doi: 10.1242/jeb.202.9.1047 .
- Isler, K. (2005). 3D-Kinematics of vertical climbing in hominoids. *American Journal of Physical Anthropology*, 126(1):66–81. doi: 10.1002/ajpa.10419 .
- Ivanenko, Y. P., d’Avella, A., Poppele, R. E., and Lacquaniti, F. (2008). On the origin of planar covariation of elevation angles during human locomotion. *Journal of Neurophysiology*, 99(4):1890–1898. doi: 10.1152/jn.01308.2007 .
- Jackson, B. E., Evangelista, D. J., Ray, D. D., and Hedrick, T. L. (2016). 3D for the people: multi-camera motion capture in the field with consumer-grade cameras and open source software. *Biology Open*, 5(9):1334–1342. doi: 10.1242/bio.018713 .
- Jackson, K. M. (1979). Fitting of Mathematical Functions to Biomechanical Data. *IEEE Transactions on Biomedical Engineering*, BME-26(2):122–124. doi: 10.1109/TBME.1979.326551 .
- Jenkins, F. A. (1971). Limb posture and locomotion in the Virginia opossum (*Didelphis marsupialis*) and in other non-cursorial mammals. *Journal of Zoology*, 165(3):303–315. doi: 10.1111/j.1469-7998.1971.tb02189.x .
- Johnson, C. J., Parker, K. L., Heard, D. C., and Gillingham, M. P. (2002). Movement parameters of ungulates and scale-specific responses to the environment. *Journal of Animal Ecology*, 71(2):225–235. doi: 10.1046/j.1365-2656.2002.00595.x .
- Jones, M. F. and Hasiotis, S. T. (2023). Terrestrial locomotor behaviors of the big brown bat (Vespertilionidae: *Eptesicus fuscus*). *Mammal Research*, 68(2):253–262. doi: 10.1007/s13364-022-00669-9 .
- Kabsch, W. (1976). A solution for the best rotation to relate two sets of vectors. *Acta Crystallographica Section A*, 32(5):922–923. doi: 10.1107/S0567739476001873 .
-

-
- Karashchuk, P., Rupp, K. L., Dickinson, E. S., Walling-Bell, S., Sanders, E., Azim, E., Brunton, B. W., and Tuthill, J. C. (2021). Anipose: A toolkit for robust markerless 3D pose estimation. *Cell Reports*, 36(13):109730. doi: 10.1016/j.celrep.2021.109730 .
- Kashiha, M. A., Bahr, C., Ott, S., Moons, C. P., Niewold, T. A., Tuytens, F., and Berckmans, D. (2014). Automatic monitoring of pig locomotion using image analysis. *Livestock Science*, 159:141–148. doi: 10.1016/j.livsci.2013.11.007 .
- Kemeny, J. G., Snell, J. L., Knapp, A. W., and Griffeath, D. (1976). Introduction to random fields. *Denumerable Markov Chains: with a chapter of Markov Random Fields by David Griffeath*, pages 425–458.
- Kendall, D. G. (1989). A survey of the statistical theory of shape. *Statistical Science*, 4(2):87–99. doi: 10.1214/ss/1177012582 .
- Kenwright, B. (2015). Quaternion Fourier Transform for Character Motions. In Jaillet, F., Zara, F., and Zachmann, G., editors, *Workshop on Virtual Reality Interaction and Physical Simulation*. The Eurographics Association. doi: 10.2312/vrphys.20151328 .
- Kirtley, C., Whittle, M., and Jefferson, R. (1985). Influence of walking speed on gait parameters. *Journal of Biomedical Engineering*, 7(4):282–288. doi: 10.1016/0141-5425(85)90055-X .
- Kneller, G. R. (1991). Superposition of molecular structures using quaternions. *Molecular Simulation*, 7(1-2):113–119. doi: 10.1080/08927029108022453 .
- Knörlein, B. J., Baier, D. B., Gatesy, S. M., Laurence-Chasen, J. D., and Brainerd, E. L. (2016). Validation of XMALab software for marker-based XROMM. *Journal of Experimental Biology*, 219(23):3701–3711. doi: 10.1242/jeb.145383 .
- Korr, A. L. and Hyer, P. (1962). A trifilar pendulum for the determination moments of inertia. Technical report, Frankford Arsenal Philadelphia PA Pitman-Dunn Research Labs.
- Kramer, P. A. and Sylvester, A. D. (2013). Humans, geometric similarity and the Froude number: is "reasonably close" really close enough? *Biology Open*, 2(2):111. doi: 10.1242/bio.20122691 .
- Kuhl, F. P. and Giardina, C. R. (1982). Elliptic Fourier features of a closed contour. *Computer Graphics and Image Processing*, 18(3):236–258. doi: 10.1016/0146-664X(82)90034-X .
- Kumar, S., Stecher, G., Suleski, M., and Hedges, S. B. (2017). TimeTree: A Resource for Timelines, Timetrees, and Divergence Times. *Molecular Biology and Evolution*, 34(7):1812–1819. doi: 10.1093/molbev/msx116 .
- Kurz, M. J., Stuber, W., and DeJong, S. L. (2011). Body weight supported treadmill training improves the regularity of the stepping kinematics in children with Cerebral Palsy. *Developmental Neurorehabilitation*, 14(2):87–93. doi: 10.3109/17518423.2011.552459 .
- Laczo, J. and Latash, M. L. (2016). *Progress in Motor Control: Theories and Translations*, volume 957. Springer.

-
- Lamb, P. F. and Stöckl, M. (2014). On the use of continuous relative phase: Review of current approaches and outline for a new standard. *Clinical Biomechanics*, 29(5):484–493. doi: 10.1016/j.clinbiomech.2014.03.008 .
- Latash, M. L. (1998). *Progress in motor control: Bernstein's traditions in movement studies*, volume 1. Human kinetics.
- Latash, M. L. (2008). *Synergy*. Oxford University Press.
- Latash, M. L. (2010). Motor synergies and the equilibrium-point hypothesis. *Motor Control*, 14(3):294 – 322. doi: 10.1123/mcj.14.3.294 .
- LaVallee, K. T., Maus, T. P., Stock, J. D., Stalder, K. J., Karriker, L. A., Murthy, N. S., Kanwar, R., Beutler, A. S., and Unger, M. D. (2020). Quantitation of Gait and Stance Alterations Due to Monosodium Iodoacetate-induced Knee Osteoarthritis in Yucatan Swine. *Comparative medicine*, 70(3):248–257. doi: 10.30802/AALAS-CM-19-000075 .
- Lawrence, J., Bernal, J., and Witzgall, C. (2019). A purely algebraic justification of the kabsch-umeyama algorithm. *Journal of research of the National Institute of Standards and Technology*, 124:1. doi: 10.6028/jres.124.028 .
- Le Dividich, J., Charneca, R., and Thomas, F. (2017). Relationship between birth order, birth weight, colostrum intake, acquisition of passive immunity and pre-weaning mortality of piglets. *Spanish Journal of Agricultural Research*, 15(2):e0603. doi: 10.5424/sjar/2017152-9921 .
- Lees, J., Gardiner, J., Usherwood, J., and Nudds, R. (2016). Locomotor preferences in terrestrial vertebrates: An online crowdsourcing approach to data collection. *Scientific Reports*, 6:28825. doi: 10.1038/srep28825 .
- Lejeune, T. M., Willems, P. A., and Heglund, N. C. (1998). Mechanics and energetics of human locomotion on sand. *Journal of Experimental Biology*, 201(13):2071. doi: 10.1242/jeb.201.13.2071 .
- Lennie, E., Tsoumpas, C., and Sourbron, S. (2021). Multimodal phantoms for clinical PET/MRI. *EJNMMI physics*, 8(1):1–24. doi: 10.1186/s40658-021-00408-0 .
- Lewandowski, D., Kurowicka, D., and Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, 100(9):1989–2001. doi: 10.1016/j.jmva.2009.04.008 .
- Lewin, W. (1999a). 8.01x - Lect 19 - Rotating Objects, Moment of Inertia, Rotational KE, Neutron Stars. MIT / via youtube.com
url: https://www.youtube.com/watch?v=fDJeVR0o__w. accessed 2023-07-11.
- Lewin, W. (1999b). 8.01x - lect 20 - angular momentum, torques, conservation of angular momentum. MIT / via youtube.com
url: <https://www.youtube.com/watch?v=sNaaL19opxw>. accessed 2023-06-15.
- Lewin, W. (1999c). 8.01x - Lect 24 - Rolling Motion, Gyroscopes, VERY NON-INTUITIVE. MIT / via youtube.com
url: https://youtu.be/XPUuF_dECVI. accessed 2023-11-02.
-

-
- Lewin, W. and Goldstein, W. (2012). *For the Love of Physics: From the End of the Rainbow to the Edge of Time-A Journey through the Wonders of Physics*. Simon and Schuster.
- Leyshon, B. J., Radlowski, E. C., Mudd, A. T., Steelman, A. J., and Johnson, R. W. (2016). Postnatal Iron Deficiency Alters Brain Development in Piglets. *The Journal of Nutrition*, 146(7):1420–1427. doi: 10.3945/jn.115.223636 .
- Liang, F., Yu, S., Pang, S., Wang, X., Jie, J., Gao, F., Song, Z., Li, B., Liao, W.-H., and Yin, M. (2023). Non-human primate models and systems for gait and neurophysiological analysis. *Frontiers in Neuroscience*, 17. doi: 10.3389/fnins.2023.1141567 .
- LibreTexts (2022a). 15.5: Pendulums. LibreTexts: University Physics I - Mechanics, Sound, Oscillations, and Waves
url: <https://phys.libretexts.org/@go/page/4065>. accessed 2023-07-17.
- LibreTexts (2022b). 29. Introduction to Quantum Physics: Probability and The Heisenberg Uncertainty Principle. LibreTexts College Physics 1e
url: <https://phys.libretexts.org/@go/page/2763>. accessed 2023-07-17.
- Lind, N. M., Moustgaard, A., Jelsing, J., Vajta, G., Cumming, P., and Hansen, A. K. (2007). The use of pigs in neuroscience: Modeling brain disorders. *Neuroscience & Biobehavioral Reviews*, 31(5):728–751. doi: 10.1016/j.neubiorev.2007.02.003 .
- Litten, J., Drury, P., Corson, A., Lean, I., and Clarke, L. (2003). The influence of piglet birth weight on physical and behavioural development in early life. *Neonatology*, 84(4):311–318. doi: 10.1159/000073640 .
- Liu, L. P., Hwang, M., Hung, M., Soulen, M. C., Schaer, T. P., Shapira, N., and Noël, P. B. (2023). Non-invasive mass and temperature quantifications with spectral ct. *Scientific Reports*, 13(1):6109. doi: 10.1038/s41598-023-33264-2 .
- Loscher, D. M., Meyer, F., Kracht, K., and Nyakatura, J. A. (2016). Timing of head movements is consistent with energy minimization in walking ungulates. *Proceedings of the Royal Society B: Biological Sciences*, 283(1843):20161908. doi: 10.1098/rspb.2016.1908 .
- Lynch, K. M. and Park, F. C. (2017). *Modern Robotics*. Cambridge University Press.
- Macaulay, S., Hutchinson, J. R., and Bates, K. T. (2017). A quantitative evaluation of physical and digital approaches to centre of mass estimation. *Journal of Anatomy*, 231(5):758–775. doi: 10.1111/joa.12667 .
- MacLaren, J. A. and Nauwelaerts, S. (2016). A three-dimensional morphometric analysis of upper forelimb morphology in the enigmatic tapir (Perissodactyla: Tapirus) hints at subtle variations in locomotor ecology. *Journal of Morphology*, 277(11):1469–1485. doi: 10.1002/jmor.20588 .
- MacLellan, M. J. and McFadyen, B. J. (2010). Segmental control for adaptive locomotor adjustments during obstacle clearance in healthy young adults. *Experimental Brain Research*, 202(2):307–318. doi: 10.1007/s00221-009-2133-6 .

-
- MacLeod, N. (2007). PalaeoMath101: Part 5 - PCA, Eigenanalysis & Regression V. Palaeontology Newsletter Nr. 64
url: <https://www.palass.org/publications/newsletter/palaeomath-101/palaeomath-part-5-pca-eigenanalysis-regression-v>. accessed 2019-04-05.
- Magrini, S. H., Mossor, A. M., German, R. Z., and Young, J. W. (2023). Developmental factors influencing bone strength in precocial mammals: An infant pig model. *Journal of Anatomy*, 243(1):174–181. doi: 10.1111/joa.13848 .
- Marchant, J. N., Rudd, A. R., Mendl, M. T., Broom, D. M., Meredith, M. J., Corning, S., and Simmins, P. H. (2000). Timing and causes of piglet mortality in alternative and conventional farrowing systems. *Veterinary Record*, 147(8):209–214. doi: 10.1136/vr.147.8.209 .
- Marey, E.-J. (1878). *La méthode graphique dans les sciences expérimentales et particulièrement en physiologie et en médecine*. G. Masson.
- Marteniuk, R. G. and Romanow, S. K. (1983). Human Movement Organization and Learning As Revealed by Variability of Movement, use of Kinematic Information, and Fourier Analysis. In Magill, R. A., editor, *Memory and Control of Action*, volume 12 of *Advances in Psychology*, pages 167–197. North-Holland. doi: 10.1016/S0166-4115(08)61991-3 .
- Mathis, A., Mamidanna, P., Cury, K. M., Abe, T., Murthy, V. N., Mathis, M. W., and Bethge, M. (2018). DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nature Neuroscience*, 21(9):1281–1289. doi: 10.1038/s41593-018-0209-y .
- Mathis, A., Schneider, S., Lauer, J., and Mathis, M. W. (2020). A Primer on Motion Capture with Deep Learning: Principles, Pitfalls, and Perspectives. *Neuron*, 108(1):44–65. doi: 10.1016/j.neuron.2020.09.017 .
- Maxwell, S. E., Delaney, H. D., and Kelley, K. (2017). *Designing Experiments and Analyzing Data: a Model Comparison Perspective*. Routledge, New York.
- Mayerl, C., Gould, F., Adjerd, K., Edmonds, C., and German, R. (2023a). The Pathway from Anatomy and Physiology to Diagnosis: A Developmental Perspective on Swallowing and Dysphagia. *Dysphagia*, 38(1):33–41. doi: 10.1007/s00455-022-10449-x .
- Mayerl, C. J., Dial, T. R., Mainwaring, M. C., Heers, A. M., and German, R. Z. (2023b). Birth and the pathway to adulthood: integration across development, environment, and evolution. *Integrative and Comparative Biology*, page icad101. doi: 10.1093/icb/icad101 .
- McElreath, R. (2018). *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC.
- McGibbon, C. A. (2003). Toward a better understanding of gait changes with age and disablement: Neuromuscular adaptation. *Exercise and Sport Sciences Reviews*, 31(2).
- McHenry, M. J. and Hedrick, T. L. (2023). The science and technology of kinematic measurements in a century of Journal of Experimental Biology. *Journal of Experimental Biology*, 226(Suppl1):jeb245147. doi: 10.1242/jeb.245147 .
-

-
- McMahon, T. A. (1975). Allometry and biomechanics: Limb bones in adult ungulates. *The American Naturalist*, 109(969):547–563. doi: 10.1086/283026 .
- Meijer, E., Bertholle, C. P., Oosterlinck, M., van der Staay, F. J., Back, W., and van Nes, A. (2014). Pressure mat analysis of the longitudinal development of pig locomotion in growing pigs after weaning. *BMC veterinary research*, 10(1):1–11. doi: 10.1186/1746-6148-10-37 .
- Mellor, D. J. and Cockburn, F. (1986). A comparison of energy metabolism in the new-born infant, piglet and lamb. *Quarterly Journal of Experimental Physiology*, 71(3):361–379. doi: 10.1113/expphysiol.1986.sp002995 .
- Micheline, A., Eshraghi, A., and Andrysek, J. (2020). Two-dimensional video gait analysis: A systematic review of reliability, validity, and best practice considerations. *Prosthetics and Orthotics International*, 44(4):245–262. doi: 10.1177/030936462092129 .
- Mielke, F. (2020). Notebook series on inverse dynamics. "PhD Notes" blog by the author url: http://mielke-bio.info/falk/posts/10.id_lmx0a_wrenches. accessed 2023-06-22.
- Mielke, F. (2021). Inverse dynamics procedure in python and blender. Git repository by the author url: https://git.sr.ht/~falk/flying_femur. accessed 2023-06-22.
- Mielke, F. (2024). Linear models: Can out-of-sample prediction yield novel results? "PhD Notes" blog by the author url: <http://mielke-bio.info/falk/posts/33.linearmodels>. accessed 2024-01-31.
- Mielke, F., Amson, E., and Nyakatura, J. A. (2018a). Morpho-functional analysis using Procrustes superimposition by static reference. *Evolutionary Biology*, 45(4):449–461. doi: 10.1007/s11692-018-9456-9 .
- Mielke, F., Schunke, V., Wölfer, J., and Nyakatura, J. A. (2018b). Motion analysis of non-model organisms using a hierarchical model: Influence of setup enclosure dimensions on gait parameters of Swinhoe’s striped squirrels as a test case. *Zoology*, 129:35–44. doi: 10.1016/j.zool.2018.05.009 .
- Mielke, F., Van Ginneken, C., and Aerts, P. (2019). Quantifying intralimb coordination of terrestrial ungulates with Fourier Coefficient Affine Superimposition. *Zoological Journal of the Linnean Society*, 189(3):1067–1083. doi: 10.1093/zoolinlean/zlz135 .
- Mielke, F., Van Ginneken, C., and Aerts, P. (2023). A Workflow for Automatic, High Precision Livestock Diagnostic Screening of Locomotor Kinematics. *Frontiers in Veterinary Science*, 10. doi: 10.3389/fvets.2023.1111140 .
- Mielke, M., Aerts, P., Van Ginneken, C., Van Wassenbergh, S., and Mielke, F. (2020). Progressive tracking: a novel procedure to facilitate manual digitization of videos. *Biology Open*, 9(11). doi: 10.1242/bio.055962 .
- Mirkiani, S., Roszko, D. A., O’Sullivan, C. L., Faridi, P., Hu, D. S., Fang, D., Everaert, D. G., Toossi, A., Robinson, K., and Mushahwar, V. K. (2021). Overground Gait Kinematics and Muscle Activation Patterns in the Yucatan Mini Pig. *bioRxiv*. doi: 10.1101/2021.10.19.465020 .

-
- Mitteroecker, P. and Bookstein, F. (2011). Linear discrimination, ordination, and the visualization of selection gradients in modern morphometrics. *Evolutionary Biology*, 38(1):100–114. doi: 10.1007/s11692-011-9109-8 .
- Mohling, C. M., Johnson, A. K., Coetzee, J. F., Karriker, L. A., Abell, C. E., Millman, S. T., and Stalder, K. J. (2014). Kinematics as objective tools to evaluate lameness phases in multiparous sows. *Livestock Science*, 165:120–128. doi: 10.1016/j.livsci.2014.04.031 .
- Montuelle, S. J. and Williams, S. H. (2015). In Vivo Measurement of Mesokinesis in Gekko gecko: The Role of Cranial Kinesis during Gape Display, Feeding and Biting. *PLOS ONE*, 10(7):1–23. doi: 10.1371/journal.pone.0134710 .
- Mostayed, A., Mynuddin, M., Mazumder, G., Kim, S., and Park, S. J. (2008). Abnormal Gait Detection Using Discrete Fourier Transform. In *2008 International Conference on Multimedia and Ubiquitous Engineering (mue 2008)*, pages 36–40.
- Mota-Rojas, D., Orozco-Gregorio, H., Villanueva-Garcia, D., Bonilla-Jaime, H., Suarez-Bonilla, X., Hernandez-Gonzalez, R., Roldan-Santiago, P., Trujillo-Ortega, M., et al. (2011). Foetal and neonatal energy metabolism in pigs and humans: a review. *Veterinari Medicina*, 56(5):215–225. doi: 10.17221/1565-VETMED .
- Mudd, A. T. and Dilger, R. N. (2017). Early-life nutrition and neurodevelopment: Use of the piglet as a translational model. *Advances in Nutrition*, 8(1):92–104. doi: 10.3945/an.116.013243 .
- Muk, T., Leto, A., Brunse, A., Stensballe, A., Thymann, T., Sangild, P., and Nguyen, D. N. (2023). Neonatal prophylactic antibiotics after preterm birth affect plasma proteome and immune development in pigs. *Pediatric Research*. doi: 10.1038/s41390-023-02492-7 .
- Mull, R. (1984). Mass estimates by computed tomography: physical density from CT numbers. *American Journal of Roentgenology*, 143(5):1101–1104. doi: 10.2214/ajr.143.5.1101 .
- Müller, A. (2018). Screw and Lie group theory in multibody dynamics: recursive algorithms and equations of motion of tree-topology systems. *Multibody System Dynamics*, 42(2):219–248. doi: 10.1007/s11044-017-9583-6 .
- Muns, R., Manzanilla, E. G., Sol, C., Manteca, X., and Gasa, J. (2013). Piglet behavior as a measure of vitality and its influence on piglet survival and growth during lactation. *Journal of Animal Science*, 91(4):1838–1843. doi: 10.2527/jas.2012-5501 .
- Muybridge, E. (1893). *Descriptive zoopraxography, or, the science of animal locomotion made popular*. University of Pennsylvania.
- Myers, C. A., Laz, P. J., Shelburne, K. B., and Davidson, B. S. (2015). A probabilistic approach to quantify the impact of uncertainty propagation in musculoskeletal simulations. *Annals of biomedical engineering*, 43:1098–1111. doi: 10.1007/s10439-014-1181-7 .
- Nardone, R., Florea, C., Höller, Y., Brigo, F., Versace, V., Lochner, P., Golaszewski, S., and Trinka, E. (2017). Rodent, large animal and non-human primate models of spinal cord injury. *Zoology*, 123:101–114. doi: 10.1016/j.zool.2017.06.004 .
-

-
- NASA (2021). Conservation of momentum. NASA, ed. Nancy Hall
url: <https://www.grc.nasa.gov/www/k-12/airplane/conmo.html>. accessed 2023-07-18.
- Nelder, J. A. and Mead, R. (1965). A Simplex Method for Function Minimization. *The Computer Journal*, 7(4):308–313. doi: 10.1093/comjnl/7.4.308 .
- Netukova, S., Duspivova, T., Tesar, J., Bejtíc, M., Baxa, M., Ellederova, Z., Szabo, Z., and Krupicka, R. (2021). Instrumented pig gait analysis: State-of-the-art. *Journal of Veterinary Behavior*, 45:51–59. doi: 10.1016/j.jveb.2021.06.006 .
- Newell, K. M. and Liu, Y.-T. (2021). Collective Variables and Task Constraints in Movement Coordination, Control and Skill. *Journal of Motor Behavior*, 53(6):770–796. doi: 10.1080/00222895.2020.1835799 .
- Newton, I. (1687). *Philosophiae naturalis principia mathematica*. Trin. Coll. Cantab. Soc.
- Nguyen, V., Sanctorem, J. G., Van Wassenbergh, S., Dirckx, J. J. J., Sijbers, J., and De Beenhouwer, J. (2021). Geometry calibration of a modular stereo cone-beam x-ray ct system. *Journal of Imaging*, 7(3). doi: 10.3390/jimaging7030054 .
- Nicolas, G., Multon, F., Berillon, G., and Marchal, F. (2006). Plausible Locomotion for Bipedal Creatures Using Motion Warping and Inverse Kinematics. In Nishita, T., Peng, Q., and Seidel, H.-P., editors, *Advances in Computer Graphics*, pages 586–593, Berlin, Heidelberg. Springer Berlin Heidelberg.
- Nicolas, G., Multon, F., Berillon, G., and Marchal, F. (2007). From bone to plausible bipedal locomotion using inverse kinematics. *Journal of Biomechanics*, 40(5):1048–1057. doi: 10.1016/j.jbiomech.2006.04.010 .
- Nishikawa, K., Biewener, A. A., Aerts, P., Ahn, A. N., Chiel, H. J., Daley, M. A., Daniel, T. L., Full, R. J., Hale, M. E., Hedrick, T. L., Lappin, A. K., Nichols, T. R., Quinn, R. D., Satterlie, R. A., and Szymik, B. (2007). Neuromechanics: an integrative approach for understanding motor control. *Integrative and Comparative Biology*, 47(1):16–54. doi: 10.1093/icb/icm024 .
- Noether, E. (1918). Invariante Variationsprobleme. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse*, 1918:235–257.
- Normann, B. D. (2016). Gauss’ Law of Error Propagation ‘Explained’. Norwegian University of Science and Technology (NTNU), Department of Physics
url: <https://penofben.files.wordpress.com/2016/02/gausserrorpropagation2.pdf>. accessed 2023-07-11.
- Nyakatura, J. A., Andrada, E., Grimm, N., Weise, H., and Fischer, M. S. (2012). Kinematics and center of mass mechanics during terrestrial locomotion in northern lapwings (*Vanellus vanellus*, Charadriiformes). *Journal of Experimental Zoology Part A: Ecological Genetics and Physiology*, 317(9):580–594. doi: 10.1002/jez.1750 .
- Nyakatura, J. A., Melo, K., Horvat, T., Karakasiliotis, K., Allen, V. R., Andikfar, A., Andrada, E., Arnold, P., Lauströer, J., Hutchinson, J. R., Fischer, M. S., and Ijspeer, A. J. (2019). Reverse-engineering the locomotion of a stem amniote. *Nature*, 565(7739):351–355. doi: 10.1038/s41586-018-0851-2 .
-

-
- Nyakatura, J. A., Petrovitch, A., and Fischer, M. S. (2010). Limb kinematics during locomotion in the two-toed sloth (*Choloepus didactylus*, Xenarthra) and its implications for the evolution of the sloth locomotor apparatus. *Zoology*, 113(4):221–234. doi: 10.1016/j.zool.2009.11.003 .
- Nyquist, H. (1928). Certain topics in telegraph transmission theory. *Transactions of the American Institute of Electrical Engineers*, 47(2):617–644. doi: 10.1109/T-AIEE.1928.5055024 .
- Ogihara, N., Oku, T., Andrada, E., Blickhan, R., Nyakatura, J. A., and Fischer, M. S. (2014). Planar covariation of limb elevation angles during bipedal locomotion in common quails (*Coturnix coturnix*). *Journal of Experimental Biology*, 217(22):3968. doi: 10.1242/jeb.109355 .
- Osgood, B. G. (2007). Lecture Notes for EE261: The Fourier Transform and its Applications.
- Osgood, B. G. (2019). *Lectures on the Fourier transform and its applications*, volume 33 of *Pure and Applied Undergraduate Texts*. American Mathematical Soc.
- Parisi Dutra, R., Casali, D. d. M., Missagia, R. V., Gasparini, G. M., Perini, F. A., and Cozzuol, M. A. (2017). Phylogenetic systematics of peccaries (Tayassuidae: Artiodactyla) and a classification of South American tayassuids. *Journal of Mammalian Evolution*, 24(3):345–358. doi: 10.1007/s10914-016-9347-8 .
- Pataky, T. C. (2010). Generalized n-dimensional biomechanical field analysis using statistical parametric mapping. *Journal of Biomechanics*, 43(10):1976–1982. doi: 10.1016/j.jbiomech.2010.03.008 .
- Pataky, T. C. (2020). Online lecture: Discrete vs Continuous Biomechanical Data Analysis. Sports Biomechanics Lecture Series (host: Stuart McErlain-Naylor) url: <https://www.youtube.com/live/isLD00iE6kQ>. accessed 31/10/2023.
- Pataky, T. C., Caravaggi, P., Savage, R., Parker, D., Goulermas, J. Y., Sellers, W. I., and Crompton, R. H. (2008). New insights into the plantar pressure correlates of walking speed using pedobarographic statistical parametric mapping (pSPM). *Journal of Biomechanics*, 41(9):1987–1994. doi: 10.1016/j.jbiomech.2008.03.034 .
- Pataky, T. C., Robinson, M. A., and Vanrenterghem, J. (2013). Vector field statistical analysis of kinematic and force trajectories. *Journal of Biomechanics*, 46(14):2394–2401. doi: 10.1016/j.jbiomech.2013.07.031 .
- Patua, R., Muchhal, T., and Basu, S. (2021). Gait-Based Person Identification, Gender Classification, and Age Estimation: A Review. In Panigrahi, C. R., Pati, B., Mohapatra, P., Buyya, R., and Li, K.-C., editors, *Progress in Advanced Computing and Intelligent Engineering*, pages 62–74, Singapore. Springer Singapore.
- Pech, P., Bergström, K., Rauschnig, W., and Haughton, V. M. (1987). Attenuation values, volume changes and artifacts in tissue due to freezing. *Acta Radiologica*, 28(6):779–782. doi: 10.3109/02841858709177444 .
- Peham, C., Scheidl, M., and Licka, T. (1996). A method of signal processing in motion analysis of the trotting horse. *Journal of Biomechanics*, 29(8):1111–1114. doi: 10.1016/0021-9290(95)00179-4 .
-

-
- Pennestri, E. and Valentini, P. P. (2010). Dual quaternions as a tool for rigid body motion analysis: a tutorial with an application to biomechanics. *Archive of Mechanical Engineering*, LVII(2):187–205.
- Perrot, O., Laroche, D., Pozzo, T., and Marie, C. (2011). Kinematics of obstacle clearance in the rat. *Behavioural Brain Research*, 224(2):241–249. doi: 10.1016/j.bbr.2011.05.027 .
- Phelps, M. E., Gado, M. H., and Hoffman, E. J. (1975). Correlation of effective atomic number and electron density with attenuation coefficients measured with polychromatic x rays. *Radiology*, 117(3):585–588. doi: 10.1148/117.3.585 .
- Phillips, D. and Lannutti, J. (1997). Measuring physical density with X-ray computed tomography. *NDT & E International*, 30(6):339–350. doi: 10.1016/S0963-8695(97)00020-0 .
- Pike, A. V. L. and Alexander, R. M. (2002). The relationship between limb-segment proportions and joint kinematics for the hind limbs of quadrupedal mammals. *Journal of Zoology*, 258(4):427–433. doi: 10.1017/S0952836902001577 .
- Platt, J. R. (1964). Strong inference. *Science*, 146(3642):347–353. doi: 10.1126/science.146.3642.347 .
- Plocek, M. R. and Dunham, N. T. (2023). Spatiotemporal walking gait kinematics of semi-arboreal red pandas (*Ailurus fulgens*). *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology*. doi: 10.1002/jez.2725 .
- Polk, J. (2004). Influences of limb proportions and body size on locomotor kinematics in terrestrial primates and fossil hominins. *Journal of Human Evolution*, 47(4):237–252. doi: 10.1016/j.jhevol.2004.07.003 .
- Polk, J. D. (2002). Adaptive and phylogenetic influences on musculoskeletal design in cercopithecine primates. *Journal of Experimental Biology*, 205(21):3399. doi: 10.1242/jeb.205.21.3399 .
- Polk, J. D., Spencer-Smith, J. B., DiBerardino III., L. A., Ellis, D., Downen, M., and Rosengren, K. S. (2008). Quantifying variability in phase portraits: Application to gait ontogeny. *Infant Behavior and Development*, 31(2):302–306. doi: 10.1016/j.infbeh.2007.10.005 .
- Polly, P. D., Lawing, A. M., Fabre, A.-C., and Goswami, A. (2013). Phylogenetic principal components analysis and geometric morphometrics. *Hystrix, the Italian Journal of Mammalogy*, 24(1):33–41. doi: 10.4404/hystrix-24.1-6383 .
- Popper, K. (2002). *The logic of scientific discovery*. Routledge, 2 edition.
- Porges, S. W., Bohrer, R. E., Cheung, M. N., Drasgow, F., McCabe, P. M., and Keren, G. (1980). New time-series statistic for detecting rhythmic co-occurrence in the frequency domain: the weighted coherence and its application to psychophysiological research. *Psychological bulletin*, 88(3):580. doi: 10.1037/0033-2909.88.3.580 .
-

-
- Price, S. A. and Bininda-Emonds, O. R. P. (2009). A comprehensive phylogeny of extant horses, rhinos and tapirs (Perissodactyla) through data combination. *Zoosystematics and Evolution*, 85(2):277–292. doi: 10.1002/zoos.200900005 .
- Pronost, N., Dumont, G., Berillon, G., and Nicolas, G. (2006). Morphological and stance interpolations in database for simulating bipedalism of virtual humans. *The Visual Computer*, 22:4–13. doi: 10.1007/s00371-005-0350-y .
- Quiniou, N., Dagorn, J., and Gaudré, D. (2002). Variation of piglets’ birth weight and consequences on subsequent performance. *Livestock Production Science*, 78(1):63–70. doi: 10.1016/S0301-6226(02)00181-1 .
- Radlowski, E. C., Conrad, M. S., Lezmi, S., Dilger, R. N., Sutton, B., Larsen, R., and Johnson, R. W. (2014). A neonatal piglet model for investigating brain and cognitive development in small for gestational age human infants. *PLOS ONE*, 9(3):1–12. doi: 10.1371/journal.pone.0091951 .
- Raichlen, D. A., Pontzer, H., and Shapiro, L. J. (2013). A new look at the dynamic similarity hypothesis: the importance of swing phase. *Biology Open*, 2(10):1032. doi: 10.1242/bio.20135165 .
- Reilly, S. M., McElroy, E. J., and Biknevicius, A. R. (2007). Posture, gait and the ecological relevance of locomotor costs and energy-saving mechanisms in tetrapods. *Zoology*, 110(4):271–289. doi: 10.1016/j.zool.2007.01.003 .
- Ren, L., Jones, R. K., and Howard, D. (2007). Predictive modelling of human walking over a complete gait cycle. *Journal of Biomechanics*, 40(7):1567–1574. doi: 10.1016/j.jbiomech.2006.07.017 .
- Revell, L. J. (2009). Size-correction and principal components for interspecific comparative studies. *Evolution*, 63(12):3258–3268. doi: 10.1111/j.1558-5646.2009.00804.x .
- Rezende, E., Gomes, F., Chappell, M., and Garland Jr., T. (2009). Running behavior and its energy cost in mice selectively bred for high voluntary locomotor activity. *Physiological and Biochemical Zoology*, 82(6):662–679. doi: 10.1086/605917 .
- Riley, R. D., Ensor, J., Snell, K. I. E., Harrell, F. E., Martin, G. P., Reitsma, J. B., Moons, K. G. M., Collins, G., and van Smeden, M. (2020). Calculating the sample size required for developing a clinical prediction model. *BMJ*, 368. doi: 10.1136/bmj.m441 .
- Robertson, D. G. E., Caldwell, G. E., Hamill, J., Kamen, G., and Whittlesey, S. (2013). *Research methods in biomechanics*. Human kinetics. p. 115.
- Robertson, G. E., Caldwell, G. E., Hamill, J., Kamen, G., and Whittlesey, S., editors (2018). *Research Methods in Biomechanics*. Human Kinetics, 2nd edition.
- Roehe, R. and Kalm, E. (2000). Estimation of genetic and environmental risk factors associated with pre-weaning mortality in piglets using generalized linear mixed models. *Animal Science*, 70(2):227–240. doi: 10.1017/S1357729800054692 .
- Roelofs, S., Meijer, E., Baas, V., Dobrovolski, M., van der Staay, F. J., and Nordquist, R. E. (2019). Neurological functioning and fear responses in low and normal birth weight piglets. *Applied Animal Behaviour Science*, 220:104853. doi: 10.1016/j.applanim.2019.104853 .
-

-
- Rohlf, F. J. and Corti, M. (2000). Use of Two-Block Partial Least-Squares to Study Covariation in Shape. *Systematic Biology*, 49(4):740–753. doi: 10.1080/106351500750049806 .
- Rohlf, F. J. and Slice, D. (1990). Extensions of the Procrustes Method for the Optimal Superimposition of Landmarks. *Systematic Biology*, 39(1):40–59. doi: 10.2307/2992207 .
- Rosengren, K. S., Deconinck, F. J. A., DiBerardino III., L. A., Polk, J. D., Spencer-Smith, J. B., De Clercq, D., and Lenoir, M. (2009). Differences in gait complexity and variability between children with and without developmental coordination disorder. *Gait & Posture*, 29(2):225–229. doi: 10.1016/j.gaitpost.2008.08.005 .
- Ruina, A., Bertram, J. E. A., and Srinivasan, M. (2005). A collisional model of the energetic cost of support work qualitatively explains leg sequencing in walking and galloping, pseudo-elastic leg behavior in running and the walk-to-run transition. *Journal of Theoretical Biology*, 237(2):170–192. doi: 10.1016/j.jtbi.2005.04.004 .
- Ryder, O. A. and Steiner, C. C. (2011). Molecular phylogeny and evolution of the Perissodactyla. *Zoological Journal of the Linnean Society*, 163(4):1289–1303. doi: 10.1111/j.1096-3642.2011.00752.x .
- Røraas, T., Sandberg, S., Aarsand, A. K., and Støve, B. (2019). A Bayesian Approach to Biological Variation Analysis. *Clinical Chemistry*, 65(8):995–1005. doi: 10.1373/clinchem.2018.300145 .
- Salvatier, J., Wiecki, T. V., and Fonnesbeck, C. (2016). Probabilistic programming in Python using PyMC3. *PeerJ Computer Science*, 2:e55. doi: 10.7717/peerj-cs.55 .
- Sanctorum, J. G., Van Wassenbergh, S., Nguyen, V., De Beenhouwer, J., Sijbers, J., and Dirckx, J. J. (2020). Projection-angle-dependent distortion correction in high-speed image-intensifier-based x-ray computed tomography. *Measurement Science and Technology*, 32(3):035404. doi: 10.1088/1361-6501/abb33e .
- Sangild, P. T., Siggers, R. H., Schmidt, M., Elnif, J., Bjornvad, C. R., Thymann, T., Grondahl, M. L., Hansen, A. K., Jensen, S. K., Boye, M., Moelbak, L., Buddington, R. K., Weström, B. R., Holst, J. J., and Burrin, D. G. (2006). Diet- and Colonization-Dependent Intestinal Dysfunction Predisposes to Necrotizing Enterocolitis in Preterm Pigs. *Gastroenterology*, 130(6):1776–1792. doi: 10.1053/j.gastro.2006.02.026 .
- Schedlinski, C. and Link, M. (2001). A survey of current inertia parameter identification methods. *Mechanical Systems and Signal Processing*, 15(1):189–211. doi: 10.1006/mssp.2000.1345 .
- Scheidt, A., Ditzel, P. C., Geiger, S. M., Wagner, F. C., Mülling, C. K. W., and Nyakatura, J. A. (2022). A therian mammal with sprawling kinematics? Gait and 3D forelimb X-ray motion analysis in tamanduas. *Journal of Experimental Biology*, 225(12):jeb243625. doi: 10.1242/jeb.243625 .
- Schendel, M., Dekutoski, M., Ogilvie, J., Olsewski, J., and Wallace, L. (1995). Kinematics of the canine lumbar intervertebral joint: An in vivo study before and after adjacent instrumentation. *Spine*, 20(23):2555–2564.

-
- Schmidt, M. (2008). Forelimb proportions and kinematics: how are small primates different from other small mammals? *Journal of Experimental Biology*, 211(24):3775. doi: 10.1242/jeb.019802 .
- Schneider, E. and Chao, E. (1983). Fourier analysis of ground reaction forces in normals and patients with knee joint disease. *Journal of Biomechanics*, 16(8):591–601. doi: 10.1016/0021-9290(83)90109-4 .
- Schulte, J. A., Losos, J. B., Cruz, F. B., and Núñez, H. (2004). The relationship between morphology, escape behaviour and microhabitat occupation in the lizard clade *Liolaemus* (Iguanidae: Tropidurinae*: Liolaemini). *Journal of Evolutionary Biology*, 17(2):408–420. doi: 10.1046/j.1420-9101.2003.00659.x .
- Schwartz, M. H., Rozumalski, A., and Trost, J. P. (2008). The effect of walking speed on the gait of typically developing children. *Journal of Biomechanics*, 41(8):1639–1650. doi: 10.1016/j.jbiomech.2008.03.015 .
- Schwarz, D., Konow, N., Roba, Y. T., and Heiss, E. (2020). A salamander that chews using complex, three-dimensional mandible movements. *Journal of Experimental Biology*, 223(5):jeb220749. doi: 10.1242/jeb.220749 .
- Scott, B., Seyres, M., Philp, F., Chadwick, E. K., and Blana, D. (2022). Healthcare applications of single camera markerless motion capture: a scoping review. *PeerJ*, 10:e13517. doi: 10.7717/peerj.13517 .
- Seethapathi, N., Wang, S., Saluja, R., Blohm, G., and Kording, K. P. (2019). Movement science needs different pose tracking algorithms. *arXiv*, 1907.10226. doi: 10.48550/arXiv.1907.10226 .
- Shanno, D. F. (1970). A family of variable-metric methods derived by variational means. *Mathematics of Computation*, 24:647–656. doi: 10.1090/S0025-5718-1970-0274029-X .
- Shannon, C. (1949). Communication in the Presence of Noise. *Proceedings of the IRE*, 37(1):10–21. doi: 10.1109/JRPROC.1949.232969 .
- Shepard, E. L. C., Wilson, R. P., Rees, W. G., Grundy, E., Lambertucci, S. A., and Vosper, S. B. (2013). Energy landscapes shape animal movement ecology. *The American Naturalist*, 182(3):298–312. doi: 10.1086/671257 .
- Shikano, S. (2019). Hypothesis testing in the bayesian framework. *Swiss Political Science Review*, 25(3):288–299. doi: 10.1111/spsr.12375 .
- Shmueli, G. (2010). To Explain or to Predict? *Statistical Science*, 25(3):289 – 310. doi: 10.1214/10-STS330 .
- Shmueli, G. and Koppius, O. R. (2011). Predictive Analytics in Information Systems Research. *MIS Quarterly*, 35(3):553–572. doi: 10.2307/23042796 .
- Sijbers, J. and Postnov, A. (2004). Reduction of ring artefacts in high resolution micro-ct reconstructions. *Physics in Medicine & Biology*, 49(14):N247. doi: 10.1088/0031-9155/49/14/N06 .
-

-
- Six, N., De Beenhouwer, J., and Sijbers, J. (2019). poly-DART: A discrete algebraic reconstruction technique for polychromatic X-ray CT. *Opt. Express*, 27(23):33670–33682. doi: 10.1364/OE.27.033670 .
- Skejø, S. D., Lund, M. E., Stensvig, M., Kaae, N. M., and Rasmussen, J. (2021). Running in circles: Describing running kinematics using Fourier series. *Journal of Biomechanics*, 115:110187. doi: 10.1016/j.jbiomech.2020.110187 .
- Soudan, K. (1982). Standardization of Gait Kinematic Data Using a Gait Symmetry Index and Fourier Analysis. In Huiskes, R., van Campen, D. H., and de Wijn, J. R., editors, *Biomechanics: Principles and Applications: Selected Proceedings of the 3rd General Meeting of the European Society of Biomechanics Nijmegen, The Netherlands, 21–23 January 1982*, pages 135–140. Springer Netherlands. doi: 10.1007/978-94-009-7678-8_10 .
- Spengler, D., Rintz, N., and Krause, M. F. (2019). An Unsettled Promise: The Newborn Piglet Model of Neonatal Acute Respiratory Distress Syndrome (NARDS). Physiologic Data and Systematic Review. *Frontiers in Physiology*, 10. doi: 10.3389/fphys.2019.01345 .
- Sponberg, S., Abbott, E., and Sawicki, G. S. (2023). Perturbing the muscle work loop paradigm to unravel the neuromechanics of unsteady locomotion. *Journal of Experimental Biology*, 226(7):jeb243561. doi: 10.1242/jeb.243561 .
- Stamos, P. A. and Alemseged, Z. (2023). Hominin locomotion and evolution in the Late Miocene to Late Pliocene. *Journal of Human Evolution*, 178:103332. doi: 10.1016/j.jhevol.2023.103332 .
- Stankowich, T. and Caro, T. (2009). Evolution of weaponry in female bovids. *Proceedings of the Royal Society B: Biological Sciences*, 276(1677):4329–4334. doi: 10.1098/rspb.2009.1256 .
- Stavarakakis, S., Guy, J., Warlow, O., Johnson, G., and Edwards, S. (2014). Walking kinematics of growing pigs associated with differences in musculoskeletal conformation, subjective gait score and osteochondrosis. *Livestock Science*, 165:104–113. doi: 10.1016/j.livsci.2014.04.008 .
- Stein, B. R. and Casinos, A. (1997). What is a cursorial mammal? *Journal of Zoology*, 242(1):185–192. doi: 10.1111/j.1469-7998.1997.tb02939.x .
- Steudel-Numbers, K. and Weaver, T. D. (2006). Froude number corrections in anthropological studies. *American Journal of Physical Anthropology*, 131(1):27–32. doi: 10.1002/ajpa.20404 .
- Stiffler-Joachim, M. R., Wille, C., Kliethermes, S., and Heiderscheit, B. (2020). Factors Influencing Base of Gait During Running: Consideration of Sex, Speed, Kinematics, and Anthropometrics. *Journal of Athletic Training*, 55(12):1300–1306. doi: 10.4085/1062-6050-565-19 .
- Stockwell, R., Mansinha, L., and Lowe, R. (1996). Localization of the complex spectrum: the S transform. *IEEE Transactions on Signal Processing*, 44(4):998–1001. doi: 10.1109/78.492555 .

-
- Stoessel, A. and Fischer, M. S. (2012). Comparative intralimb coordination in avian bipedal locomotion. *Journal of Experimental Biology*, 215(23):4055. doi: 10.1242/jeb.070458 .
- Sullivan, S., Friess, S. H., Ralston, J., Smith, C., Propert, K. J., Rapp, P. E., and Margulies, S. S. (2013). Improved Behavior, Motor, and Cognition Assessments in Neonatal Piglets. *Journal of Neurotrauma*, 30(20):1770–1779. doi: 10.1089/neu.2013.2913 .
- Tanabe, S., Fujita, S., Yoshihara, S., Sakamoto, A., and Yamamoto, S. (2005). YAG glass-ceramic phosphor for white LED (II): luminescence characteristics. In Ferguson, I. T., Carrano, J. C., Taguchi, T., and Ashdown, I. E., editors, *Fifth International Conference on Solid State Lighting*, volume 5941, page 594112. International Society for Optics and Photonics, SPIE.
- Ting, L. H., Chiel, H. J., Trumbower, R. D., Allen, J. L., McKay, J. L., Hackney, M. E., and Kesar, T. M. (2015). Neuromechanical principles underlying movement modularity and their implications for rehabilitation. *Neuron*, 86(1):38–54. doi: 10.1016/j.neuron.2015.02.042 .
- Tipler, P. A. and Mosca, G. (2007). *Physics for scientists and engineers*. Macmillan, 6th edition.
- Triche, B. L., Nelson, J. T., McGill, N. S., Porter, K. K., Sanyal, R., Tessler, F. N., McConathy, J. E., Gauntt, D. M., Yester, M. V., and Singh, S. P. (2019). Recognizing and Minimizing Artifacts at CT, MRI, US, and Molecular Imaging. *RadioGraphics*, 39(4):1017–1018. doi: 10.1148/rg.2019180022 .
- van de Schoot, R., Depaoli, S., King, R., Kramer, B., Märtens, K., Tadesse, M. G., Vannucci, M., Gelman, A., Veen, D., Willemsen, J., et al. (2021). Bayesian Statistics and Modelling. *Nature Reviews Methods Primers*, 1(1):1–26. doi: 10.1038/s43586-020-00001-2 .
- Van Ginneken, C., Ayuso, M., Van Bockstal, L., and Van Cruchten, S. (2022). Prewaning performance in intrauterine growth-restricted piglets: Characteristics and interventions. *Molecular Reproduction and Development*. doi: 10.1002/mrd.23614 .
- Van Gompel, G., Van Slambrouck, K., Defrise, M., Batenburg, K. J., de Mey, J., Sijbers, J., and Nuyts, J. (2011). Iterative correction of beam hardening artifacts in ct. *Medical Physics*, 38(S1):S36–S49. doi: 10.1118/1.3577758 .
- Van Tichelen, K., Prims, S., Ayuso, M., Van Kerschaver, C., Vandaele, M., Degroote, J., Van Cruchten, S., Michiels, J., and Van Ginneken, C. (2021). Handling Associated with Drenching Does Not Impact Survival and General Health of Low Birth Weight Piglets. *Animals*, 11(2). doi: 10.3390/ani11020404 .
- van Weeren, P. R., van den Bogert, A. J., and Barneveld, A. (1992). Correction models for skin displacement in equine kinematics gait analysis. *Journal of Equine Veterinary Science*, 12(3):178–192. doi: 10.1016/S0737-0806(06)81478-4 .
- Vanden Hole, C. (2019). Locomotor Development in the Newborn Pig. *PhD Thesis*.
- Vanden Hole, C., Aerts, P., Prims, S., Ayuso, M., Van Cruchten, S., and Van Ginneken, C. (2018a). Does intrauterine crowding affect locomotor development? A comparative

-
- study of motor performance, neuromotor maturation and gait variability among piglets that differ in birth weight and vitality. *PLOS ONE*, 13(4):1–21. doi: 10.1371/journal.pone.0195961 .
- Vanden Hole, C., Ayuso, M., Aerts, P., Prims, S., Van Cruchten, S., and Van Ginneken, C. (2019). Glucose and glycogen levels in piglets that differ in birth weight and vitality. *Heliyon*, 5(9):e02510. doi: 10.1016/j.heliyon.2019.e02510 .
- Vanden Hole, C., Ayuso, M., Aerts, P., Van Cruchten, S., Thymann, T., Sangild, P. T., and Van Ginneken, C. (2021). Preterm Birth Affects Early Motor Development in Pigs. *Frontiers in Pediatrics*, 9:1003. doi: 10.3389/fped.2021.731877 .
- Vanden Hole, C., Cleuren, S., Van Ginneken, C., Prims, S., Ayuso, M., Van Cruchten, S., and Aerts, P. (2018b). How does intrauterine crowding affect locomotor performance in newborn pigs? A study of force generating capacity and muscle composition of the hind limb. *PLOS ONE*, 13(12):1–18. doi: 10.1371/journal.pone.0209233 .
- Vanden Hole, C., Goyens, J., Prims, S., Fransen, E., Ayuso Hernando, M., Van Cruchten, S., Aerts, P., and Van Ginneken, C. (2017). How innate is locomotion in precocial animals? A study on the early development of spatio-temporal gait variables and gait symmetry in piglets. *Journal of Experimental Biology*, 220(15):2706–2716. doi: 10.1242/jeb.157693 .
- Vanhooydonck, B., James, R. S., Tallis, J., Aerts, P., Tadic, Z., Tolley, K. A., Measey, G. J., and Herrel, A. (2014). Is the whole more than the sum of its parts? Evolutionary trade-offs between burst and sustained locomotion in lacertid lizards. *Proceedings of the Royal Society of London B: Biological Sciences*, 281:2132677. doi: 10.1098/rspb.2013.2677 .
- Vaughan, C. L. and O'Malley, M. J. (2005). Froude and the contribution of naval architecture to our understanding of bipedal locomotion. *Gait & Posture*, 21(3):350–362. doi: 10.1016/j.gaitpost.2004.01.011 .
- von Wachenfelt, H., Pinzke, S., Nilsson, C., Olsson, O., and Ehlorsson, C.-J. (2008). Gait analysis of unprovoked pig gait on clean and fouled concrete surfaces. *Biosystems Engineering*, 101(3):376–382. doi: 10.1016/j.biosystemseng.2008.09.002 .
- Wang, J., Yang, M., Cao, M., Lin, Y., Che, L., Duraipandiyan, V., Al-Dhabi, N. A., Fang, Z., Xu, S., Feng, B., Liu, G., and Wu, D. (2016). Moderately increased energy intake during gestation improves body condition of primiparous sows, piglet growth performance, and milk fat and protein output. *Livestock Science*, 194:23–30. doi: 10.1016/j.livsci.2016.09.012 .
- Webb, D. and Sparrow, W. A. (2007). Description of joint movements in human and non-human primate locomotion using Fourier analysis. *Primates*, 48(4):277–292. doi: 10.1007/s10329-007-0043-4 .
- Weber, W. and Weber, E. (1836). *Die Mechanik Der Menschlichen Gerverzeuge*.
- Welch, P. (1967). The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE Transactions on Audio and Electroacoustics*, 15(2):70–73. doi: 10.1109/TAU.1967.1161901 .

-
- Wells, J. P. and DeMenthon, D. F. (1987). Measurement of body segment mass, center of gravity, and determination of moments of inertia by double pendulum in lemur fulvus. *American Journal of Primatology*, 12(3):299–308. doi: 10.1002/ajp.1350120307 .
- Wheat, J. S. and Glazier, P. S. (2006). Measuring coordination and variability in coordination. In Davids, K., Bennett, S., and Newell, K. M., editors, *Movement System Variability*, chapter 9. Human Kinetics.
- Widdowson, E. (1971). Intra-Uterine Growth Retardation in the Pig. I. Organ Size and Cellular Development at Birth and after Growth to Maturity: I. Organ Size and Cellular Development at Birth and after Growth to Maturity. *Biologia Neonatorum*, 19(4-6):329–340. doi: 10.1159/000240427 .
- Wiecki, T. (2013). This world is far from Normal(ly distributed): Bayesian Robust Regression in PyMC3. While My MCMC Gently Samples (blog)
url: <https://twiecki.io/blog/2013/08/27/bayesian-glms-2>. accessed 2023-07-25.
- Wikipedia/Moment of Inertia (2023). Moment of inertia tensor. Wikimedia Foundation
url: https://en.wikipedia.org/wiki/Moment_of_inertia#Inertia_tensor.
accessed 2023-07-09.
- Wikipedia/Newton-Euler equations (2012). Moment of inertia tensor. Wikimedia Foundation
url: https://en.wikipedia.org/wiki/Newton-Euler_equations. accessed 2023-07-21.
- Wikipedia/Torque (2012). Moment of inertia, resultant torque. Wikimedia Foundation
url: https://en.wikipedia.org/wiki/Moment_of_inertia#Resultant_torque.
accessed 2023-07-21.
- Williams, M. L. and Galvis, A. G. (1974). Pulmonary complications in infants. symposium on pulmonary problems in surgery. *Surgical Clinics of North America*, 54(5):1137–1154. doi: 10.1016/S0039-6109(16)40443-3 .
- Wischner, D., Kemper, N., and Krieter, J. (2009). Nest-building behaviour in sows and consequences for pig husbandry. *Livestock Science*, 124(1):1–8. doi: 10.1016/j.livsci.2009.01.015 .
- Wootton, R., Flecknell, P. A., Royston, J. P., and John, M. (1983). Intrauterine growth retardation detected in several species by non-normal birthweight distributions. *Reproduction*, 69(2):659 – 663. doi: 10.1530/jrf.0.0690659 .
- Worsley, K. J., Taylor, J. E., Tomaiuolo, F., and Lerch, J. (2004). Unified univariate and multivariate random field theory. *NeuroImage*, 23:S189–S195. doi: 10.1016/j.neuroimage.2004.07.026 . Mathematics in Brain Imaging.
- Wu, G., Bazer, F. W., Wallace, J. M., and Spencer, T. E. (2006). Intrauterine growth retardation: Implications for the animal sciences. *Journal of Animal Science*, 84(9):2316–2337. doi: 10.2527/jas.2006-156 .
- Yang, Q., Chakraborty, N., Lakshtanov, D., Sheppard, A., and Kingston, A. (2021). Density estimation in x-ray computed tomography using the Alvarez-Macovski model. In *Developments in X-Ray Tomography XIII*, volume 11840, pages 64–77. SPIE Digital Library.
-

-
- Young, J. W., Mayerl, C. J., Mannava, A., Lewis, C., Fan, T., Nair, M., Mamone, C., Schapker, N. M., Mossor, A. M., and German, R. Z. (2023a). Evaluating The Precocial-altricial Axis of Motor Skill at Birth in A Preterm Pig Model. *Integrative and Comparative Biology*, page icad013. doi: 10.1093/icb/icad013 .
- Young, J. W. and Shapiro, L. J. (2018). Developments in development: What have we learned from primate locomotor ontogeny? *American Journal of Physical Anthropology*, 165(S65):37–71. doi: 10.1002/ajpa.23388 .
- Young, M. W., McKamy, A. J., Dickinson, E., Yarbrow, J., Ragupathi, A., Guru, N., Avey-Arroyo, J. A., Butcher, M. T., and Granatosky, M. C. (2023b). Three toes and three modes: Dynamics of terrestrial, suspensory, and vertical locomotion in brown-throated three-toed sloths (Bradypodidae, Xenarthra). *Journal of Experimental Zoology Part A: Ecological and Integrative Physiology*, 339(4):383–397. doi: 10.1002/jez.2684 .
- Zehr, E. P., Barss, T. S., Dragert, K., Frigon, A., Vasudevan, E. V., Haridas, C., Hundza, S., Kaupp, C., Klarner, T., Klimstra, M., et al. (2016). Neuromechanical interactions between the limbs during human locomotion: an evolutionary perspective with translation to rehabilitation. *Experimental brain research*, 234:3059–3081. doi: 10.1007/s00221-016-4715-4 .
- Zhu, C., Byrd, R. H., Lu, P., and Nocedal, J. (1997). Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on mathematical software (TOMS)*, 23(4):550–560.
- Zscheile, F. P. (1934). An Improved Method for the Purification of Chlorophylls A and B; Quantitative Measurement of Their Absorption Spectra; Evidence for the Existence of a Third Component of Chlorophyll. *Botanical Gazette*, 95(4):529–562. doi: 10.1086/334429 .
- Zurano, J. P., Magalhães, F. M., Asato, A. E., Silva, G., Bidau, C. J., Mesquita, D. O., and Costa, G. C. (2019). Cetartiodactyla: updating a time-calibrated molecular phylogeny. *Molecular Phylogenetics and Evolution*, 133:256–262. doi: 10.1016/j.ympev.2018.12.015 .

9.2 Video Digitization

To perform any kinematic analysis, one has to extract kinematic data. Therefore, one has to “track” the trajectory of points of interest on a digital video.

The standard way to do this has changed dramatically during the course of this PhD project: whereas the gold standard used to be having a student assistant to click points frame-wise on the videos (“video tracking” of points), there are now deep learning algorithms to replace those students (“pose estimation” of whole animals). Some relevant publications are listed here for the interested reader. More details can be found in the articles marked as “review”. The author was involved in one of the studies (Mielke et al., 2020).

- (review) DLTv: Hedrick (2008)
- Argus: Jackson et al. (2016)
- (review) Progressive Tracking: Mielke et al. (2020)
- DeepLabCut: Mathis et al. (2018, 2020)
- ThruTracker: Corcoran et al. (2021)
- AniPose: Karashchuk et al. (2021)
- (review) overview of deep learning methods: Cronin (2021)

9.3 Statistical Parametric Mapping

A classical purpose of kinematics research is the identification of differences between study groups. For example, one might look at low- and normal birth weight piglets, record hip joint angle profiles for approximately a hundred walking strides of each group, and ask at what temporal intervals they differ. One key property of joint angle profiles is that temporal samples are not independent. In other words: joint angles do not change at random, and the joint angle at a certain frame in a video will be more similar to the previous and next frame than what would be expected by chance.

This is a well-known property. Naïvely, one might choose to overcome this by selecting key points in the joint angle profile (e.g. the peak flexion) and compare those events between study groups. However, this immediately raises questions about the number and choice/relevance of key points. Choice and relevance are rather philosophical questions; potentially they are motivated by a desired outcome (focusing points that differ). The number of key points has effects on significance thresholds (multiple testing), and it has been shown that comparing key points will increase the chance of false positives in difference detection (due to true measurement uncertainty).

Instead of points, one might look at intervals, yet again it would be desirable to have a non-subjective, automatic criterium of how to choose those. This can be provided by random field theory (Brett et al., 2003; Kemeny et al., 1976). In random field theory, the interdependence of measurement samples can be quantified by measuring how random changes between samples are. Randomness and smoothness are on two sides of a continuum, and the less random the changes between samples are, the “wider” the range of dependent data points can be considered. This concept has found its way into the framework of “Statistical Parametric Mapping” (SPM), which originates in neuroscience and functional imaging of the brain (Flandin and Friston, 2008; Friston, 2003; Friston et al., 1994; Worsley et al., 2004), and was adapted for biomechanic research (Pataky, 2010; Pataky et al., 2008, 2013).

With some limitations, one can say that SPM enables the hypothesis-bases statistical comparison of joint angle profiles. Some of the limitations are the sensitivity to registration (especially non-linear registration techniques), to processing (e.g. smoothing), and the normality assumption. It retains all the limitations of the chosen hypothesis tests (e.g. normality assumption often not confirmed). It must be added that SPM does not generally account for the cyclicity of joint angle profiles (edge effects), though this can certainly be overcome. As the authors of the method put it (Pataky, 2020), the only purpose of SPM is the transfer of statistical methods to a multidimensional, interdependent data situation.

This relates to FCAS on several flanks. FCAS can be used for spatiotemporal alignment (removing the phase differences), thereby reducing the dependency on more or less arbitrarily chosen key points in the joint angles. FCAS quantifies the temporal interrelation of sample points by translating it to harmonic components. However, FCAS is a transformation, whereas SPM is an analysis method: in fact, SPM can be used to compare the “coordination” residuals of superimposed joint angle profiles, or to compare relative angle profiles. In that regard, they are complementary.

9.4 List of Publications

The following resources with contributions by the author are publicly available. Updated information can be found on the author’s “orcid” profile.

9.4.1 Journal Articles

- **Mielke F**, Schunke V, Wölfer J and Nyakatura JA (2018), “Motion analysis of non-model organisms using a hierarchical model: Influence of setup enclosure dimensions on gait parameters of Swinhoe’s striped squirrels as a test case”, *Zoology*. Vol. 129, pp. 35-44. <https://doi.org/10.1016/j.zool.2018.05.009>
- **Mielke F**, Amson E and Nyakatura JA (2018), “Morpho-Functional Analysis Using Procrustes Superimposition by Static Reference”, *Evolutionary Biology*, December, 2018. Vol. 45(4), pp. 449-461. <https://doi.org/10.1007/s11692-018-9456-9>
- **Mielke F**, Van Ginneken C and Aerts P (2019), “Quantifying intralimb coordination of terrestrial ungulates with Fourier Coefficient Affine Superimposition”, *Zoological Journal of the Linnean Society*, 11, 2019. Vol. 189(3), pp. 1067-1083. <https://doi.org/10.1093/zoolinlean/zl135>
- Mielke M, Aerts P, Van Ginneken C, Van Wassenbergh S and **Mielke F** (2020), “Progressive tracking: a novel procedure to facilitate manual digitization of videos”, *Biology Open*, 11, 2020. Vol. 9(11). <https://doi.org/10.1242/bio.055962>
- Nguyen V, Alves Pereira LF, Liang Z, **Mielke F**, Van Houtte J, Sijbers J and De Beenhouwer J (2022), “Automatic landmark detection and mapping for 2D/3D registration with BoneNet”, *Frontiers in Veterinary Science*. 9, 923449. <https://doi.org/10.3389/fvets.2022.923449>
- **Mielke F**, Van Ginneken C and Aerts P (2023), “A Workflow for Automatic, High Precision Livestock Diagnostic Screening of Locomotor Kinematics”, *Frontiers in Veterinary Science*. Vol. 10. <https://doi.org/10.3389/fvets.2023.1111140>
- Aerts P, **Mielke F**, Vanden Hole C, Van Gorp MJ and Van Ginneken C (2023), “Early development of locomotion in the term piglet model: does size matter?”, *Integrative and Comparative Biology*. pp. icad054. <https://doi.org/10.1093/icb/icad054>

9.4.2 Conferences and Public Events

(first author contributions only)

- Annual Meeting of the German Zoological Society (DZG, 2019): “The Shape of Locomotion: A Method for Large Scale Analyses of Intra-Limb Coordination”. Recording of pre-talk available.
- Annual Meeting of the Society for Integrative and Comparative Biology (SICB, 2021): “Swing it like a piglet!”. Recording available.
- Annual Meeting of the Society of Experimental Biology (SEB, 2021): “CT Density Approximation to Integrate Inverse Dynamics into the XROMM Workflow”. Recording available.

9.4.3 Code Repositories

Here you find the code and code history of recent projects around this thesis.

- FCAS minimal code in various programming languages.
https://git.sr.ht/~falk/fcas_code
- Free and Open Source XROMM workflow, using blender and avoiding XMALab (for reasons of frustration and ideology).
https://git.sr.ht/~falk/foss_xromm
- Piglet kinematics analysis using FCAS and probabilistic modeling.
https://git.sr.ht/~falk/piglet_fcas
- Application of FCAS to baboon data.
https://git.sr.ht/~falk/papio_fcas
- Code of the “flying femur” project.
https://git.sr.ht/~falk/flying_femur
- Automatic video/image segmentation.
https://git.sr.ht/~falk/piglet_image_autosegmentation
- Thesis defence presentation.
<https://git.sr.ht/~falk/defence>
- PhD thesis wrapping.
<https://git.sr.ht/~falk/thesis>

9.4.4 Blog Posts

All listed posts are self-published and can be found on the website <http://mielke-bio.info/falk>.

Force Plate Calibration

- equipment considerations
<http://mielke-bio.info/falk/posts/01.fpequipment>
- basic calculations
<http://mielke-bio.info/falk/posts/02.fpcalculations>
- calibration of force magnitude
<http://mielke-bio.info/falk/posts/03.fpcalibration1>
- calibration of contact point
<http://mielke-bio.info/falk/posts/04.fpcalibration2>

Video Analysis

- landmark tracking
<http://mielke-bio.info/falk/posts/07.tracking>
- contrast enhancement
http://mielke-bio.info/falk/posts/08.image_contrast
- image undistortion
<http://mielke-bio.info/falk/posts/09.undistortion>
- stereo camera calibration
http://mielke-bio.info/falk/posts/17.camera_calibration
- automatic stride cycle extraction
http://mielke-bio.info/falk/posts/27.cycle_extraction
- image segmentation
http://mielke-bio.info/falk/posts/30.diff_lineseg

Statistics

- probabilistic statistical modeling
<http://mielke-bio.info/falk/posts/06.patella>
- linear model prediction capability
http://mielke-bio.info/falk/posts/33.linear_models

Inverse Dynamics

- wrenches
http://mielke-bio.info/falk/posts/10.id_lmx0a_wrenches
- fictitious forces - overview
http://mielke-bio.info/falk/posts/11.id_lmx0b_fforces
- Euler force
http://mielke-bio.info/falk/posts/12.id_lmx1_stutter
- Centrifugal force
http://mielke-bio.info/falk/posts/13.id_lmx2_vinyl
- Coriolis force
http://mielke-bio.info/falk/posts/14.id_lmx3_slider
- D'Alembert force
http://mielke-bio.info/falk/posts/15.id_lmx4_arm
- **Generalization** - n-link arm
http://mielke-bio.info/falk/posts/16.id_lmx5_nlink

CT Analysis

- density approximation from CT data
http://mielke-bio.info/falk/posts/23.ct_density

Miscellaneous

- Procrustes superimposition
<http://mielke-bio.info/falk/posts/18.procrustes>
- quaternions
<http://mielke-bio.info/falk/posts/21.quaternions>
- photogrammetry
<http://mielke-bio.info/falk/posts/22.photogrammetry>
- custom wildlife cameras
<http://mielke-bio.info/falk/posts/24.funcameras>
- academic writing software use
<http://mielke-bio.info/falk/posts/32.scird2023>

9.5 List of Abbreviations

BFGS	– Broyden, Fletcher, Goldfrab, Shanno (Algorithm)
BMD	– Bone Mineral Density
BMI	– Body Mass Index
CAD	– Computer Aided Design (models)
CC	– Coordination Component
$\vec{COM}$	– Center Of Mass
CRP	– Continuous Relative Phase
CT	– computed tomography
DECT	– Dual-Energy CT
DFT	– Discrete Fourier Transform
DLC	– DeepLabCut (video digitization software)
DOF	– Degrees Of Freedom
eROM	– effective Range Of Motion
FBP	– Filtered Backprojection (CT reconstruction algorithm)
FCAS	– Fourier Coefficient Affine Superimposition
FFT	– Fast Fourier Transform
FSD	– Fourier Series Decomposition
$\vec{g}$, g	– gravitational acceleration or its magnitude
HDI	– Highest probability Density Interval
HDS	– High Duty factor Stride
HMC	– Hamiltonian Monte Carlo
HMPE	– high-molecular-weight polyethylene
I	– Mass Moment of Inertia
ICA	– Independent Component Analysis
ICS	– Inertial Coordinate System
iff	– if, and only if
LBW	– (piglet with) Low Birth Weight
LDS	– Low Duty factor Stride
LED	– Light Emitting Diode
LF	– left frontlimb
LH	– left hindlimb
LKJ	– Lewandowski-Kurowicka-Joe prior distribution

LOO	– Leave One Out
MCMC	– Markov Chain Monte Carlo
NBW	– (piglet with) Normal Birth Weight
NUTS	– No U-Turn Sampler
$O(\dots)$	– in the Order of ... (magnitude)
PA6	– Nylon 6 / Polycaprolactam
PC	– Polycarbonate
PC, PCi	– (i^{th}) Principal Component
PCA	– Principal Component Analysis
PCs	– Principal Components (see PCA)
PD	– Procrustes Distance
PE	– Polyethylene
PEEK	– Polyether ether ketone
PP	– Polypropylene
PTFE	– Polytetrafluoroethylene
PVC	– Polyvinylchloride
RF	– right frontlimb
RH	– right hindlimb
ROM	– Range Of Motion
SPM	– Statistical Parametric Mapping
WAIC	– Widely Applicable Information Criterion
XROMM	– X-ray Reconstruction of Moving Morphology

9.6 Acknowledgements

I am grateful for my time at Campus Drie Eiken, and wholeheartedly enjoyed being around for so long. In all the research groups which accommodated me, I was welcomed with a supportive, friendly, collaborative, and creative atmosphere, which is all due to the great people I met there.

As seems to be the nature of these things, my project took a bit longer, and so I saw many fellow students and researchers leave, and others join. I am glad to have met all of you inspiring, young and creative people, and I wish each of you all the best! I will not attempt to list all of you PhD students and postdocs, nor will I pick individuals. The reasons are your sheer number (three different departments!), an approaching print deadline, and the fear of missing any important person. And I am humbled by those of you who did take or will take the time to write individual acknowledgements, how cool is that?

Kudos, my friends! Keep on the great research!

However, I must point out my “office roommates”: Miriam Ayuso Hernando, Lieselotte Van Bockstal, Sara Prims, Joaquim Sanctorum, Simon Baeckens, and Maja Mielke. Thank you for being mild (or at least pretending to not care about) my chaotic desks and loud keyboards.

A special shout-out to the phenomenal technical and administrative support staff for whom I caused a lot of work: Katty Huybrechts, Gunther Vrolix, Steve Van Bienst, Josie Meaney-Ward, Jan Scholliers and Jorrit Mertens. I appreciate your great efforts and your open ear to my problems and frustrations. In contrast to some other “supportive” and “administrative” departments, you were actually and personally there, and oftentimes I would have been lost without you. You are who keep the labs running – modern pastoral care, in a world full of compliance bullshit and supply chain crackdowns. Thank you!

Also a technical staff by now, and a great support, I am thankful for the brief period I teamed up with Joaquim Sanctorum to operate the “DynXLab”. I am glad that you took it into your able hands, and wish you all the best, there or elsewhere!

I wave my virtual hand and bow to the participants of our virtual COVID Coffee Club, the “UA Code Clinic”, the 2019 research day committee, and regular TGIF guests; we should do all that again! And I am glad that the “Vrienden van de Ark”, the caretakers (mainly Denise, An, Ellen), and Harlee were there to teach me various cool skills such as handling cows, taking a Lama’s temperature, or how to operate a “potstall”. Matter of fact, I enjoyed “arkdienst”, and occasionally have to force myself to think about that monstrous dung heap, to counteract a wistful memory creeping in.

None of my studies and research, as meaningless as they might be, would not have been possible without the inspiring work of Alexandra Elbakyan (sci-hub) and the developers of “yt-dlp”. There are a tremendous number of people who depend on your enormous efforts, and think that you deserve all our thanks and praises for your courage.

A collegial “Thank You!” to all the reviewers who took time to read and improve my often cryptic or technical texts. And, not less, one to those reviewers who did not care or did not have time, and just let me pass. I appreciate the efforts of all my jury members: I was struggling in my pre-defence, yet in a good way, and the reason is that all of you took the time to carefully assess my lengthy research reports. I acknowledge that I am not the most “readable” writer, and thus I much appreciate your efforts and feedback, thank you!

There is one special person to fit many of the above categories. Of course, I mean Sam Van Wassenbergh, who, besides other roles, is a professor and still has at least half a foot in D-1.39.

Do not be too sad, Sam. You cannot always be torn in two. You will have to be one and whole, for many years. You have so much to enjoy and to be, and to do.

It was not quite “Mount Doom”, but I remember us standing in the fully prepared recording room, a foot deep in water, between a high-voltage closet, a self-soldered “trogger” (shorthand for “Trigger-Logger”, an add-on device necessary because the force plate recordings would require all of my computer’s capacity), and a unique piece of Jan’s phenomenal balsa wood craftsmanship. Someone had removed a critical pump, and our x-ray room was a swimming pool. It must have been a nightmare come true. Yet we focused. We cleaned up. And we made it. Thank you so much!

I will not forget my supervisors!

Chris. You protected me when I was threatened with administrative punishment (e.g. “no show arbeidsarts”) or disaster (e.g. “gelijkstelling FELASA”). I wholeheartedly thank you for the more memorable (computer-free) days of my PhD, involving unforgettable trips to remote farms full of farrowing sows, or the scanning of a giraffe skeleton. By far the most memorable event was a trip to the zoo; I am so glad that I was allowed to join. You were always there with sharp wits; you showed an incredible stamina to read texts which were full of jargon from other fields, and nonetheless you saved me from many misformulations and mistakes. You are a great person, and a role model, and I wish you all thinkable luck and success for the possible new role as vice rector, or any other challenge you take beside or beyond it!

Peter. For some reason, I feel that we had a bumpy start on the professional side, which involved several occasions where we would exchange page-long mails (sorry!) or hold intense discussions for hours (“referentierooster!”). Then we would usually clarify some basic terminology to find out that we are actually on the same page. I must take those on me, because by now I got to know you as a knowledgeable, generous and patient supervisor; I enjoy our scientific discussions and hope that we will have many more chances. I am sorry for the one or the other night which you might have been awake contemplating one or the other dead end of this project. ...or my incapability to obey university rules and authorities (e.g. “opleidingskrediet”, “edumaster enrolment”). And I admire your attitude to climb the attic to convince me that fictitious forces don’t exist (or, do they?), to write very useful code (though personally I would never choose Matlab), and to keep on pursuing the flamingo project. I sincerely hope that we do not “lose” you to pension, but “win” you time for the fun research projects which have been lingering in the back of your head (if you like). Enjoy, anyways, and all the best for it!

Both of you have found a good mode to manage the more Don-Quijote-like aspects of my personality, supportive when it was adequate, always fair and reasonable; for that I am glad. I could not have wished for better supervisors.

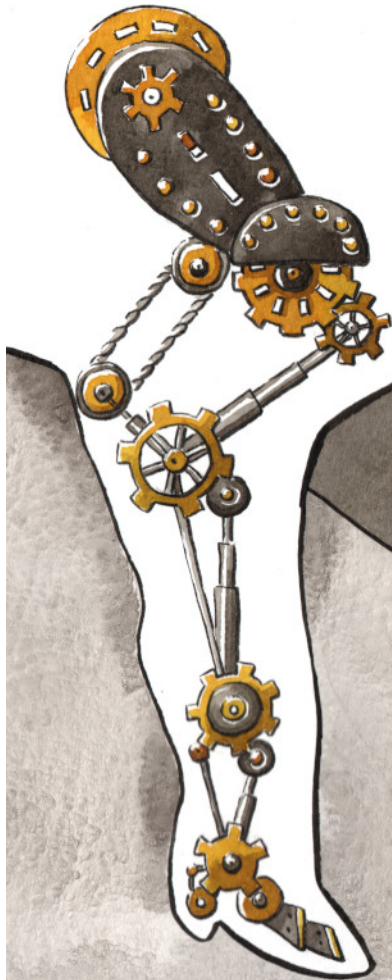
Finally, I am full of love and thankfulness for the support of my family. Maja, Mona, Ronja, my parents, grandparents, (grand)parents in law; all my direct and acquired siblings and their families; with special mention of Alex, Andreza, and Leonardo. I think that the more personal appreciation is best shared in person (another good reason to meet more often). Thank you! I hope to see you all, soon.

Nevertheless, the final words must be dedicated to Maja, who, besides being my wife and friend and lover and the best possible mother for our kids and everything I could ever wish to find in a partner, also jumps in with professional advice on biomechanics, with an ear for my swearing, or just some of her time to keep the little troubles of life at distance. Your contribution to this project is most invaluable, my dear. Danke, dass Du bei mir bist!

$$\ddot{\theta} \approx -\frac{mgL\theta}{I + mL^2}$$

$$\ddot{x} = x \left(\frac{L}{\sqrt{x^2 + y^2}} - 1 \right)$$

$$c_n = \frac{1}{T} \sum_{t=0}^T e^{-2\pi i n \frac{t}{T}} f(t) \quad \forall n \geq 0$$



2019
mm