This extended version of Wernicke et al. (Psychological Methods, 2026) reinstates sections of the paper that were moved to online supplements to meet journal page requirements. The APA journal publication remains the version of record and can be cited as:

Wernicke S, de Falco J, Gironés Z, Jungerius C, Tomić I, Triabhall A, Wang X & Bays PM. Psychological Methods. Advance Online Publication. https://doi.org/10.1037/met0000861

Working With Circular Data: A Tutorial for Cognitive and Behavioral Research

Stella Wernicke[1*], Julie de Falco[1], Zahara Gironés[1,2,3], Chris Jungerius[1], Ivan Tomić[4], Adam Triabhall[1], Xiaolu Wang[1,5], and Paul M. Bays[1]

[1]Department of Psychology, University of Cambridge, United Kingdom
[2]Department of Engineering, University of Cambridge, United Kingdom
[3]Department of Clinical Neuroscience, University of Cambridge, United Kingdom
[4]Department of Psychology, University of Zagreb, Croatia
[5]Department of Psychology, Sun Yat-sen University, China

Abstract Introductory statistics courses and conventional methods in behavioral and cognitive science are typically restricted to Euclidean space without explicitly stating this, leaving many researchers unaware of how to handle data with circular or periodic properties (e.g., time of day or orientation on a screen). This tutorial provides an accessible yet rigorous introduction, designed to bridge the gap between statistical theory and practical application. After an overview of common applications, we discuss different ways of representing circular data mathematically and fundamental principles for working with them. We explain angular and vector representations and operations, including key considerations for implementing them in computer code. We then survey methods for visualizing circular data, including linear and polar plots and presentation of summary statistics. Next, we discuss circular measures of central tendency, dispersion, and shape: how they relate to their Euclidean counterparts and why they differ. Finally, we turn to more advanced topics of the kind needed to construct models of circular data. We introduce the most common circular distribution families, including circular analogs of the Gaussian distribution and their properties, as well as flexible distribution families that can describe skewed and heavy-tailed data. The tutorial takes a method-oriented approach, not tied to specific software, but highlighting factors relevant to experimental design, data analysis, and statistical modeling. A consistent theme is the need to rigorously account for the inherent periodicity of circular data. We also provide a consistent notation for circular statistics and parameters, addressing the lack of established conventions in the field.

Translational Abstract Many areas of behavioral and cognitive research involve circular data, such as periodic temporal data or orientations in space. Yet researchers often apply "conventional" Euclidean statistical methods that can distort conclusions. This tutorial is intended for researchers who are familiar with statistical measures and distributions in Euclidean space and want to understand their circular equivalents. It is also suitable for readers who want to refresh their knowledge of circular statistics or deepen their understanding of more advanced topics related to circular data. We start with an overview of common applications of circular statistics, including spatial, temporal, and neural data that are naturally periodic or can be reconceptualized as such. Next, we introduce different ways of representing circular data and fundamental principles for operating with them, such as arithmetic operations. We then discuss options for visualizing circular data, including the benefits and limitations of different representations. Afterward, we introduce descriptive circular measures of central tendency, spread, and shape of empirical distributions, drawing parallels to the Euclidean analogs where helpful. The topics become progressively more advanced, concluding with an introduction to a range of circular statistical models. In this section, we also discuss circular analogs of the Gaussian (normal) distribution. In contrast, to many previous methodological publications on circular statistics that emphasize statistical tests, this concise and hands-on tutorial focuses on handling circular data in the context of behavioral and cognitive research, experimental design, and statistical modeling. While being software-agnostic, the tutorial highlights common pitfalls to watch out for during practical implementations.

Keywords: Circular statistics; directional data; periodicity; behavior; cognition; experimental design

*Correspondence concerning this article should be addressed to Stella Wernicke via saiw2@cam.ac.uk.

Introduction

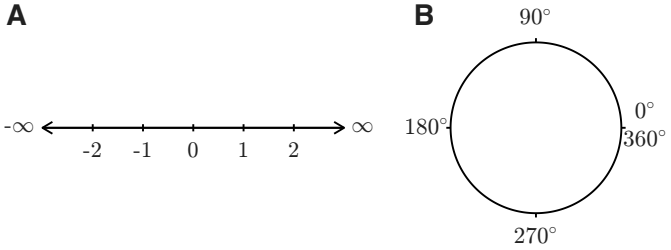
Mathematical training for biological and psychological scientists typically focuses on the analysis of data in Euclidean spaces, particularly measurements that can be described as points on the real number line (referred to as $\mathbb{R}^1$; Figure 1A), or as coordinates in a plane ($\mathbb{R}^2$) or in three-dimensional space ($\mathbb{R}^3$). However, many biological and behavioral variables have a periodic or cyclical nature that requires different methods of analysis. Measurements in a circular space (denoted $\mathbb{S}^1$) can be described by a single number (as on the real line), but moving far enough through the value range in either direction returns you to the point where you started (Figure 1B). Such a space is topologically equivalent to a closed loop or circle and fundamentally different from the real line (which stretches to infinity in either direction) or a bounded interval (a section of the real line with two ends). A simple example of a circular space is the directions of a compass: If you describe directions in terms of angles (e.g., measured clockwise from North), then 0°, 360° and 720° all refer to the same direction, and the distance between 0° and 359° is the same as the distance between 0° and 1°. The topic of circular statistics describes methods for analyzing random variables in circular spaces. It is a subset of directional statistics, which extends to variables that can be represented as points on a sphere ($\mathbb{S}^2$, which includes directions in three-dimensional space), and analogs in higher dimensions. This tutorial focuses on methods for circular data, as they are widely applicable to behavioral research and also provide a foundation for learning about directional statistics more generally. For an introduction to higher-dimensional directional statistics, refer to Mardia and Jupp (2009) and Pewsey and García-Portugués (2021).

This tutorial targets researchers who are familiar with statistical measures and distributions in Euclidean space and want to understand the corresponding circular equivalents. It develops progressively, from the basic properties and operations needed to present and analyze circular data to more advanced concepts, including the foundations for formal generative models that output circular variables.

Conventions for mathematical notation in circular statistics are less standardized than in Euclidean statistics. In cases where there is a strong convention, we have followed it, but for cases where there is not, we have developed a notation (Table 1 and Appendix A) that aims to be clear and self-consistent, although it should not be regarded as canonical. The abbreviations used in this tutorial are defined in Appendix B. Our goal is to give a concise introduction to circular statistics that allows readers to get started with employing them in research involving circular data. For a thorough theoretical introduction to the topic including mathematical derivations and proofs, we recommend Mardia and Jupp (2009). This tutorial will not cover significance tests for circular variables, as there are numerous other sources for these (e.g., Pewsey and García-Portugués, 2021, Chapter 5; Cremers et al., 2021).

Figure 1

Value Ranges in One-Dimensional Euclidean and Circular Space



Note. Comparison of the infinite line of real numbers in one-dimensional Euclidean space (A) with the closed set of angles in circular space (B).

Table 1

List of Symbols for Circular Sample and Population Parameters

Parameter	Sample	Population
Circular mean	$\bar{\theta}$	μ
Circular median	θ_{med}	μ_{med}
Circular mode	θ_{mode}	μ_{mode}
Circular variance	$\widehat{V}_\theta$	V_θ
Circular standard deviation	s_θ	σ_θ
Circular skewness	$g_{1\theta}$	$\gamma_{1\theta}$
Circular kurtosis	$g_{2\theta}$	$\gamma_{2\theta}$
Resultant vector	$\mathbf{R}$	-
p-th circular moment of a population	$\widehat{m}_p$	m_p
Mean resultant vector	$\widehat{m}_1$	m_1
Resultant length	R	-
Mean resultant length	$\bar{R}$	ρ

We will begin with a survey of use cases for circular statistics in behavioral and cognitive research (Section Applications). This section gives an overview of the wide range of applications of circular statistics; the publications cited in this section may also help orient the reader to how circular statistics can be applied to specific topics and types of data. In Section Working With Circular Data, we will introduce key characteristics of circular data, different ways of conceptualizing circular values and describing them numerically, and elementary operations on circular variables. Section Visualizations of Circular Data gives an overview of different options for visualizing circular data and the methodology behind them.

Section Descriptive Statistics of Circular Data will cover the most important statistical measures for circular data, including analogs of mean and variance, and why they differ

from their better-known Euclidean counterparts. We will then explore circular statistical models (Section Probability Theory and Models for Circular Variables), including the vexed question of circular equivalents to the normal distribution, and concepts of importance to constructing ideal observer and Bayesian models. We will conclude with an overview of further topics in circular statistics that fall outside the scope of this tutorial, providing references to the most useful sources (Section Beyond This Tutorial).

Applications

Many areas of behavioral and cognitive research involve data that are either inherently circular or benefit from being conceptualized as circular in order to design experiments or analyze data. To illustrate its broad practical relevance, this section presents use cases of circular statistics, including for spatial, temporal, and neural data.

Spatial and Directional Data

Circular statistics can be a valuable tool for behavioral research using spatial data. Many aspects of spatial cognition and behavior naturally involve circular variables, while in other cases spatial features that are not inherently circular may be constrained to a circular subspace for purposes of experiment or analysis. In the following, we will look at examples of both.

As described in the introduction, compass direction is an inherently circular spatial feature, with animal navigation and migration as a classic use case. Factors supporting navigation include the relative direction of the sun (Giraldo et al., 2018; Lindecke et al., 2019), moonlight and stars (Foster et al., 2019; Freas et al., 2024), and, in some animals, the orientation of magnetic fields (Packmor et al., 2024; Panagopoulos et al., 2024). Circular statistics have been used to study movement directions in migration routes of birds (Bolón et al., 2025; Ghosh et al., 2024; Packmor et al., 2024), insects (Huang et al., 2024; Vijayan & Somanathan, 2024), and mammals (Thompson et al., 2024). Circular statistics have also been employed to study movement patterns on smaller scales, such as bee waggle dances. This complex form of non-verbal, symbolic communication allows forager bees to inform each other about food sources: the angle of the waggle run relative to gravity encodes the angle of the food source in relation to the sun. Circular statistics allow for an appropriate analysis of these complex dances (Deverell, 2003; Schürch & Couvillon, 2013; Vijayan & Somanathan, 2024).

Location in space is an example of a feature that, although not circular by nature, may be structured circularly to support experimental design. In human vision research, for instance, it is common to record responses of experimental participants to visual stimuli presented via a computer display. If stimuli were presented at random locations across the whole display, differences in their retinal eccentricity (distance of the image from the central foveal region of the retina) would influence how they are processed by the visual system (Olivers, 2025; Staugaard et al., 2016). Additionally, the stimuli's varying proximity to the borders of the screen could influence how they are perceived (Fortenbaugh et al., 2012; May & Zhaoping, 2009; Schulz-Hildebrandt, 2025). Confining stimulus locations to an invisible circle on the display, while instructing participants to fix their gaze on a point at its center (see Stimulus Presentation in Figure 2A; Eriksen and Colegate, 1971), is a common experimental technique that reduces variability by ensuring all stimuli are viewed at the same eccentricity and comparable distances from the screen edges. In this design, stimulus location can be incorporated into the behavioral analysis as a circular variable, equal to its angular position on the circle (Favila et al., 2022; Schneegans & Bays, 2017).

The psychophysical method of analog reproduction on a circle, developed by Prinzmetal et al. (1998), was subsequently extended by Wilken and Ma (2004) to study visual working memory – the cognitive resource that retains and manipulates information on the order of seconds. In contrast to previous methods that made a binary measurement of whether or not an object was remembered, their cued recall task (see Figure 2) allowed for assessing how precisely visual information can be recalled, inspiring new models of working memory capacity (Bays et al., 2009; Fougnie et al., 2012; Ma et al., 2014; van den Berg et al., 2012). A wide range of visual features that permit circular representations have been used in this task, including color hue (Wilken & Ma, 2004), motion direction (Ester et al., 2014), and even shape (Salmela et al., 2012).

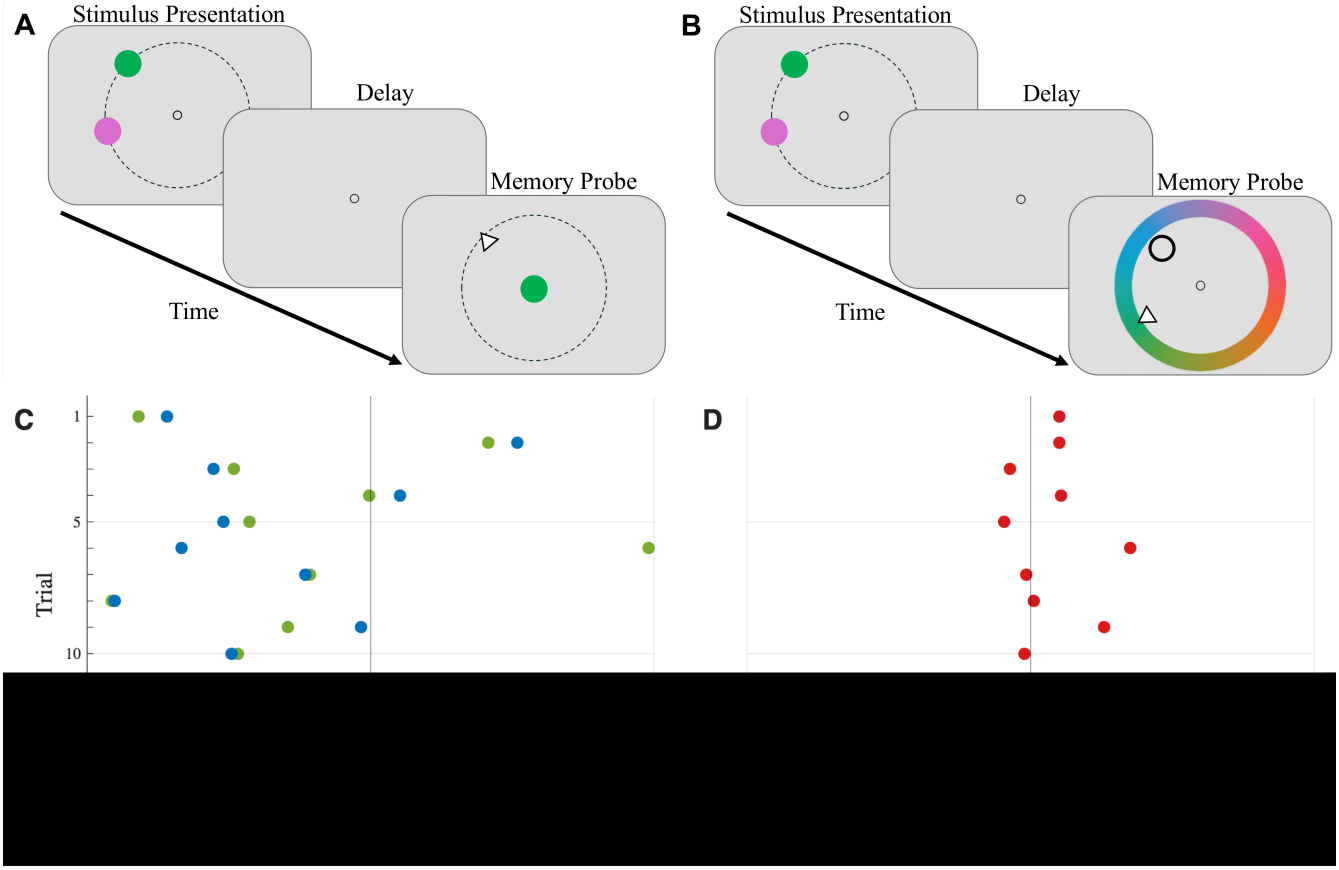
Temporal Data

Human and animal behavior is often structured in recurring patterns, many of which follow approximately or exactly repeating cycles. Some of these cycles have a biological basis (e.g., hormonal rhythms), while others are shaped by environmental factors or, in humans, by cultural conventions (e.g., weekly schedules). Often, biological and cultural factors jointly determine how humans experience time and its influence on behavior (Patke et al., 2020).

A key feature for classifying and distinguishing periodic systems is the length of a single cycle. The most commonly studied chronobiological rhythm is the circadian rhythm (from Latin *circa* — approximately, *dies* — day), which has a period of about 24 hours. Numerous physiological and behavioral processes in humans and other animals vary systematically across the day. Among others, circadian rhythms influence sleep-wake cycles (Saper et al., 2005; Zisapel, 2018), hormone secretion (Nelson et al., 2025; Zisapel, 2018), body temperature (Refinetti, 2020), and activity patterns (Lautenschlager et al., 2024; Palmeirim et al., 2025). In humans, the internal circadian cycle slightly exceeds 24 hours, but external

Figure 2

Circular Data in Visual Memory Tasks



Note. (A, B) Schematic of a visual memory task where participants are presented with colored disks. After a brief delay, one stimulus is either cued by its color (green disk in the center), and participants are asked to recall its location (A), or one stimulus is cued by its location (black circle), and participants are asked to report its color by selecting the corresponding point on the color wheel (B). Note that the dashed circles are only for illustration purposes and are not presented to the participants. (C) Uniformly distributed locations of cued stimuli (green) and corresponding reports from memory of these locations (blue) by trial for the task in (A). (D) Reported locations re-centered on stimulus locations, that is, errors in memory report. All data were simulated.

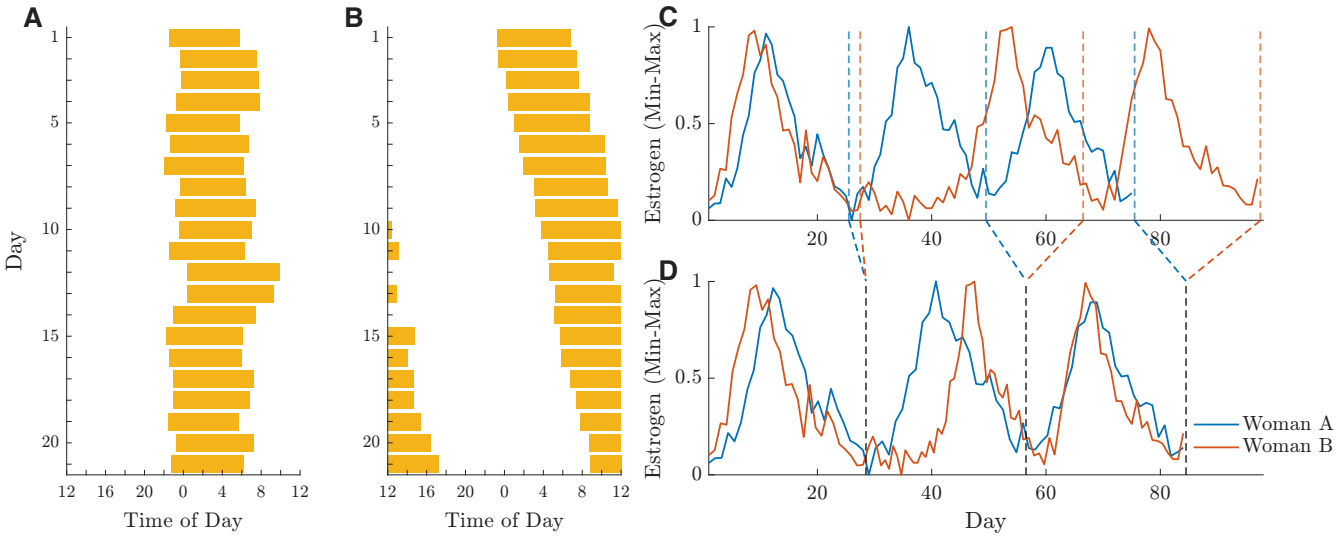
cues – such as daylight and social structures – entrain it to a 24-hour rhythm. Figures 3A and 3B show a schematic of sleep-wake patterns in the presence and absence of time-of-day cues, respectively. In the former, sleep is entrained to a 24-hour rhythm with only random day-to-day variability, whereas in the latter, sleep times progressively drift to later hours in the absence of external cues.

In other periodic systems, cycle lengths vary between individuals or across repetitions, without external constraints enforcing a fixed duration. Even in such cases, circular statistics can be used after an appropriate rescaling of the data. For instance, the human menstrual cycle averages around 28 days but shows considerable inter- and intraindividual variability (Bull et al., 2019; Cunningham et al., 2024). To enable circular analysis, each cycle can be interpolated to a standardized length – typically 28 days (Alzueta et al., 2024; Gombert-Labedens et al., 2024; Sides et al., 2023). Interpolation is conducted by rescaling each cycle's data proportionally to the reference duration, allowing for meaningful comparisons across individuals and repetitions. Figure 3C shows a schematic of estrogen levels of two hypothetical cis-women across three menstrual cycles. Note that the two people display more similar patterns after the data has been normalized (Figure 3D).

Some behavioral rhythms are entirely culturally constructed. For example, the seven-day week, which structures work-rest patterns, primarily originated from religious traditions. Research employing circular statistics has indicated that the weekly cycle influences stress levels (Karoly et al., 2018; Liu & West, 2016) and even the onset of the menstrual cycle (Ecochard et al., 2024). Certain aspects of behavior and cognition also display annual rhythms. This includes hormone levels (Tendler et al., 2021) and brain activity (Meyer et al., 2016) in humans, as well as hibernation and breeding in other animals (Kondo et al., 2006; Paul et al., 2008; Stouffer et al., 2013).

Figure 3

Human Temporal Cycles



Note. (A, B) Schematic of a human sleep cycle with (A) and without (B) daytime cues over three weeks. Yellow bars indicate sleep periods. (C, D) Estrogen (E2) levels across three menstrual cycles of two hypothetical cis-women. (C) E2 levels aligned to the raw timeline of each individual, with cycle boundaries indicated by dashed vertical lines in the respective person's color. (D) Same data as in (C) rescaled so that each menstrual cycle spans 28 days; black dashed vertical lines mark the rescaled cycle boundaries. Note that the estrogen patterns are more similar across individuals and cycles than in (C). All hormone values were normalized to the [0, 1] range within each person (min–max scaling). Rescaling was performed cycle-by-cycle by mapping the day indices of each cycle linearly onto a standardized 28-day axis. Cycle boundaries indicate the first day of menstruation. All data were simulated.

Note that temporal circular data may display autocorrelation structure (e.g., if consecutive cycles are more alike than cycles further apart in time), which violates the independence assumptions of many circular statistical tests.

Neural Data

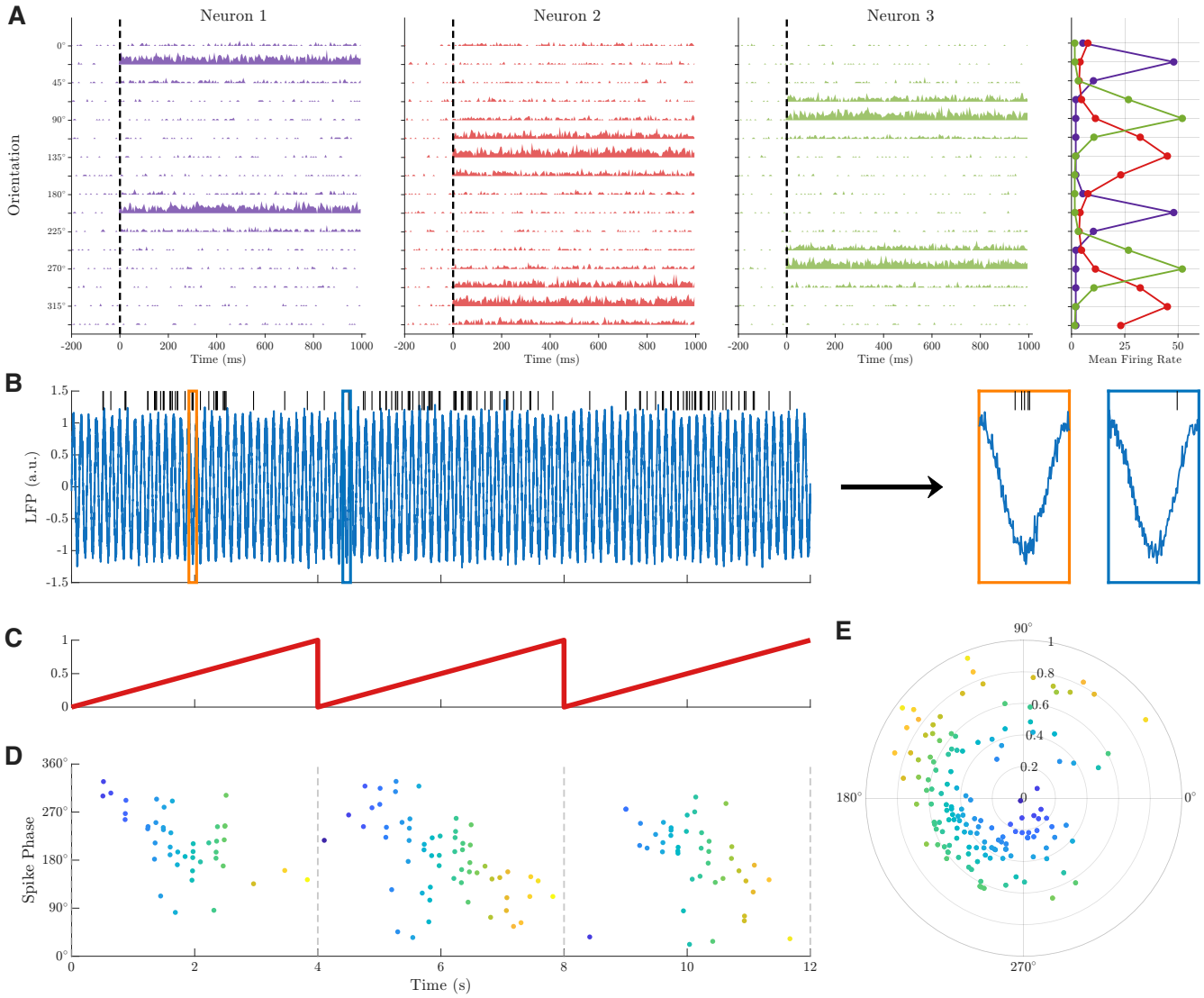
In the previous subsections, we discussed behavioral studies involving spatial and temporal features that are intrinsically circular, or constrained experimentally to be. In the following, we examine how circular statistics have been applied to investigate neural processes related to circular features or that have periodic properties themselves.

The discovery of the columnar organization of the primary visual cortex (V1) by Hubel and Wiesel (1962) was a major advance in sensory neuroscience, recognized by a Nobel Prize. Through electrophysiological recordings in V1 in cats and monkeys, they identified neurons that were selectively responsive to specific orientations of a visible bar or grating presented on a flat screen. Each cell had a "preferred" orientation that caused it to be most active (Figure 4A), and neurons with similar orientation preferences were grouped into vertical columns (Hubel & Wiesel, 1962, 1968; Hubel et al., 1977). These orientation columns form orderly, repeating patterns across the cortical surface, often described as "pinwheels", because preferred orientations of neurons rotate continuously around a central point, covering the full range of orientations (Blasdel & Salama, 1986; Bonhoeffer & Grinvald, 1991; Bonhoeffer et al., 1995; Maldonado et al., 1997; Ts'o et al., 1990). While Hubel and Wiesel and early follow-up studies were primarily descriptive or employed Euclidean statistics, later work recognized the circular nature of orientation data and used circular statistical methods (Blasdel, 1992; Maldonado et al., 1997; Olson & Grossberg, 1998; Ringach et al., 2002). More recent visual-processing research has applied circular statistics to investigate how many neurons are required to represent continuous spatial variables reliably (Noorman et al., 2024), to explore the neural mechanisms underlying visual illusions (L.-Q. Zhang et al., 2025), and to model how different visual features of an object could be jointly represented in neural activity (Schneegans & Bays, 2017).

The neural control of movement is another area where analysis of circular variables has been fruitfully employed. Georgopoulos et al. (1983, 1984) used circular statistics to study firing patterns of motor cortical neurons. They investigated the activity of neurons in the primary motor cortex (M1) of monkeys making reaching movements in the horizontal plane. They found neurons whose firing rate varied as an approximately cosine function of movement direction, each with a preferred direction that induced maximum activity. By treating each neuron's instantaneous firing rate as the length of a vector pointing in its preferred direction, and summing across all neurons, they obtained a "population" vector whose direction closely matched the actual reach direction on that trial. Re-conceptualizing the motor cortex as a network that encodes higher-order features (rather than a simple

Figure 4

Neural Processes With Circular Features



Note. (A) Simulated response properties of single units recorded at the center of a pinwheel in the primary visual cortex of a hypothetical cat. The three left panels show peristimulus time histograms (PSTH) from three neurons that were stimulated by visual display of an oriented drifting square wave grating. Each row shows the PSTH for a different stimulus orientation ranging from 0° to 337.5° in 22.5° steps. Recordings are presented from 200 ms before stimulus display and 1000 ms of stimulus display, accumulated over 20 trials for each orientation. The right panel shows the mean firing rate in spikes per second for each cell and stimulus orientation. Note that the neurons vary in their orientation preferences. (B–E) Theta phase precession in place cells. (B) Simulated local field potential (LFP, 8 Hz theta rhythm) in arbitrary units (a.u.) as a hypothetical rat repeatedly traverses a spatial field. Black ticks indicate spike times from one hippocampal place cell. The two framed theta phases are enlarged on the right, illustrating how spike timings and frequencies vary for different positions in the place field. (C) Normalized position of the rat within the place field. (D) Theta phase of each spike over time. Each dot shows a single spike, color-coded by its within-field position. Note that the spike phase advances systematically within each pass. (E) Polar representation of the data in (C) with spike phases and normalized positions within the place field represented by angles and radii, respectively. Data were simulated using methods adopted from Kempter et al. (2012).

map of individual muscles or body parts), Georgopoulos and colleagues laid the foundation for our current understanding of how motor cortical neurons represent abstract aspects of movement (Drew & Doucet, 1991; Georgopoulos et al., 1986; Perkins et al., 2025).

Circular statistics are also a key methodological tool in neuroscience research on spatial navigation (Moore et al., 2021; Nau et al., 2020). For instance, Seeber et al. (2025) combined brain imaging methods and circular statistics to compare real and imagined navigation in humans. Participants first navigated real-world routes and later imagined retracing them. Neurons in the hippocampus (a brain region involved in spatial navigation) showed similar directional tuning patterns in both conditions, with different neurons responding most strongly to different movement directions, and movement directions could be reconstructed from brain activity during

both real and imagined navigation.

Analysis of oscillatory neural activity ("brainwaves") is a common use case for circular statistics. A striking example of periodic neural representation of linear space is phase precession (Figure 4B–4E), first observed in single-unit recordings referenced to the hippocampal theta local field potential (LFP) in rats navigating in an enclosure. In this phenomenon, certain neurons in the hippocampus (so-called place cells) fire at times progressively earlier in the theta wave (~4–12 Hz) cycle as the rat moves through the cell's preferred location (its place field; O'Keefe & Recce, 1993; Skaggs et al., 1996). For analyzing phase-precession data, a combination of a circular representation of the periodic phase variable and a Euclidean representation of position in space has proven reliable (Kempter et al., 2012), and subsequently provided evidence for phase precession in human episodic memory (Doeller et al., 2010; Zheng et al., 2024).

Other Applications

Most examples we discussed above reflect physical measurements of time or space that have a circular or periodic structure. However, there are also more abstract perceptual, cognitive, and behavioral variables for which a circular representation is appropriate or has proven useful. For instance, circular statistics have been employed in social psychology to examine interpersonal interactions. This includes the synchronization of body movements to those of an interaction partner (Koul et al., 2023; Macpherson et al., 2024), interactions in romantic relationships (Sadikaj et al., 2015; Vowels et al., 2018), and interpersonal behavior of teachers (Cremers et al., 2021; van der Lans et al., 2020). The Interpersonal Circumplex (IPC) is an established model of interpersonal relationships that describes social behavior and traits along two orthogonal dimensions: agency and communion, referring to the degree of power or control and the degree of friendliness or affiliation an individual conveys in interaction with others, respectively. Together, these two dimensions form a circular space in which interpersonal behaviors are located as angles on a circle (Gurtman, 2009; Leary, 1957; Wiggins, 1991). Figure 5A displays the IPC with anchor words describing the type of behavior located in each part of the interpersonal circle. While agency and communion are often analyzed separately, Cremers et al. (2021) argue that a circular representation integrating both dimensions more appropriately captures the angular structure of the IPC, where interpersonal behaviors are expressed as positions on a circle reflecting blends of agency and communion. The identical agency and communion scores are represented in two-dimensional Euclidean space in Figure 5B and as one circular variable in Figure 5A.

Another example is in human perception of color. Color is detected in the retina by three types of cone cells which respond preferentially to different wavelengths of light. However, in perceptual models of color, such as CIELAB (Schanda & International Commission on Illumination, 2007), different hues are arranged in a circle around a grayscale axis. Color wheels obtained as circles of constant luminance and chroma (distance from the gray axis) have been used in human psychophysical research to account for the continuous nature of perceptual and mnemonic data (Fairchild, 2013). To obtain a continuous measure of recall from visual memory, for instance, participants can be asked to select the color of a previously presented stimulus from a color wheel (Figure 2B; Wilken and Ma, 2004). Operationalizing color as a circular variable simplifies analysis, as angular error of reproduction on the color wheel can be treated as independent of the target color, whereas a linear representation in a bounded space would have to take into account the proximity of the target color to the boundary of the space.

Working With Circular Data

In this section, we describe the fundamental characteristics of circular data to provide the foundation for subsequent sections. We explain the advantages of describing circular values as vectors of unit length in a two-dimensional plane and three standard forms of this vector representation: in polar coordinates, Cartesian coordinates, and on the complex plane. There is no universally correct choice for representing a given circular dataset; the optimal representation depends on the context and purpose of the analysis. For instance, angular measurements are often more intuitive for data visualization. If the methods described here are correctly applied, conclusions drawn from circular data will remain valid regardless of the chosen representation.

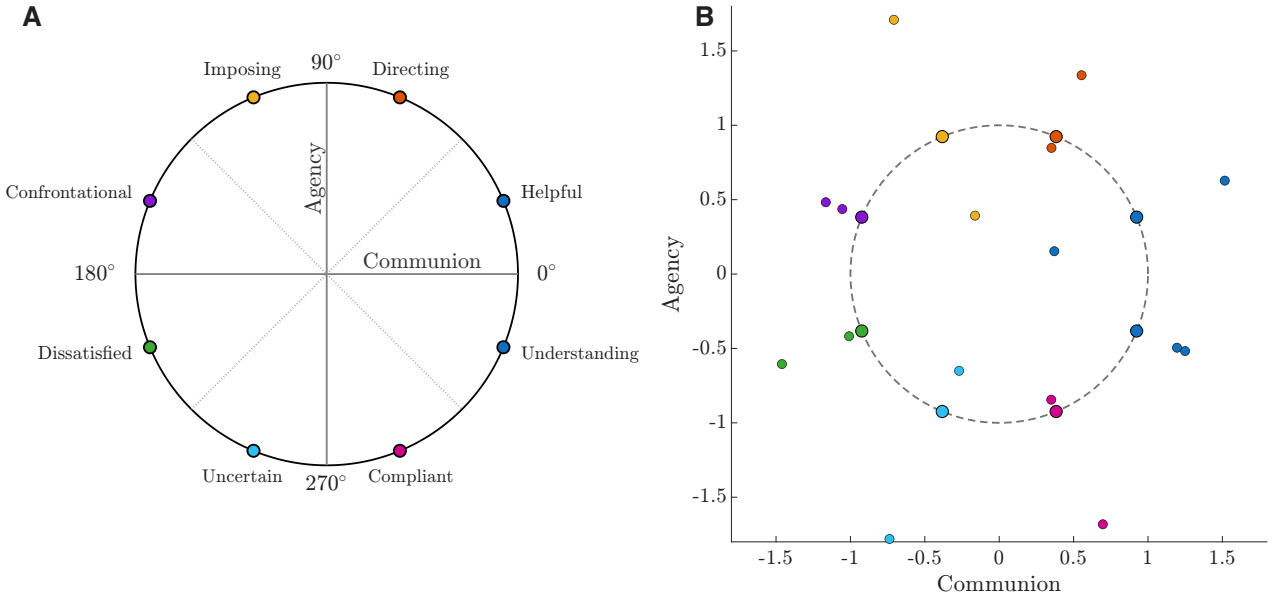
Identifying the Periodicity of Data

Finding the appropriate cycle length requires careful consideration, as an unsuitable choice might bias results. One cycle needs to contain every distinct point in the space of possible data values exactly once. Points that are separated by a whole number of cycles must be identical or at least equivalent for the purpose of the analysis. Selecting too narrow a period risks obscuring important aspects of the data by treating data points as equivalent when they are not: For instance, if temporal data are analyzed on a daily level, but they actually displays weekly rhythms. On the other hand, selecting too wide a period will mean the periodic nature of the data is not properly accounted for. Liu and West (2016) discuss a method for identifying an appropriate cycle length for circular analysis of temporal data.

Particular care is needed when working with data that in their raw form are angular or directional, because the appropriate period for circular analysis may not correspond to 360° in the units of the data. If we rotate the black shape in Figure 6A around the black dot in its center, we will need to rotate 360° before it looks the same again. This stimulus

Figure 5

Interpersonal Circumplex Model in Circular and Euclidean Space



Note. (A) Interpersonal Circumplex Model in circular space with anchor points (colored dots). (B) Euclidean representation of the same points as in A (large colored dots) and alternative points (small dots) that would result in the same points when constraining the data to the circle. The dashed line in (B) indicates the unit circle to which values are constrained in (A). Note that the model conceptualizes social behavior and traits continuously; the anchor points only provide orientation for interpreting values in different segments of the circle.

therefore has no (or onefold) *rotational symmetry*, and the circular space of its orientations has a period of 360°. The shape in Figure 6B, in contrast, will first look identical after being rotated through 180°, and again every subsequent 180° of rotation. This shape therefore displays twofold rotational symmetry. Analogously, Figures 6C and 6D show shapes with three- and fourfold symmetry. For shapes with twofold or higher rotational symmetry, the full space of possible orientations covers less than 360°. It is this smaller period that must be mapped to the full circle for the purposes of circular analysis, and as this will typically be measured in angular units, it is important to keep clear the distinction between the angles of physical orientation and the angles in the circular space of possible values. To illustrate, Figure 6E shows the orientations corresponding to the cardinal points in the circular space for an object with onefold rotational symmetry (top) and twofold rotational symmetry (bottom). Further care is needed in reporting the results of circular analysis of such data to make clear whether they are in the units of the circular space or of the original data. Figures 6F–6I show examples of common experimental stimuli used in vision research with one-, two-, three-, and fourfold symmetry, respectively.

Representations of Circular Data

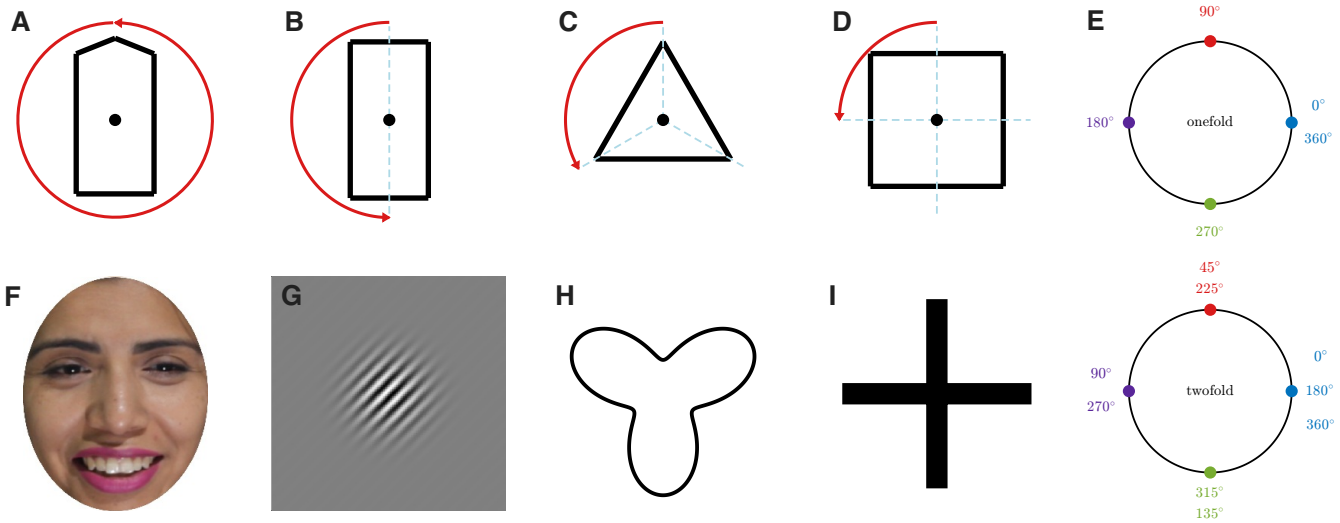
Having identified the periodicity of our circular data, we can map each observation to its corresponding angle or *phase*, indicating where it falls within one full period or cycle. This requires us to select a data value that will correspond to an angle of 0, and a direction in the data space that will correspond to the positive direction of increasing angle. In some cases these choices are obvious (e.g., for temporal data with a period of one day, one would usually assign 12 o'clock midnight to the zero angle and increasing positive angular values to later times in the day), but in others they may be arbitrary, or the researcher may have to choose between several viable alternatives.

Angles are typically measured in degrees (°), in which case one cycle covers a range of 360°, or radians (rad), in which case (for reasons explained below) one cycle covers 2π ($\approx$ 6.28) rad. As discussed above, for data consisting of directions or orientations in space, the angular units of the circular space may differ from those of the raw data. The units in which angular values are reported are a matter of choice, subject to conventions of the field, but some mathematical operations are simplified by using radians. To convert an angle in degrees to radians, multiply by the ratio $\frac{\pi}{180}$, and divide by the same ratio to convert the other way. Lowercase Greek letters, typically θ (theta), ϕ (phi), or ψ (psi), are often used for angular variables in a similar way to x or y for Euclidean variables. Similarly, uppercase Greek letters – Θ (Theta), Φ (Phi), or Ψ (Psi) – are often used for sets or collections of elements analogous to X or Y in Euclidean space.

Representing circular data as angles provides a common framework in which to describe and work with different kinds

Figure 6

Rotational Symmetries in Circular Space



Note. (A–D) Abstract stimuli with no or onefold (A), twofold (B), threefold (C), and fourfold (D) rotational symmetry. Black dots show the points around which the stimuli can be rotated; symmetry axes are perpendicular to the page passing through these points. Red arrows indicate one complete rotation in the space of possibilities. (E) Cardinal angles in the circular space of possible orientations marked with corresponding physical rotations in degrees for shapes with onefold (top) and twofold (bottom) rotational symmetry. (F–I) Sample stimuli used in perceptual and memory research. (F) Oval-cropped face with onefold (i.e., no) rotational symmetry. Image from Mitre-Ortiz and Mitre-Hernandez (2019). (G) Gabor patch with twofold symmetry. (H) Radial frequency pattern with a frequency of three cycles per 360° (threefold symmetry) and an amplitude of 50%. Stimuli with different amplitudes can be used to manipulate perceptual and mnemonic discriminability. (I) Fixation cross used for gaze stabilization, with fourfold rotational symmetry.

of periodic data. However, for many of the calculations we will want to perform on circular data, a vector representation (i.e., one combining angle and distance) turns out to be even more powerful. There are three mathematically equivalent ways to represent such a vector, summarized in Figure 7, and each offers distinct advantages depending on the context of application.

Polar Coordinates

Representing a circular value in polar coordinates simply involves supplementing the angular representation of data, θ, with a distance or *norm*, r, resulting in the polar coordinate pair (r, θ). By convention the norm is set to 1, giving $(1, \theta)$. Now each observation is identified not just with a direction but also with a specific point in the two-dimensional plane, and the space of all possible observations is described by a circle of unit radius. Technically, this is an *embedding* of the circular space ($\mathbb{S}^1$) in the real plane ($\mathbb{R}^2$). We can now appreciate the basis for radian units: an angle of X rad corresponds to a distance X measured on the circumference of the circle, with a full period corresponding to the full circumference of 2π. The polar vector representation is illustrated in the left panel of Figure 7.

Cartesian Coordinates

We can also describe circular observations as vectors in a Cartesian coordinate system, with origin O and two perpendicular axes, X and Y, passing through O. By convention, the 0° direction is positive on the x-axis. In this reference frame, any point on the plane can be represented as a pair of rectangular coordinates $[x, y]$, with unit vectors constrained to have $x^2 + y^2 = 1$. Compared with a polar representation, this simplifies vector operations (which are used to define statistics like the circular mean; see Section Descriptive Statistics of Circular Data). For example, addition of two vectors simply involves pairwise addition of their Cartesian coordinates. On the other hand, it makes angular operations, such as adding two angles together, more difficult. The Cartesian vector representation is the most easily extended to higher-dimensional directional statistics. For example, spherical data can be represented by unit vectors in three dimensions, described in $[x, y, z]$ coordinates.

To convert from polar coordinates to Cartesian coordinates, we use the trigonometric functions sine and cosine. For a vector with polar representation (r, θ), the corresponding Cartesian coordinates are

$$[x, y] = [r\cos(\theta), r\sin(\theta)] \tag{1}$$

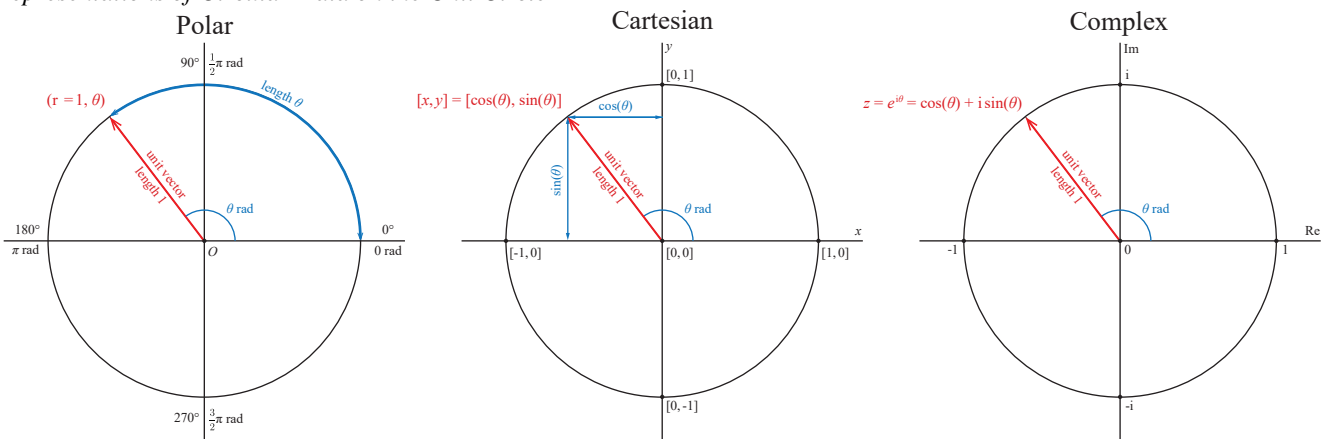
and for the inverse,

$$(r, \theta) = (\sqrt{x^2 + y^2}, \text{atan2}(y, x)) \tag{2}$$

where atan2 is the two-argument arctangent function. These trigonometric functions are available in most programming languages and typically default to using radians as angular units, with atan2 usually returning angles in the range $-\pi < \theta \leq \pi$ rad. The Cartesian vector representation is illustrated in the middle panel of Figure 7.

Figure 7

Representations of Circular Data on the Unit Circle



Note. Three mathematically equivalent vector representations of the same circular value θ: the polar vector representation (left), the Cartesian vector representation (middle), and the complex vector representation (right). Representations are shown on the unit circle with radius 1. For the polar representation, cardinal angles are annotated in radians (rad) and degrees (°). Note that it is a mathematical convention to plot circular data with the zero reference point on the right side of the horizontal axis and a counterclockwise sense of rotation. The angle of θ rad corresponds to an arc length θ on the circumference of the circle. The circumference of the circle is 2π, and a full rotation corresponds to 2π rad and 360°. In both Cartesian and complex space, a unit vector (or point on the unit circle) at angle θ has horizontal and vertical components sin θ and cos θ, respectively.

Complex Coordinates

Finally, each data point can be represented by a unit complex number, z, interpreted as a vector in the complex plane ($\mathbb{C}^1$). The angle θ, measured in radians, and the unit complex number z are related by Euler's formula:

$$z = e^{\mathrm{i}\theta} = \cos\theta + \mathrm{i}\sin\theta \tag{3}$$

where $\mathrm{i} = \sqrt{-1}$ is the imaginary unit. Note that we write the imaginary unit as i denoted in upright Roman to distinguish it from the indexing variables i denoted in italics within the later equations (e.g., summing from $i = 1$ to n). While it may seem esoteric, the complex vector representation of circular observations allows for particularly compact descriptions of both angular and vector operations, as will become clear in subsequent sections.

The complex plane is usually depicted with the horizontal axis as the real component, whereas the vertical axis represents the imaginary component. The unit complex number $e^{\mathrm{i}\theta}$ evaluates to 1 at $\theta = 0$, then moves into the upper half-plane as θ grows, tracing the unit circle counterclockwise. This is the basis of the mathematical convention of plotting circular data with the zero reference point on the right side of the horizontal axis and a counterclockwise sense of rotation (see Figure 7), a convention we follow in this tutorial. Nonetheless, many fields deviate from this standard, adopting domain-specific conventions to facilitate data interpretation (e.g., in meteorology, with North = 0° plotted up and clockwise rotation). A Cartesian coordinate $[x, y]$ corresponds to a complex coordinate $z = x + \mathrm{i}y$. A polar coordinate (r, θ) can be converted to a complex coordinate by either $z = re^{\mathrm{i}\theta}$ or equivalently $z = r(\cos\theta + \mathrm{i}\sin\theta)$. The complex vector representation is illustrated in the right panel of Figure 7.

Periodicity, Wrapping, and the Modulo Operation

As described above, the defining characteristic of circular data is its inherent periodicity, meaning that measurements that differ by a whole number of cycles are considered equivalent. This property fundamentally distinguishes circular from Euclidean space, where proximity and ordering follow conventional numerical relationships. For example, in Euclidean space, 1 is always closer to 3 than 10. In circular spaces, values that may appear numerically distant can in fact be close, for example, 1° and 359° differ by only 2° because 1° and 361° are equivalent angles.

While conceptually any observation in a circular space can be represented by infinitely many angular values, differing by whole numbers of cycles, for practical purposes, it is usually constrained to one unique value within a specific periodic interval. According to the mathematical convention, angular values are constrained to a positive interval $[0, 2\pi)$ rad. Note that the square bracket indicates the inclusion of the lower bound and the round bracket the exclusion of the upper bound. This specification ensures that the same point is not included twice, which would occur if both bounds were closed, that is, with this definition, an observation can be assigned to 0 but not 2π.

The term "wrapping" is often used for the process of mapping circular observations to their representative values within a specified interval (imagine wrapping the real number line around the unit circle, like winding thread onto a bobbin). The process of wrapping angle θ to the interval $[0, 2\pi)$ can

be described by a modulo operation,

$$\theta_{\text{wrapped}} = \theta \bmod 2\pi, \tag{4}$$

where $a \bmod b$ is the *least nonnegative residue of a modulo* b, that is, the smallest number ≥ 0 that can be obtained by subtracting an integer multiple of b from a.

An alternative convention is to constrain angular values to the symmetric signed interval $[-\pi, \pi)$, centered on 0. For this, the wrapping operation is

$$\theta_{\text{wrapped}} = ((\theta + \pi) \bmod 2\pi) - \pi. \tag{5}$$

The same operations can be used to convert angles constrained to one of these conventional intervals into the other. Importantly, these operations preserve the circular topology so that individual observations are not shifted relative to one another; only their numerical representation is changed.

Using the signed interval $[-\pi, \pi)$ can facilitate data interpretation and visualization when investigating signed deviations or errors relative to a reference angle (e.g., Figure 2D), or other observations that tend to cluster around 0. It is also the convention for certain sources of angular data. For example, in sound localization tasks, azimuthal (horizontal) directions are represented with 0 corresponding to the listener's frontal midline and positive values indicating deviations to the listener's right-hand side.

Most programming languages have a built-in modulo function, but, importantly, the implementation for negative inputs frequently differs from the least nonnegative residue assumed above. As of the software versions available at the time of writing (MATLAB R2026a, R-4.6.1, Python 3.14.7, NumPy 2.5.2, ES2026, C++23), MATLAB's `mod(a,b)`, R's `a%%b`, and Python's `a%b`, or `numpy.mod(a,b)` all return a result with the same sign as the divisor `b`. This is acceptable for wrapping operations, because the divisor is the length of the desired interval (e.g., 2π) and so positive. In contrast, JavaScript's `a%b` (ES2025), C++'s `a%b` and `std::fmod(a,b)` (C++23), and Python's `math.fmod(a,b)` return an output equivalent to $-((-a) \bmod b)$ when a is negative. One solution is to define your own `mod` function based on the `floor` function, for example, in JavaScript:

```
function mod(a,b) { return(a -
    b*Math.floor(a/b)) };
```

Alternatively, you can use the following sign-safe versions of Equations 4 and 5, which produce acceptable results across programming languages:

$$\theta_{\text{wrapped}} = ((\theta \bmod 2\pi) + 2\pi) \bmod 2\pi \tag{6}$$
$$\implies \theta_{\text{wrapped}} \in [0, 2\pi)$$
$$\theta_{\text{wrapped}} = ((\theta \bmod 2\pi) + 3\pi) \bmod 2\pi - \pi \tag{7}$$
$$\implies \theta_{\text{wrapped}} \in [-\pi, \pi)$$

Operations on Angles

When working with Euclidean variables, there exist many operations that map real numbers to other real numbers. By contrast, for circular variables, the need to respect periodic equivalence means there are very few operations that map angles to other angles: Essentially, these are limited to addition, subtraction, and multiplication by integer quantities (equivalent to repeated addition or subtraction). While it is possible to perform other operations on numerical representations of angles, the results will depend on the particular interval you have chosen for representation. For instance, 0° and 360° refer to the same angle, but if you attempt to divide them by 2 you will get 0° and 180°, which are different angles.

Having specified a range of angles in which to represent circular data, we will want to represent the results of angular addition and subtraction within the same interval. Practically, this just means applying a "wrapping" or modulo operation, as described above, after the usual arithmetic operation of addition or subtraction. A familiar example of modular arithmetic is the hour hand on a clock dial. If we subtract 2 hours from 1 o'clock, we get 11 o'clock, not −1 o'clock. Similarly, if we add 5 hours to 10 o'clock, we get 3 o'clock and not 15 o'clock.

The modular addition of two angles θ and ϕ in the standard interval $[0, 2\pi)$ is given by

$$(\theta + \phi) \bmod 2\pi \tag{8}$$

and modular subtraction by

$$(\theta - \phi) \bmod 2\pi \tag{9}$$

The modulo operation ensures that the result always lies in $[0, 2\pi)$, regardless of the values of θ and ϕ. Modular subtraction can be used to recenter data on a particular reference direction, or compute the angular deviation or error of an estimate θ from a true value ϕ. In these cases, it may be preferable to wrap the result to the signed interval $[-\pi, \pi)$ using Equation 5.

Importantly, because a modulo operation changes only the numerical representation of a circular value, it can be applied at every step of an arithmetic calculation or only to the final output without changing the result, for example,

$$((\theta + \phi) \bmod 2\pi) + \psi) \bmod 2\pi = (\theta + \phi + \psi) \bmod 2\pi. \tag{10}$$

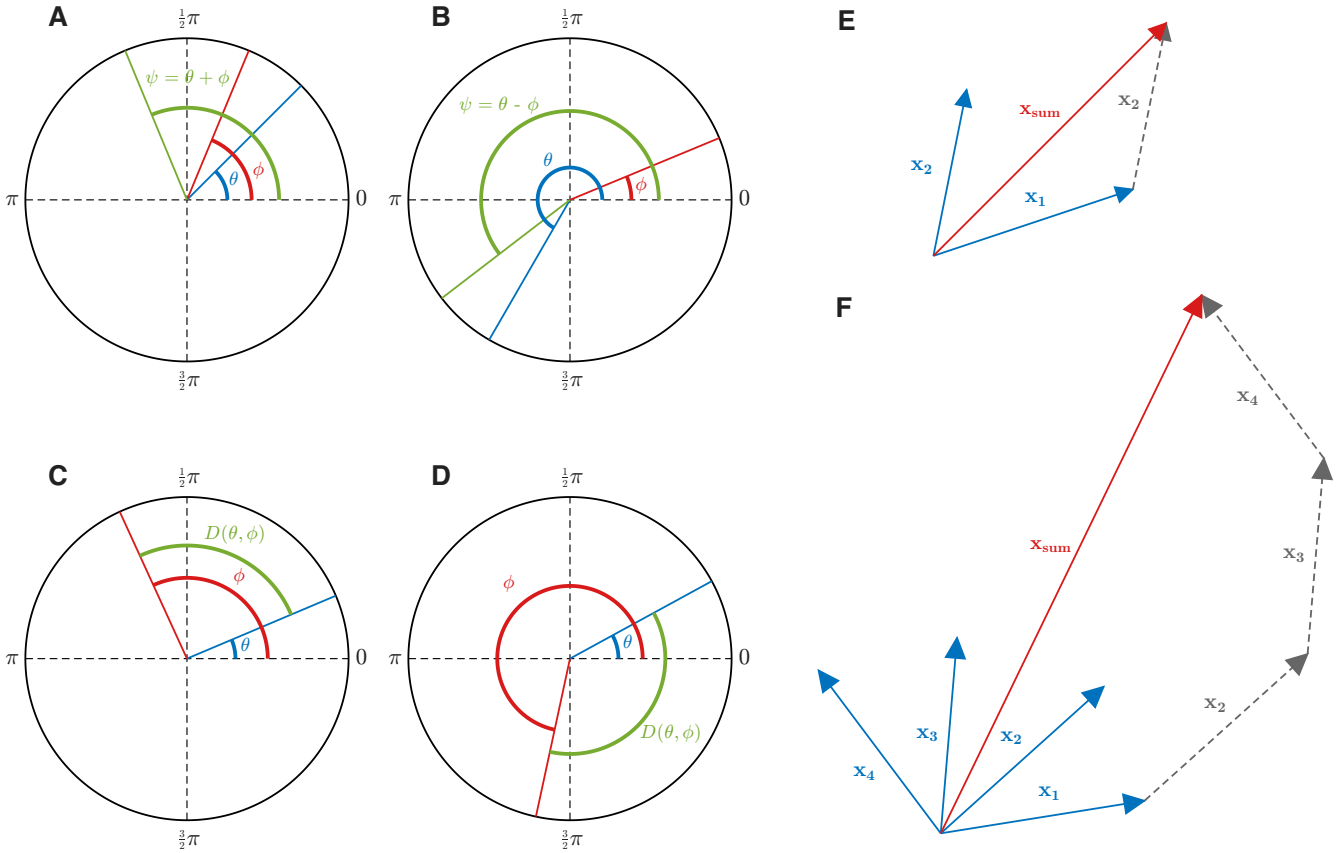
The symbols $\oplus$ and $\ominus$ are sometimes used for modular addition and subtraction, but this is not an established standard, so their intended meaning should always be explicitly stated, including the interval in which their outputs are to be represented.

Vector Equivalents of Angular Arithmetic

The modular arithmetic formulation is useful for working directly with angles. It is also possible, and often preferable,

Figure 8

Angular Arithmetic and Vector Summation



Note. (A) Modular addition (as defined in Equation 8) of angles θ (blue) and ϕ (red) yielding sum ψ (green). (B) Modular subtraction (as defined in Equation 9) of angle ϕ (red) from angle θ (blue) yielding difference ψ (green). (C, D) Angular distance $D(\theta, \phi)$ (green) between angles θ (blue) and ϕ (red). Although any pair of two angles can be connected by rotation in either direction around the circle, counterclockwise and clockwise rotation lead to the smaller angle in (C) and (D), respectively. (E, F) Vector addition (as defined in Equation 16) of two (E) or multiple (F) vectors. Blue and red arrows show the original vectors $\mathbf{x_i}$ and summed vectors $\mathbf{x_{sum}}$, respectively. Gray dashed arrows show the tip-to-tail visualization where summing vectors can be viewed as connecting individual vectors sequentially.

to implement angular addition and subtraction in a vector representation of circular data. As shown in Figure 8A, in the standard polar coordinate system, adding an angle ϕ to an angle θ corresponds to a counterclockwise (positive) rotation of the unit vector corresponding to θ. Conversely, subtracting an angle ϕ from θ corresponds to a clockwise (negative) rotation (Figure 8B).

In Cartesian coordinates, rotation of vectors is achieved by multiplying by a *rotation matrix*. A rotation of a vector with coordinates $[x, y]$ through an angle ϕ about the origin O can be expressed in matrix form as:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{11}$$

The rotation is counterclockwise in standard coordinates if ϕ is positive, and clockwise if it is negative. If $[x, y]$ is a unit vector corresponding to angle θ, then $[x', y']$ is another unit vector corresponding to angle $\theta + \phi$. Rotation matrices are a relatively cumbersome way of implementing angular arithmetic, but they find considerable application in computer graphics, including 3D graphics where rotations in three dimensions are expressed analogously as 3×3 matrices. Importantly, in computer graphics software, pixel locations on a two-dimensional display are usually represented in a nonstandard "left-handed" Cartesian system, with y increasing downward instead of upward, x increasing to the right, and the origin at the top-left corner of the display. In this system, all coordinates lie in the first quadrant where $x \geq 0$ and $y \geq 0$. The inversion of the y-axis reverses the sense of rotation, so for screen coordinates $[x, y]$, the rotation described by Equation 11 is clockwise for positive θ.

For circular values represented as unit vectors in the complex plane, the addition and subtraction of angles correspond to multiplication and division, respectively. That is, if z_1 and z_2 are two unit complex numbers corresponding to the angles θ and ϕ, then their product is

$$z_1 \cdot z_2 = e^{i\theta} \cdot e^{i\phi} = e^{i(\theta+\phi)}, \tag{12}$$

yielding a new unit vector representing the angle $\theta + \phi$. Similarly, the subtraction $\theta - \phi$ corresponds to dividing the vector representation of θ by the vector representation of ϕ:

$$\frac{z_1}{z_2} = \frac{e^{\mathrm{i}\theta}}{e^{\mathrm{i}\phi}} = e^{\mathrm{i}(\theta-\phi)} \tag{13}$$

The complex representation therefore provides a simple and mathematically elegant implementation of angular arithmetic, in which the "wrapping around" of large angle changes is taken care of automatically by the geometry of the complex plane. Many programming languages have support for complex numbers built-in (e.g., in both MATLAB and Python you can create complex numbers by typing `1j` for the imaginary unit i and they will work with all the usual math operators and functions), whereas other languages have well-established libraries that implement them (e.g., math.js for JavaScript). Note, however, that the complex representation of circular variables does not have convenient extensions to spherical or higher-dimensional directional variables: The polar and Cartesian representations provide better starting points for working with those kinds of data.

Measuring Distance

The periodic equivalence of circular data means careful consideration must be given to the measurement of distances. The Euclidean distance between two points on the real number line is just the absolute value of their difference, calculated equivalently as either $|x - y|$ or $|y - x|$. By contrast, given two angles θ and ϕ corresponding to two points A and B on the unit circle, their angular separation can be measured either clockwise or counterclockwise, yielding two different distances that add up to one complete cycle. The *angular distance* is typically taken as the smaller of the two, equivalent to the shortest arc length between A and B (Figures 8C and 8D). This can be calculated as

$$\begin{aligned} D(\theta, \phi) &= \min(|\theta - \phi|, 2\pi - |\theta - \phi|) \\ &= \pi - |\pi - |\theta - \phi||. \end{aligned} \tag{14}$$

where the circular variables θ and ϕ must be represented in the same interval of range 2π, for example, $[0, 2\pi)$ or $[-\pi, \pi)$. The resulting distance always lies in $[0, \pi]$ and it captures the magnitude of the angular separation, without regard to its direction. Note that the angular distance is not itself a circular variable.

An alternative measure of separation between circular variables, and one that is in a sense more "natively" circular, is the *cosine dissimilarity* (also *cosine error*), defined as

$$D_{\cos}(\theta, \phi) = 1 - \cos(\theta - \phi) = 1 - \cos(\phi - \theta). \tag{15}$$

The cosine dissimilarity lies in the range $[0, 2]$, with 0 and 2 corresponding to the minimum and maximum possible separation of 0 and π rad, respectively. This measure of distance can be conceptualized as a circular analogue of the squared Euclidean distance and plays a similarly important role in circular statistics, including in the definition of circular variance, as we will discuss in Section Circular Variance. Indeed, with respect to the vector representation of angles, the cosine dissimilarity is equal to half the squared straight-line distance between points on the unit circle, that is, $\frac{1}{2}||\mathbf{A} - \mathbf{B}||^2$ in the Cartesian representation or $\frac{1}{2}|z_1 - z_2|^2$ in the complex representation. For unit vectors with small differences in angle δ, the straight-line distance and the angular distance are approximately equal, so the cosine dissimilarity $1 - \cos(\delta)$ converges to $\frac{\delta^2}{2}$. Note that, like angular distance, the cosine dissimilarity is not a circular variable.

In study design, circular distance becomes relevant when a minimum or maximum separation between angular values must be ensured. Such constraints allow continuous rather than discrete variables while keeping points neither too close nor too far apart. For instance, experimental stimuli could be sampled with a minimum circular distance so they remain perceptually distinct, or ambulatory-assessment surveys scheduled at random times subject to a maximum gap between them. A method for sampling from a uniform distribution in circular space while ensuring a minimum separation between values is provided in Appendix C.

Vector Summation

We described above how addition and subtraction of angles can be represented as rotation of unit vectors in the real or complex plane. In subsequent sections we will see that vector representations provide a general foundation for the analysis of circular data. This includes the definition of circular analogues of mean and variance, which are based on vector summation.

The vector sum of two unit vectors, $\mathbf{x_1}$ and $\mathbf{x_2}$, is illustrated geometrically in Figure 8E, with vectors represented by arrows. If we shift one vector arrow, for instance $\mathbf{x_2}$, without rotating it, from the origin O to a point where its tail coincides with the tip of the other vector arrow $\mathbf{x_1}$, then we can draw a vector sum, also called the *resultant vector*, as an arrow connecting the origin O with the tip of the shifted vector. The three vectors $\mathbf{x_1}$, $\mathbf{x_2}$, and $\mathbf{x_1} + \mathbf{x_2}$ form the so-called vector triangle.

To sum vectors defined in Cartesian coordinates, add up x and y coordinates separately,

$$\mathbf{x_1} + \mathbf{x_2} = \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 + x_2 \\ y_1 + y_2 \end{bmatrix}. \tag{16}$$

For a complex vector representation, real and imaginary parts sum separately,

$$z_1 + z_2 = (a + \mathrm{i}b) + (c + \mathrm{i}d) = a + c + \mathrm{i}(b + d). \tag{17}$$

When summing more than two vectors, the order of addition does not affect the result. For example, we have: $(\mathbf{x_1} + \mathbf{x_2}) + \mathbf{x_3} = \mathbf{x_1} + (\mathbf{x_2} + \mathbf{x_3})$. Geometrically, each vector can be placed tip-to-tail in sequence. The resulting chain of vectors forms a vector polygon, and the resultant vector is the one that connects the origin to the tip of the final vector in the sequence (Figure 8F).

Visualizations of Circular Data

Visualizing circular data requires approaches that respect its inherent periodicity and directional nature. A key decision to make is whether data are unwrapped onto a line and represented in a Euclidean coordinate system or represented in a two-dimensional polar coordinate system – where angle is mapped to direction (θ) and magnitude to radius (r). In some cases, linear and polar visualizations convey similar information, whereas in others, one format provides a much clearer or more intuitive representation of the data. This section will outline common approaches and potential pitfalls in the visualization of circular data, concluding with a discussion of methods for visualizing summary statistics. We also provide practical guidance on producing these visualizations in commonly used programming languages.

Linear Representations of Circular Data

One common way to visualize circular data is by unwrapping it into a *linear histogram* along the circular dimension, using binned counts (absolute or relative frequencies) or continuous density estimates. This requires two key decisions: selecting the interval and – as for Euclidean data – defining the bins. Both the standard interval $[0, 2\pi)$ and the signed interval $[-\pi, \pi)$, are common choices for linear visualizations. The latter is especially useful if there is a meaningful zero point, such as for errors relative to a reference angle. The choice of the interval determines the wrap-around point, which critically affects how clearly the structure of the data is displayed. When data are unimodal and highly concentrated, a linear histogram can closely resemble a standard histogram of non-circular data and often provides a simple, interpretable summary. In such cases, the interval should be chosen so that the region of highest density lies in the center of the x-axis. For multimodal distributions, boundaries should, where possible, be positioned in regions of low density to ensure that all modes are clearly represented. For instance, in Figure 9A, the shape of the distribution around the dominant mode is clearly presented, while the secondary mode is positioned too close to the boundary, leading to density spilling over to the opposite side of the plot.

As for Euclidean data, the choice of bin size further shapes the appearance of the histogram: wider bins "smooth" over noise but may conceal fine structure in the data, whereas narrower bins reveal more detail but are susceptible to spurious fluctuations (Figure 9B). Common rules for selecting the size and number of bins in Euclidean space (e.g., Freedman & Diaconis, 1981; Scott, 1979, 2010; Sturges, 1926) are sometimes applied to circular data by treating angular values as lying on the real number line. However, heuristics based on dispersion measures may perform poorly in circular space and even yield pathological results, such as vanishing bin widths for near-uniform data. Currently, no universally accepted bin-selection rule exists for circular histograms. In practice, researchers often use a fixed number of bins. Crucially, bin widths and locations should be selected so that edges coincide with the boundaries of the presentation interval. Another consequence of mapping circular data to a linear axis is that subgroup visualization no longer requires any circular-specific treatment; the same methods used for Euclidean data apply directly, including grouped, overlaid, or stacked histograms (e.g., Figure 9I).

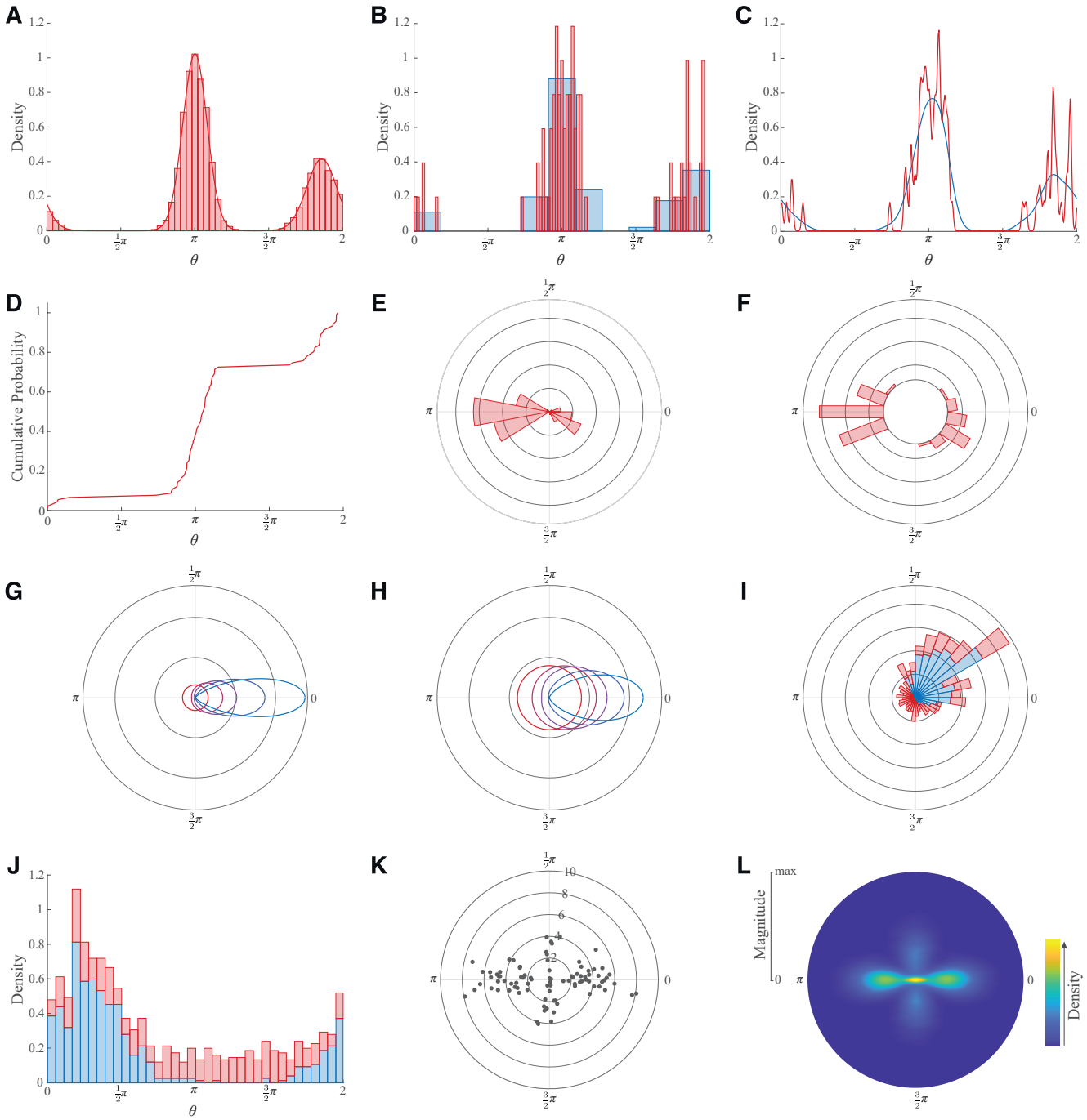
A related method is to plot a *linear continuous density curve*, which provides a smoothed representation of the distribution. For sufficiently dense data, continuous density curves intuitively summarize the overall structure of the distribution without being tied to arbitrary bin edges. However, as for bin size in histograms, the choice of smoothing parameter strongly influences the result: too much smoothing can mask structure, while too little smoothing exaggerates noise (Figure 9C). Additionally, a smoothing algorithm that takes into account the connectedness of the two ends of the interval must be chosen. As for the Euclidean case, circular continuous density plots are not appropriate for discrete or heavily rounded data or for sparse data, where density estimates can be unstable or misleading. For a short review of nonparametric methods for smooth density estimation of circular data, see Chaubey (2022).

The *empirical cumulative distribution function* (eCDF) summarizes the data without requiring a choice of bin size or smoothing parameter (Figure 9D). It shows the cumulative proportion of observations along the circular axis and can help highlight changes in density while avoiding some of the visual noise of histograms. Like other linear representations, however, eCDFs still require choosing an interval, and selecting the interval so that the wrap-around point lies opposite the region of highest concentration helps to center the distribution.

For producing linear plots of circular data, readers can use the same software and functions familiar from visualizing Euclidean data. For this, it is essential that the data are correctly wrapped (i.e., mapped to their representative values within the interval selected for visualization), and if applicable, that density estimation is conducted using methods accounting for the periodicity of the data. Once the circular data are appropriately preprocessed, linear plots can be generated by using the same software for (binning and) plotting as for Euclidean data.

Figure 9

Visualization of Circular Data in Linear and Polar Spaces



Note. (A) Linear histogram of a bimodal circular distribution. Bars show the empirical density estimated from a random sample (N = 9,000), and the line shows the corresponding theoretical distribution. (B–F) All plots show the same sample data (N = 90) drawn from the theoretical distribution shown as the line in (A). (B) Effect of bin widths on the representation of circular data. Both histograms show the same data but with coarse (blue, 11 bins) and fine (red, 111 bins) binning. (C) Effect of smoothing strength on circular density estimation. Both curves represent the same data with stronger (blue) and weaker (red) smoothing. (D) Empirical cumulative distribution function. (E, F) Polar histograms showing density as a function of angle: polar histogram with wedges (E) and wrapped bar histogram with rectangular bars (F). (G) Polar plot of von Mises distributions (circular normal distributions, see the *Probability Theory and Models for Circular Variables* section) with probability density plotted in the radial dimension. Note that dense regions appear exaggerated, whereas sparse regions are understated because area scales with the square of the radius. (H) Polar plot of the same von Mises distributions as in (G) with the square root of probability density on the radial axis. This area-proportional mapping ensures that enclosed areas are linearly proportional to density. (I, J) Stacked histograms of the same data with two subgroups (blue and red) layered one on top of another in polar (I) and linear circular space (J). Note that the red subgroup is approximately uniformly distributed. However, as area grows with the square of the radius in rose plots, bins of the red subgroup appear larger for bins where density in the blue subgroup is higher in the polar representation (I). (K, L) Representations of cylindrical data with one circular and one Euclidean dimension. (K) Polar scatter plot showing cylindrical data with the circular and Euclidean dimensions represented by angle and distance from the center, respectively. (L) Polar heat map showing a large number of samples drawn from the same distribution as data in (K). Density is represented smoothly by a change in hue across both angular and radial dimensions.

Polar Representations of Circular Data

We saw that linear visualizations such as histograms and eCDFs are intuitive and effective for concentrated, unimodal data. However, they may obscure structure in more dispersed or multimodal distributions, especially near wrap-around points. In such cases, polar visualizations often provide a clearer representation.

Polar histograms are widely used for polar representation of circular data (Figures 9E and 9F). The *rose plot* is a type of polar histogram that represents frequency data as radial wedges of equal angular width radiating from the center of a circle (Figure 9E). The height of each wedge indicates the number or proportion of observations in that angular bin. Unlike linear histograms, this approach preserves the continuity of the circular scale and avoids edge artifacts introduced by unwrapping. It also makes features such as the central tendency and directional biases easier to see. However, the representation is still influenced by the choice of bin size and the placement of the bin edges (for an illustration, see Figure 1 in Wells, 2000).

Analogous to its linear equivalent, a *polar continuous density curve* maps the smoothed density around the circle (Figures 9G and 9H). This representation conveys the same information as the linear density plot but emphasizes the cyclical structure of the data and does not require choosing a wrap-around point. As before, the choice of smoothing parameter strongly influences the result, and continuous density plots are also unsuitable for sparse data and data rounded to a coarse resolution, where density estimates can become unstable. A key consideration in polar representations is how magnitude is mapped to radial distance. Because the area of a polar region (i.e., a wedge or the area under a smooth curve) grows with the square of its radius, mapping frequency or density directly to radial height causes area to scale quadratically (Figure 9G). This distorts the apparent magnitude because small differences in height translate into large differences in area. As a result, dense regions are spatially overrepresented, while sparse areas are understated and difficult to distinguish. The problem arises for both binned counts and smooth densities. To correct this, radial height can be mapped to the square root of frequency or density (Mardia & Jupp, 2009). This ensures that the area of each sector is linearly proportional to the actual quantity, making sector areas directly comparable across angles. However, this comes at the cost of interpretability of the radial axis: because radial distance now represents the square root of frequency or density, values can no longer be read directly from the radius. Figure 9H shows continuous density plots of the same data as Figure 9G, using an area-proportional representation. For a more extensive discussion on this topic, see Xu and Wang (2020).

A version of polar histograms that avoids a distorted representation of area is *wrapped bar histograms*, which use rectangular bars arranged around a reference circle instead of wedges (Figure 9F). The base of each bar is aligned with the reference circle's circumference, and the radius of the reference circle is set arbitrarily to ensure that the bars fit around its perimeter. This method effectively wraps a linear histogram around a circle, preserving the height of each bar relative to its linear counterpart. This approach avoids the distortions of apparent magnitude typical of wedge-based rose plots, because both height and area are proportional to frequency. However, the gaps between adjacent bars complicate overlaying smooth density curves: to remain consistent with bar heights, density must be mapped to height rather than area; as a result, area under the density curve can misrepresent magnitude, with a disproportionately larger area plotted around longer bars compared to shorter bars. For a more detailed discussion of polar histograms and their design considerations, see Sanderson and Peacock (2020).

Stacked polar histograms (Figure 9I) can be used to show subgroup compositions within each angular bin. Annular sectors corresponding to different groups are layered radially one on top of another. However, this design inherits the same distortion mentioned previously: subgroups farther from the center occupy disproportionately more area, even when counts are equal. This occurs because area increases with the square of the radial distance, so visually comparing sector sizes can give a misleading impression of the underlying proportions. *Stacked linear histograms* show subgroup magnitudes in a clearer way (Figure 9J), even though the stacked representation still makes it difficult to compare the same subgroup across bins or relative subgroup contributions within a bin. For these comparisons, alternatives such as grouped polar bars, unstacked linear histograms, or proportion-normalized plots provide more reliable summaries.

To produce *polar plots*, MATLAB provides the built-in function `polarplot`. In R, a linear plot generated with `ggplot2` (Wickham, 2016) can be transformed into a polar representation by adding `coord_polar()` as a coordinate layer; alternatively, polar plots can be directly created using the `polar.plot` function from the package `plotrix` (Lemon, 2006). In Python, the visualization library `matplotlib` (Hunter, 2007) supports polar plotting via the `polar` projection in `matplotlib.pyplot.subplot`. For special considerations when creating polar plots of time-series data, refer to Lee Rodgers et al. (2014).

Visualizations of Cylindrical Data

Data consisting of paired circular and linear variables is referred to as cylindrical. Examples include the pairing of movement direction and amplitude, and pairs of visual features like edge orientation and spatial frequency. One way of visualizing cylindrical data is a *polar scatter plot*, where each point is positioned according to its angle around the circle and its distance from the center (Figure 9K). For large datasets

or continuous variables, scatter plots can become cluttered and difficult to interpret. In these cases, *polar heatmaps* offer a useful alternative (Figure 9L). Such plots summarize data density or intensity smoothly across both the circular and radial dimensions, making it easier to identify patterns in the data. Cylindrical data can also be represented using two-dimensional Cartesian plots, where one axis corresponds to one period of the circular data displayed in linear form.

Polar scatter plots can be generated using MATLAB's built-in function `polarscatter`, the `polar` projection with `scatter` in `matplotlib` (Hunter, 2007) in Python, and `geom_point()` combined with `coord_polar()` in `ggplot2` (Wickham, 2016) or `radial.plot` from `plotrix` (Lemon, 2006) in R. Polar heatmaps can be implemented by first converting values to Cartesian coordinates using the transformations defined in the *Representations of Circular Data* section and then applying standard two-dimensional binning and rendering functions familiar from generating heatmaps of Euclidean data.

Visualizing Summary Statistics

While histograms and density plots illustrate the overall distribution of data, visualizations of summary statistics convey key properties, such as the dominant direction in the data or how strongly clustered the observations are. Some of the concepts that we will introduce in this subsection (e.g., mean vector and circular median) are explained in greater detail in the *Descriptive Statistics of Circular Data* section. For now, the focus is on how these statistical summaries can be effectively represented in visual form, rather than on details of their calculation.

Rayleigh plots (Figure 10A and 10B) provide a simple and informative way to visualize both individual data points and summary statistics in a single plot. They are related to circular histograms but – in addition to individual observations – also visually summarize central tendency and spread. Instead of grouping observations into bins, Rayleigh plots represent each data point as a mark positioned at its corresponding angle on the circumference (Figure 10A). While typically used for ungrouped data, Rayleigh plots can be adapted for grouped data. In such cases, or when multiple observations share the same angle, instead of being overlaid, points can be stacked radially to show frequency (Figure 10B). Stacking points makes individual observations easier to distinguish compared to the classic Rayleigh plot, where overlapping points can obscure one another. A distinctive feature of the Rayleigh plot is the mean vector, drawn as an arrow from the center of the circle. The angle of this arrow indicates the direction of central tendency, while its length reflects the spread of the data. Thus, it provides an intuitive summary of both the main direction and the consistency of observations around it. For a more detailed mathematical definition of the mean vector, see the *Circular Mean* section.

For Euclidean data, a *box plot* typically displays lower quartile (Q1), median, upper quartile (Q3), with additional elements indicating the spread beyond the quartiles and potential outliers. Circular box plots (Figure 10C and 10D) aim to represent similar features but face unique challenges due to the periodic nature of angular data. A critical challenge is that, in circular space, the median and quartiles are not always uniquely defined, particularly in multimodal or uniform distributions (see the *Circular Median* section). If these order statistics are not uniquely defined, a circular box plot cannot be uniquely defined either. Despite this limitation, various implementations attempt to visualize the angular spread and central range using arcs or radial sectors to represent interquartile ranges, with additional markers for extreme values when identifiable. Like other visualizations of circular data, circular box plots can be displayed either on a circular (Figure 10C) or linear scale (Figure 10D). The circular format can be particularly useful for comparing multiple circular variables or groups, where box plots are positioned at increasing radii. However, comparing such box plots can be misleading: box plots plotted farther from the origin will be physically longer than those closer in, even if their angular spreads are identical. This can distort the perceived variability and exaggerate or diminish differences between groups. In the example shown in Figure 10C and 10D, the three groups appear more different in terms of spread when displayed on a circular plot compared to a linear one. To support interpretation, circular box plots can include radial grid lines and concentric circles, which serve as visual reference points for comparing angular spread and position across groups plotted at different radii.

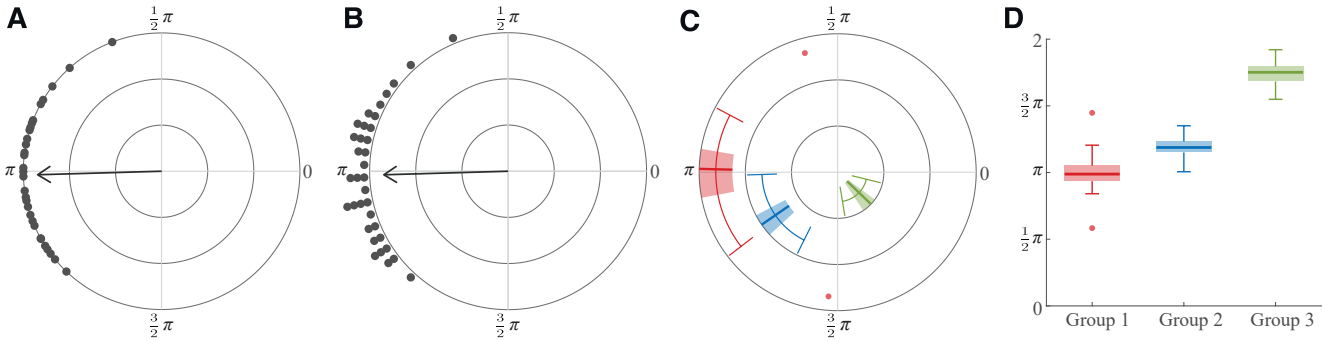
The implementation of Rayleigh plots follows the same approach as described for polar scatter plots, with the radius of each point fixed to a constant value for ungrouped data or incremented radially for each observation sharing the same angle for grouped data. The mean vector can be added as an arrow using `quiver` in MATLAB, `annotate` with `arrowprops` in `matplotlib` in Python, and `geom_segment()` in `ggplot2` in R. For a more detailed discussion of possible ways to implement circular box plots, refer to Abuzaid et al. (2012) and Buttarazzi et al. (2018).

Descriptive Statistics of Circular Data

In this section, we will review the descriptive statistics most commonly used to summarize circular datasets. For each statistic, we will give its mathematical definition, the intuition behind it, and its relationship to the corresponding Euclidean statistic. Where applicable, we will also discuss the circumstances under which Euclidean approximations of circular measures provide acceptable accuracy. As mentioned in the introduction, mathematical notation in circular statistics is less standardized than in Euclidean contexts. Some of the symbols introduced in this section represent our own

Figure 10

Visualization of Summary Statistics



Note. (A, B) Rayleigh plots. The arrow indicates the mean resultant vector (see Section Circular Mean). (A) Rayleigh plot showing 35 individual observations (dark gray circles). Note that it is not possible to see whether multiple observations share the same value. (B) Stacked Rayleigh plot of the same data shown in (A) but now binned into 5° intervals, with observations stacked radially within each bin. (C, D) Box plots summarizing three groups of circular observations with circular median (bold line), interquartile range from the 25th to the 75th percentile (opaque box), range of data not including outliers (whiskers), and outlier values (dots). Note that data dispersion is similar for Group 1 (red) and Group 3 (green), which is apparent in the linear representation (D). In the polar representation (C), the radial distance between the two groups gives the impression that Group 1 has a higher dispersion.

suggested notation (for an overview, see the middle column in Table 1). For a discussion of the corresponding population parameters and their estimation, see Section Probability Theory and Models for Circular Variables.

Central Tendency

Circular Mean

A basic requirement of any statistic calculated from circular data is that it respects the equivalence of angles separated by a whole number of cycles, for example, that it is unchanged if we describe a data point as 360° instead of 0°. The usual definition of the arithmetic mean $\frac{1}{n}\sum_{i=1}^{n} x_i$ does not fulfill this requirement. However, the problem of identifying a circular counterpart to the Euclidean mean turns out to be even more fundamental in nature. Consider a data sample consisting of the four cardinal directions (0°, 90°, 180°, and 270°). Intuitively, no single direction can be identified as its average. It is clear that a different conception of mean value is required for circular data: the centroid, or center of mass, fulfills the requirements.

In Euclidean space, the arithmetic mean is equivalent to the centroid of the set of points x_i on the real number line. This idea can be illustrated using the analogy of a seesaw: Placing equally weighted balls at each location x_i on a beam, the mean corresponds to the center of mass, or the position of a fulcrum that balances the system. We can extend this concept to circular data by treating each angle θ_i as a point on the unit circle (the embedded representation introduced in Section Polar Coordinates). In this case, the centroid lies inside the circle rather than on its circumference (purple dots in Figure 11). The analogy is to a disk with equally weighted balls placed along its edge at angles θ_i. Placing a small plinth underneath the disk to keep it balanced would require positioning it closer to the center and not on the edge. The position of the plinth reflects the location of the center of mass.

The direction of the center of mass from the origin defines the circular mean – a generalization of the linear mean that respects the periodic and directional nature of circular data. It is the same as the direction of the resultant vector (the sum of the n unit vectors corresponding to the data; see Section Vector Summation). The two-dimensional location of the centroid can be obtained by scaling the resultant vector by $\frac{1}{n}$. The length of this *mean resultant vector* contains information about the spread of the data, as we will see in Section Measures of Dispersion. If the mean resultant vector has length 1, the circular mean is undefined, as this indicates complete directional cancellation, as in a uniform distribution.

As described in Section Complex Coordinates, a two-dimensional vector can be conveniently represented by a complex number. When represented in this way, we will denote the mean resultant vector by $\widehat{m}_1$ (the origin of this notation will become clear in Section Circular Moments). It is defined as

$$\widehat{m}_1 := \frac{1}{n}\sum_{i=1}^{n} e^{\mathrm{i}\theta_i} = \bar{R}e^{\mathrm{i}\bar{\theta}}, \tag{18}$$

where $\bar{R} = |\widehat{m}_1|$ is the modulus or length of the mean resultant vector and $\bar{\theta} = \arg(\widehat{m}_1)$ is its argument or direction, which defines the circular mean.

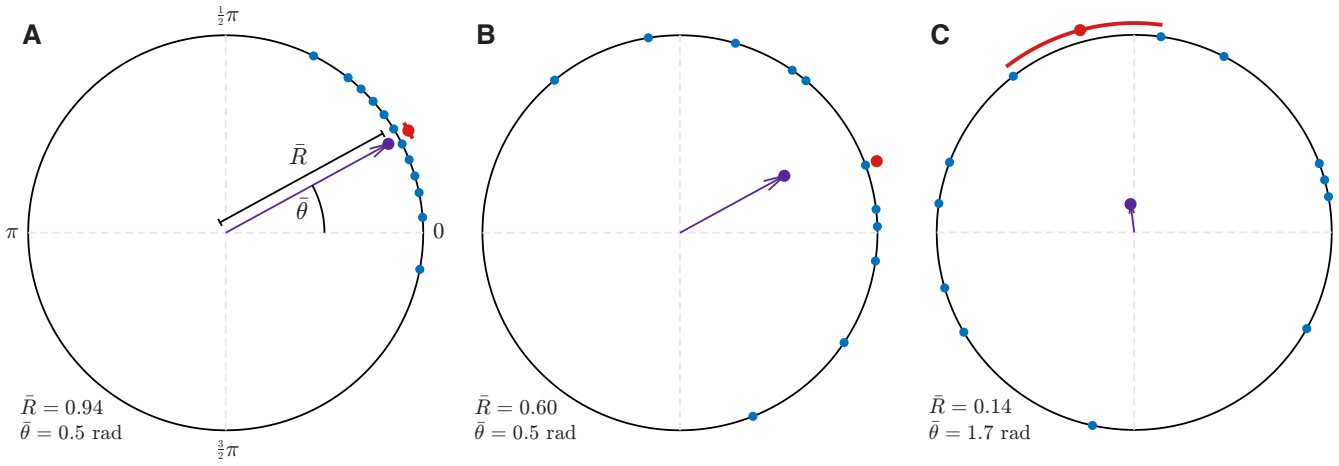
The mean resultant vector can alternatively be represented in polar coordinates, as $(\bar{\theta}, \bar{R})$, or in Cartesian coordinates as

$$\frac{1}{n}\left[\sum_{i=1}^{n}\cos\theta_i, \quad \sum_{i=1}^{n}\sin\theta_i\right]. \tag{19}$$

$\bar{R}$ can also be calculated using the Pythagorean formula with

Figure 11

Center of Mass and Central Tendency of Circular Data



Note. The panels display samples of 12 (A,C) or 11 (B) data points on the unit circle. Each data point θ_i corresponds to a complex number $e^{i\theta_i}$ with unit length. The purple dots and arrows show the center of mass and the mean resultant vector for each sample. The circular mean, $\bar{\theta}$, is the angle of this vector, and the mean resultant vector, $\bar{R}$, represents its length. The sample in (B) has the same circular mean as the one in (A), but the center of mass is closer to the center of the circle as the sample displays a higher spread. (C) displays the most dispersed sample, with the center of mass close to the center of the circle. The red points in (A–C) show possible values of the median of the different samples; the red arcs in (A) and (C) mark the range within which each median could fall. Note that the median of the odd number of data points in (B) is unique. Medians are presented outside of the circle to avoid overlapping with data points.

the x- and y-components of the Cartesian resultant vector,

$$\bar{R} = \sqrt{\left(\frac{1}{n}\sum_{i=1}^{n}\cos\theta_i\right)^2 + \left(\frac{1}{n}\sum_{i=1}^{n}\sin\theta_i\right)^2}, \tag{20}$$

$$\bar{\theta} = \operatorname{atan2}\left(\frac{1}{n}\sum_{i=1}^{n}\sin\theta_i, \frac{1}{n}\sum_{i=1}^{n}\cos\theta_i\right), \tag{21}$$

recalling from Section Representations of Circular Data that atan2 first takes the vertical and then the horizontal component.

The circular mean $\bar{\theta}$ is the angle that maximizes the sum of cosines of angular distances between each data point and that angle:

$$\bar{\theta} = \arg\max_{\phi}\sum_{i=1}^{n}\cos(\theta_i - \phi). \tag{22}$$

Equivalently, we can say it is the angle with the minimum summed cosine dissimilarity from the data (see Section Measuring Distance). This is analogous to the way the Euclidean mean minimizes the sum of squared deviations from the data.

In the special case that data points are concentrated in a small region of the circular space, some (but not all) circular statistics can be approximated by their Euclidean counterparts. Specifically, if we choose an interval in which to represent the circular data such that the data are tightly concentrated in their center (e.g., $[-\pi, \pi)$ if the data are concentrated around 0), then the "wrapping around" at the bounds of the interval will have a negligible effect. In this case, the Euclidean mean will closely approximate the circular mean.

Formally, we can approximate sine and cosine by $\sin(\theta_i) \approx \theta_i$ and $\cos(\theta_i) \approx 1$ (red dashed lines in Figure 12A and 12B). Inserting these approximations into Equation 21 yields the Euclidean mean: $\bar{\theta} \approx \operatorname{atan2}\left(\frac{1}{n}\sum_{i=1}^{n}\theta_i, 1\right) = \frac{1}{n}\sum_{i=1}^{n}\theta_i$. However, this approximation of cosine is only sufficiently precise for very small angles $\theta_i \approx 0$. As an alternative that is suitable for concentrated data that span a larger arc, we can use the second-order expansion $\cos(\theta_i - \phi) \approx 1 - \frac{1}{2}(\theta_i - \phi)^2$ (green dashed line in Figure 12B). If we apply this to Equation 22, we obtain $\bar{\theta} \approx \arg\min_{\phi}\sum_{i=1}^{n}(\theta_i - \phi)^2$. The right-hand side coincides with the variational characterization of the Euclidean mean that defines it as the value that minimizes the sum of its squared distances from all data points.

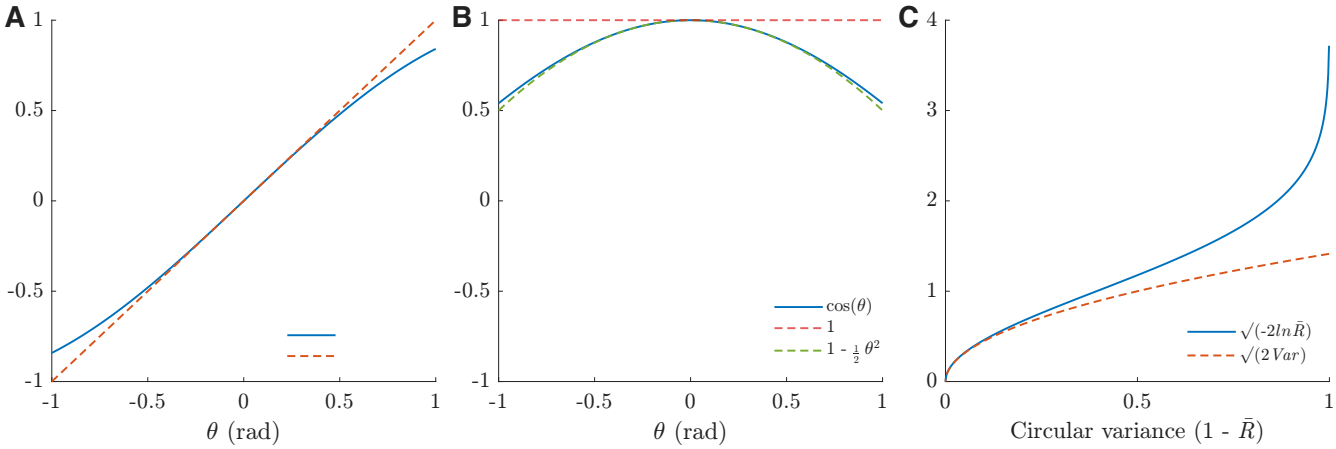
Circular Median

In Euclidean space, the median of a sample separates the data into an upper half and a lower half. Similarly, the sample *circular median* can be described as splitting the data into a clockwise half, contained in the interval $(\bar{\theta}_{\text{med}} - \pi, \bar{\theta}_{\text{med}})$, and a counterclockwise half, in $(\bar{\theta}_{\text{med}}, \bar{\theta}_{\text{med}} + \pi)$. However, in circular space, for any direction ϕ that meets this criterion, the diametrically opposite direction $\phi \pm \pi$ does so as well. To resolve this ambiguity, the circular median is defined as a point on the circle that minimizes the sum of angular distances between each data point and that angle,

$$\bar{\theta}_{\text{med}} := \arg\min_{\phi}\sum_{i=1}^{n} D(\theta_i, \phi), \tag{23}$$

Figure 12

Euclidean Approximations of Circular Measures



Note. (A, B) Approximations of sine (A) and cosine (B) for small values. (C) Approximation of circular variance; relationship between circular variance and circular standard deviation. Blue continuous lines represent exact values; dashed lines indicate approximations.

where $D(\theta_i, \theta_j)$ denotes the shortest arc length between two angles (see Section Measuring Distance). Of the two opposite directions that split the data in two, this identifies the circular median as the one more strongly in the overall direction of the data. This definition is analogous to the Euclidean median, which can be defined as a point that minimizes the summed absolute distance from the data. If the number of data points is odd, one of them satisfies Equation 23 (red dot in Figure 11B), but if the number is even, the set of solutions is an interval subtended by two neighboring data points (red arcs in Figures 11A and 11C). In this case, it is usual to take the midpoint of the interval as the circular median. This all corresponds closely to the way the median is identified in Euclidean data. Note, however, that for certain precise geometric arrangements of data points, the circular median is still not uniquely defined. As a trivial example, for any odd number of data points arranged evenly around the circle, every data point has an equal claim to be the circular median.

The circular median is less sensitive to extreme observations or *outliers* than the circular mean, and can provide a robust measure of central tendency, similarly to the median versus the arithmetic mean in Euclidean space (see Section Outlier Classification and Exclusion). When most data points are clustered in a small region in the interior of an interval of 2π, the circular median is approximately equal to the Euclidean median calculated over that interval.

Circular Mode

The circular mode of a sample is the most frequently occurring angle, taking into account the equivalence of angles separated by whole numbers of cycles. As in the Euclidean case, the circular mode is strictly only defined for discrete data, but the term may also be used in a looser sense to describe peaks in a histogram or smoothed density estimates of continuous data. Whereas the global mode is rarely used to estimate central tendency for continuous circular data, local mode estimates are relevant for several applications, including multimodality testing (Ameijeiras-Alonso et al., 2019; Bolón et al., 2025) and mode-based clustering approaches (Chang-Chien et al., 2012; Y. Zhang & Chen, 2021).

Measures of Dispersion

Circular Variance

As mentioned above and displayed in Figure 11, the mean resultant length $\bar{R}$ conveys information about the spread of a data sample, by quantifying how tightly the angles cluster around the mean direction. Specifically, $\bar{R}$ lies in the range $0 \leq \bar{R} \leq 1$, with $\bar{R} = 1$ indicating perfect alignment of all angles (i.e., maximum concentration) and $\bar{R} = 0$ implying complete directional cancellation (i.e., maximum dispersion). The sample *circular variance* is defined as

$$\widehat{V}_\theta := 1 - \bar{R}, \quad 0 \leq \widehat{V}_\theta \leq 1, \tag{24}$$

reflecting the inverse relationship between $\bar{R}$ and spread.

The circular variance is equal to the average cosine dissimilarity between the data points and their circular mean,

$$\widehat{V}_\theta = \frac{1}{n} \sum_{i=1}^{n} \left(1 - \cos(\theta_i - \bar{\theta})\right), \tag{25}$$

in the same way that the Euclidean variance is the average squared deviation from the arithmetic mean. Unlike many other circular statistics, the circular variance of clustered data is not approximately equal to the corresponding Euclidean measure. Instead, for data that are concentrated within a

small arc of the circle, circular variance is approximately *half* the value of the Euclidean sample variance (red dashed line in Figure 12C). This is a consequence of the fact that for small δ, $1 - \cos(\delta) \approx \frac{\delta^2}{2}$.

Circular Standard Deviation

In Euclidean statistics, variance is expressed in squared units of the data; to quantify dispersion in the same units as the data themselves, the standard deviation is defined as the square root of the variance. However, applying this transformation to the circular variance, which takes values in $[0, 1]$, would not yield a measure of spread in radians or any other natural angular units. Instead, the *circular standard deviation* s_θ, is defined as

$$s_\theta := \sqrt{-2 \ln \bar{R}} = \sqrt{-2 \ln(1 - \widehat{V}_\theta)}. \tag{26}$$

This definition might seem arbitrary and a bit peculiar, but it has a logic based on the Gaussian or normal distribution. Specifically, if you wrap a Gaussian probability distribution with standard deviation σ around the unit circle (Section Wrapped Normal Distribution), the resulting distribution will have a circular standard deviation equal to σ, based on this definition. In other words, s_θ is the standard deviation of a normal distribution wrapped around the circle that has the same mean resultant length as the data.

Importantly, the logarithm in Equation 26 converts the bounded range of $\bar{R}$ into an unbounded scale. As $\bar{R}$ approaches 1, s_θ approaches 0, corresponding to maximum concentration. Conversely, as $\bar{R}$ approaches 0, s_θ tends to infinity, that is, maximum spread (blue line in Figure 12C). Thus, s_θ can in principle take arbitrarily large values, just like the Euclidean standard deviation, even though the data are geometrically confined to a circle. Provided s_θ is substantially smaller than the maximum angular distance of a point from the circular mean ($s_\theta \ll \pi$), it can be interpreted as a standardized measure of dispersion in angular units, directly comparable with s in Euclidean space. However, once s_θ approaches or exceeds π, although it remains a valid measure of spread, the interpretation in angular units breaks down, because the topology of the circle means angular distances are confined to at most π rad.

For data concentrated within a small arc of the circle, s_θ can be well approximated by its Euclidean equivalent s. In this regime, $\bar{R}$ is close to 1; thus, the circular variance $\widehat{V}_\theta = 1 - \bar{R}$ is close to 0. Using the Taylor expansion $\ln(1 + x) \approx x$ for small x, we obtain $s_\theta = \sqrt{-2 \ln(1 - \widehat{V}_\theta)} \approx \sqrt{2\widehat{V}_\theta}$. As discussed in Section Circular Variance, for concentrated data, $2\widehat{V}_\theta$ is approximately equal to the Euclidean sample variance, so s_θ is approximately equal to the Euclidean sample standard deviation, s.

Circular Moments

To calculate circular mean, variance, and standard deviation, we used the mean resultant vector, which can also be characterized as the *first circular moment* of the data distribution. Circular moments, also called *trigonometric moments*, form an important class of summary statistics. They act as a "fingerprint" of the distribution, allowing complex patterns to be summarized by a few interpretable numbers.

In Euclidean statistics, the p-th moment of the sample distribution is the arithmetic mean of the p-th power of the data values, $\frac{1}{n} \sum_i x_i^p$. However, this definition does not transfer directly to circular data, because it ignores periodicity and directionality. To properly capture the geometry of the circle, powers of θ are replaced with powers of the complex exponential $e^{i\theta} = \cos\theta + i \sin\theta$, which is periodic with period 2π.

The p-th sample circular moment is therefore defined as

$$\widehat{m}_p := \frac{1}{n} \sum_{i=1}^{n} e^{ip\theta_i} = \bar{R}_p \, e^{i\bar{\theta}_p}. \tag{27}$$

The first circular moment, $\widehat{m}_1$, is the mean resultant vector (Section Circular Mean), its modulus $|\widehat{m}_1|$ is the mean resultant length $\bar{R}$, and its argument or phase $\arg(\widehat{m}_1)$ is the circular mean $\bar{\theta}$. More generally, the modulus and argument of each circular moment $\widehat{m}_p$ yield interesting statistics.

The argument of the p-th circular moment,

$$\begin{aligned} \bar{\theta}_p &:= \arg(\widehat{m}_p) \\ &= \operatorname{atan2}\left(\frac{1}{n} \sum_{i=1}^{n} \sin(p\theta_i), \frac{1}{n} \sum_{i=1}^{n} \cos(p\theta_i) \right) \end{aligned} \tag{28}$$

corresponds to the mean direction of the data when considered modulo $\frac{2\pi}{p}$, that is, when directions separated by $\frac{2\pi}{p}$ are treated as equivalent. Higher-order moment arguments generalize the circular mean to capture dominant symmetries in circular data. For instance, $\bar{\theta}_2$ gives the dominant axis when opposite directions are treated as equivalent, whereas $\bar{\theta}_3$ points to the orientation of a threefold (trimodal) arrangement.

The modulus or length of the p-th moment

$$\bar{R}_p := |\widehat{m}_p| = \sqrt{\left(\frac{1}{n} \sum_{i=1}^{n} \cos(p\theta_i) \right)^2 + \left(\frac{1}{n} \sum_{i=1}^{n} \sin(p\theta_i) \right)^2}, \quad 0 \le \bar{R}_p \le 1 \tag{29}$$

measures how strongly the data cluster when considered modulo $\frac{2\pi}{p}$. Across orders, a value of $\bar{R}_p = 0$ indicates complete uniformity around the circle. $\bar{R}_p = 1$ corresponds to perfect alignment of the data wrapped onto a circle of length $\frac{2\pi}{p}$, which has different implications for different values of p. For

$p = 1$, this means all observations fall at the same angle, for $p = 2$ they can be split between two opposite angles, and so forth.

While the argument and the modulus of the first circular moment describe central tendency and spread, higher-order circular moments allow us to characterize further aspects of a distribution's shape and symmetry. We turn to these measures in the following sections.

Circular Skewness

Skewness describes the asymmetry or lopsidedness of a distribution. In Euclidean statistics, a nonzero skewness statistic indicates that the data spread further from the mean in one direction than the other, with its sign corresponding to the direction of the long tail; it is calculated as the third moment of the z-scored (standardized) data. *Circular skewness* extends this concept to periodic data, quantifying the asymmetry of a distribution around the circular mean. The sample circular skewness is defined by Mardia (1975) as

$$g_{1\theta} := \frac{\bar{R}_2 \sin(\bar{\theta}_2 - 2\bar{\theta})}{(1-\bar{R})^{\frac{3}{2}}} = \frac{1}{\widehat{V}_\theta^{\frac{3}{2}}} \frac{1}{n} \sum_{i=1}^{n} \sin 2(\theta_i - \bar{\theta}), \quad (30)$$

where the second expression illustrates that, analogously to z-scoring in Euclidean space, the data are centered on the mean $\bar{\theta}$ and normalized by the variance $\widehat{V}_\theta$. In line with its Euclidean equivalent, a circular skewness of 0 indicates symmetry, whereas larger deviations from 0 reflect increasing asymmetry (Figure 13A). However, a positive value indicates that the long tail is in the negative (clockwise) direction from the mean (green line in Figure 13A), whereas a negative value indicates a longer positive (counterclockwise) tail (red line in Figure 13A) – the opposite sense to the Euclidean skewness. For data concentrated in the interior of a 2π interval, the circular skewness is not approximated by the Euclidean skewness g_1 directly but rather by

$$g_{1\theta} \approx -2^{\frac{3}{2}} g_1, \quad (31)$$

where the negative sign reflects the reversal of sense described above and the power-of-two scaling is due to the fact that circular variance approaches half the Euclidean variance for concentrated data (Section Circular Variance).

Circular Kurtosis

Kurtosis describes how much probability mass lies in the tails of a distribution. The Euclidean sample kurtosis b_2 is the fourth moment of the z-scored data. As with skewness, z-scoring makes the kurtosis invariant to shifts or scaling of the data. A Gaussian distribution has a kurtosis of 3, and as a result, the *excess kurtosis* of a sample in Euclidean space is defined as $g_2 = b_2 - 3$. The distribution of a typical unimodal sample with positive excess kurtosis has a narrow peak around the mean and, extending away, heavier tails than the Gaussian distribution. Conversely, the distribution of a sample with negative excess kurtosis is flatter and more spread out around the mean. The sample *circular kurtosis* is defined by Mardia (1975) as

$$\begin{aligned} g_{2\theta} &:= \frac{\bar{R}_2 \cos(\bar{\theta}_2 - 2\bar{\theta}) - \bar{R}^4}{(1-\bar{R})^{\frac{3}{2}}} \\ &= \frac{1}{\widehat{V}_\theta^2} \left(\frac{1}{n} \sum_{i=1}^{n} \cos 2(\theta_i - \bar{\theta}) - \bar{R}^4 \right). \end{aligned} \quad (32)$$

It is the circular counterpart of Euclidean excess kurtosis. The reference point for zero circular kurtosis is the wrapping of the Gaussian distribution onto the unit circle (Section Wrapped Normal Distribution), with positive values indicating heavier tails (green line in Figure 13B) and negative values lighter tails (red line in Figure 13B) than the wrapped normal, in agreement with the Euclidean excess kurtosis. For concentrated data, the circular kurtosis can be approximated by twice the Euclidean excess kurtosis,

$$g_{2\theta} \approx 2 g_2, \quad (33)$$

with the factor-of-two scaling again attributable to the fact that circular variance approaches half the Euclidean variance for concentrated data (Section Circular Variance).

Outlier Classification and Exclusion

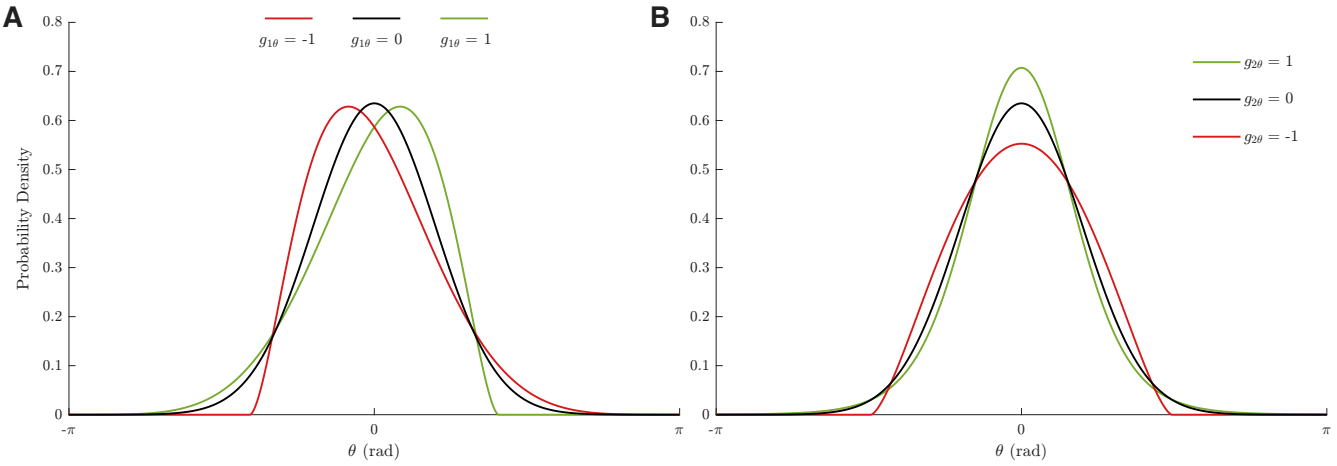
As in Euclidean statistics, an outlier in circular data is an observation that deviates markedly from others. In behavioral and cognitive research, such values may reflect measurement errors, misunderstandings of the task, lapses in attention, or meaningful extreme values. In circular statistics, detecting outliers requires special care because the ordering of the data is not unique and minimum and maximum values are not defined. Outliers should be excluded sparingly to avoid biasing the results. In this section, we briefly survey some methods for detecting outliers, from formal statistical tests to more algorithmic classification techniques.

One approach to detecting outliers in circular data is to use *statistical goodness-of-fit tests* that check whether the sample is drawn from an expected distribution. These tests can be appropriate if researchers have clear expectations about the data's distribution based on previous empirical findings or theoretical considerations. If the data deviate significantly from an expected distribution, this *may* indicate the presence of one or more outliers. Resources on suitable tests are provided in Section Beyond This Tutorial. While a significant result suggests the presence of outliers, statistical tests do not pinpoint which observations might be outliers.

Methods based on *descriptive statistics* identify outliers by analyzing how far each observation lies from a central value. In Euclidean statistics, it is an established approach to base the

Figure 13

Skewness and Kurtosis



Note. Circular distributions with different skewness (A) and kurtosis (B). In each panel, distributions have matching circular means and variances. Circular kurtosis and circular skewness are 0 for all distributions in (A) and (B), respectively. The black curves represent normal distributions wrapped around the circle (see Section Wrapping Euclidean Distributions). Contrary to the Euclidean equivalent, a positive circular skewness (green in Panel A) indicates that the long tail is in the negative (clockwise) direction from the mean, whereas a negative value (red in Panel A) indicates a longer positive (counterclockwise) tail.

threshold for outliers on z-scores. For example, outliers are sometimes defined as values that deviate from the mean by at least three standard deviations. Analogously, extreme values in circular space can be excluded using a cutoff of a maximum deviation from the circular mean in terms of circular standard deviations. However, outlier values can distort the circular mean and inflate the circular standard deviation. Therefore, alternative outlier classification methods use measures that are more robust against influences from outliers. The *median absolute deviation* is a robust measure of spread that can be used together with the circular median to identify outliers in circular data. It is calculated by computing the angular distance of each observation from the circular median, and then calculating the Euclidean median of those distances. Outliers can now be defined as values that deviate from the circular median by the median absolute deviation multiplied by some factor (e.g., three). While methods based on distance and deviation provide an intuitive approach toward identifying concrete outlier candidates, they are sometimes criticized for being exploratory and subjective.

Another approach to detecting outliers is to identify *patterns in the total data*, rather than comparing individual data points to the overall summary. These methods usually employ *clustering and density-based algorithms*. Clustering algorithms structure data into subsets, aiming for high similarity between data points in one cluster and low similarity between data points of different clusters. If most of the data form one large cluster and a few points form one or several tiny separate clusters, those few points can be deemed outliers. For circular clustering algorithms, see Section Beyond This Tutorial. While these methods allow for identifying multiple concrete outlier candidates, their results depend on the choice of the algorithm and hyperparameters.

Probability Theory and Models for Circular Variables

In this section, we introduce probabilistic foundations and modeling approaches for random variables in a circular space. The distribution of a circular random variable can be modeled by a probability density function (PDF) in a similar way to a Euclidean variable. The key difference is that a circular PDF must be periodic, such that $f(\theta) = f(\theta + 2\pi k)$ for any integer k. While a PDF in $\mathbb{R}^1$ has integral ("area under the curve") equal to one over the whole real line $(-\infty, \infty)$, a valid PDF in circular space must integrate to 1 over any interval of length 2π. Circular distributions can also be characterized by their cumulative distribution function (CDF). Given a chosen reference angle θ_0, the CDF is defined as $F(\theta) = \int_{\theta_0}^{\theta} f(\phi)\, d\phi$, so that $F(\theta_0) = 0$ and $F(\theta_0 + 2\pi) = 1$. The CDF can be constructed over any interval of length 2π; thus, the choice of starting point θ_0 is an essential part of its specification.

In this section, we will outline population counterparts of the sample statistics discussed in Section Descriptive Statistics of Circular Data, address the conditions for sample-to-population convergence in circular space, and describe some common circular population distributions. We will compare two circular analogs of the Gaussian distribution and introduce two more flexible families for fitting heavy-tailed and skewed data. We will examine the properties of circular random variables in more detail before concluding with practical guidelines for choosing between circular distributions and sampling random values from circular distributions. Throughout this section, our focus will be on continuous distributions; discussions of discrete or "lattice" distributions

on the circle can be found in Jammalamadaka and SenGupta (2001) and Mardia and Jupp (2009).

Population Parameters

As in the Euclidean case, each descriptive statistic for circular data has a corresponding population parametric form that can be used to summarize an equivalent aspect of the theoretical distribution of a random variable (for sample and population parameter notations, see Table 1). The population parameters simply replace normalized sums over the n data points in a sample with *expectations* or integrals over the circular space. So the p-th circular moment of a random variable with PDF $f(\theta)$ is

$$\mathrm{m}_p := \mathbb{E}_\theta[e^{\mathrm{i}p\theta}] = \int_\psi^{\psi+2\pi} e^{\mathrm{i}p\theta} f(\theta)\, d\theta, \tag{34}$$

where the integral can be taken over any interval of length 2π. The arguments $\mu_p = \arg(\mathrm{m}_p)$ and moduli $\rho_p = |\mathrm{m}_p|$ provide definitions for the population circular mean $\mu \equiv \mu_1$ and population mean resultant length $\rho \equiv \rho_1$. Population circular variance V_θ, standard deviation σ_θ, skewness $\gamma_{1\theta}$, and kurtosis $\gamma_{2\theta}$ all have definitions that follow from those of the corresponding sample statistics (Section Descriptive Statistics of Circular Data).

The circular median discussed in Section Circular Median is not moment-based. Like the sample version, the population circular median μ_{med} can be defined in two ways: first, as an angle that divides the circle into two semicircles $(\mu_{\mathrm{med}} - \pi, \mu_{\mathrm{med}})$ and $(\mu_{\mathrm{med}}, \mu_{\mathrm{med}} + \pi)$ that each contain equal probability mass of $\frac{1}{2}$; second, as an angle that minimizes the expected circular distance from the random variable,

$$\begin{aligned} \mu_{\mathrm{med}} &:= \arg\min_\phi \mathbb{E}_\theta[D(\theta,\phi)] \\ &= \arg\min_\phi \int_\psi^{\psi+2\pi} D(\theta,\phi) f(\theta)\, d\theta. \end{aligned} \tag{35}$$

As for the sample case, the first definition yields at least two possible solutions; therefore, the second one should be preferred. Note, however, that for certain distributions, the circular population median is not uniquely defined, even under the second definition (e.g., for symmetric, bi-, or multimodal distributions).

The properties and approximations for concentrated data described for sample statistics in the preceding section all generalize to their population counterparts.

Convergence of Sample Statistics to Population Parameters

The trigonometric moments of an independent and identically distributed (i.i.d.) sample, including the mean resultant vector, converge to their population counterparts as the sample size increases (see Section Central Limit Theorem). As a result, sample statistics that are continuous functions of the mean resultant vector (circular mean, variance, and standard deviation) and of higher-order trigonometric moments (circular skewness and kurtosis) inherit convergence via the continuous mapping theorem. The only case where convergence of these parameters is not assured is when the true parameter is not uniquely defined. For example, although we can calculate the circular mean of a sample drawn from the circular uniform distribution, the population mean resultant vector has length 0, so the mean direction is undefined and the sample mean direction has no population counterpart to converge to.

For parameters that are not moment-based, the question of convergence is more complicated. The circular median also converges when the true value is uniquely defined and the density around it is well-behaved, that is, continuous and sufficiently smooth (for a detailed discussion, see Purkayastha, 1995). Convergence results for local circular mode estimators are given in Y. Zhang and Chen (2021).

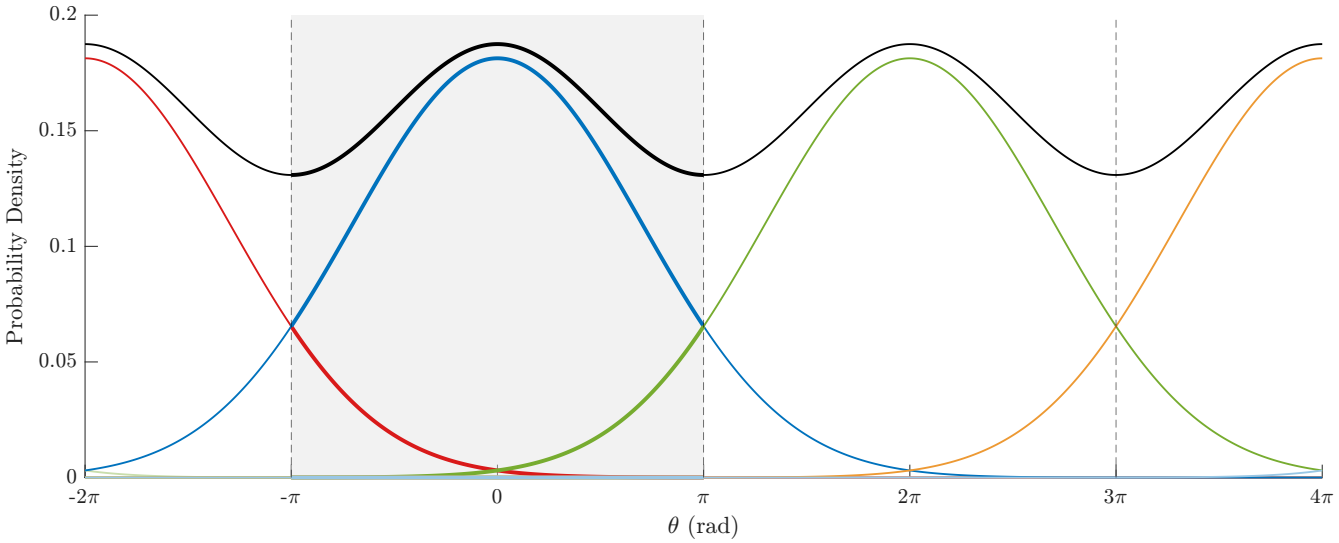
Wrapping Euclidean Distributions

One approach to modeling circular data is to "wrap" a Euclidean distribution around the circle to construct a *wrapped distribution*. For a Euclidean distribution with PDF $f(x)$, the result of wrapping is a circular PDF that can be written as the periodic sum $f_{\mathrm{W}}(\theta) = \sum_{k=-\infty}^{\infty} f(\theta + 2\pi k)$, where the sum is taken over all integers $k \in \mathbb{Z}$. Conveniently, the circular moments of the wrapped distribution are equal to the characteristic function of the unwrapped distribution evaluated at integer arguments, for example., the resultant vector (first circular moment) of a wrapped distribution is given by $\mathrm{m}_1 = \varphi(1)$, where $\varphi(t) = \mathbb{E}[e^{itx}]$ is the characteristic function of the unwrapped distribution $f(x)$. The characteristic function also provides an alternative formulation of the PDF as a Fourier series, which is an infinite sum of sinusoidal components: $f_{\mathrm{W}}(\theta) = \frac{1}{2\pi} \sum_{t=-\infty}^{\infty} \varphi(t) e^{-\mathrm{i}t\theta}$.

In practice, the infinite Fourier representation of a wrapped distribution cannot be evaluated exactly and must be approximated by truncating the series to a finite number of terms. This means evaluating the characteristic function at a limited set of integer points, which produces a smooth approximation of the true circular PDF. The resulting approximation is a finite Fourier series, and its accuracy improves as more terms are included. For an illustration of wrapping a Gaussian distribution, see Figure 14, where the blue line shows a Gaussian distribution in Euclidean space and the black line the same distribution wrapped around the circle. While their familiarity is appealing, the distributions generated by wrapping are often "false friends," proving difficult to work with analytically and lacking in desirable mathematical properties. More tractable "natively circular" distributions can be defined directly in circular space using trigonometric functions.

Figure 14

Wrapping of a Gaussian Distribution



Note. Colored lines show normal (Gaussian) distributions, with standard deviation $\sigma = 2.2$, and means centered on integer multiples of 2π, that is, $\mu = 2\pi k$. The black line shows the wrapped normal distribution with mean 0 and circular standard deviation σ_θ equal to σ, obtained by summing over all translations (for every integer k in the interval $[-5, 5]$, including those not shown) of the Gaussian distribution. The wrapped normal distribution is periodic with period 2π. The signed interval $[-\pi, \pi)$ is highlighted in gray, displaying one period of the wrapped normal distribution. Dashed gray lines indicate period boundaries.

Circular Uniform Distribution

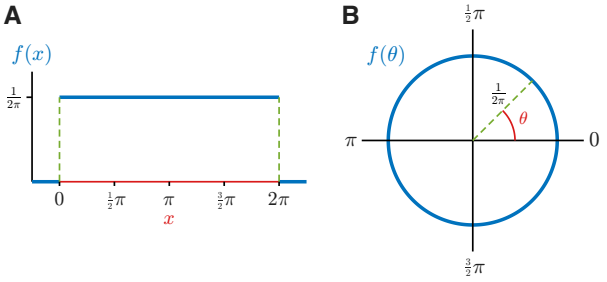
On the real line, a random variable can have a uniform distribution only with respect to a particular bounded interval, with PDF $f(x) = \frac{1}{b-a}$ for x in the interval $[a, b)$ and $f(x) = 0$ outside it (Figure 15 left). While similar distributions can be defined that have uniform probability over a specified range of circular space, the *circular uniform distribution* describes a random variable that has equal probability everywhere, with PDF $f_U(\theta) = \frac{1}{2\pi}$ for all θ (Figure 15 right). The circular uniform distribution has no well-defined mean direction: Its mean resultant is the zero vector, its circular variance is 1, and its circular standard deviation is infinite. The circular uniform distribution has the maximum entropy of any circular distribution.

Circular Analogs of the Normal Distribution

The normal or Gaussian distribution family is fundamental in Euclidean space, underpinning many statistical models due to its well-behaved mathematical properties, including closed-form expression, defined moments, and closure under translation, convolution, and scaling. For circular variables, no single family of distributions possesses all the desirable properties of the Gaussian, and some of its properties are not shared by any family of circular distributions. The closest circular equivalents to the Gaussian are the wrapped normal distribution and the von Mises distribution (sometimes referred to as the circular normal).

Figure 15

Uniform Distribution in Euclidean and Circular Space



Note. (A) Uniform distribution in Euclidean space $f(x)$ over the interval $[0, 2\pi)$. Note that density is 0 outside this interval. (B) Uniform distribution in circular space $f(\theta)$. Equivalent components are displayed in the same color in (A) and (B).

Wrapped Normal Distribution

Wrapping the Gaussian distribution $\mathcal{N}(\mu, \sigma)$ with mean μ and standard deviation σ,

$$f_{\mathrm{N}}(x; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right), \tag{36}$$

onto a unit circle results in the wrapped normal distribution $\mathrm{WN}(\mu, \sigma)$ with PDF given by

$$f_{\mathrm{WN}}(\theta; \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \sum_{k=-\infty}^{\infty} \exp\left(\frac{-(\theta-\mu+2\pi k)^2}{2\sigma^2}\right). \tag{37}$$

The circular mean of the wrapped normal distribution is μ and its circular standard deviation is σ (as discussed in Section Circular Standard Deviation, this motivates the definition of circular standard deviation). Its circular variance is not σ^2 but $1 - \bar{R} = 1 - \exp(\frac{-\sigma^2}{2})$. Its circular kurtosis and skewness are 0.

For small σ, the probability density is concentrated enough that the circularity of the space becomes irrelevant, and the wrapped normal PDF approximates the Gaussian PDF with mean μ and standard deviation σ. As $\sigma \to \infty$ the wrapped normal distribution approaches the circular uniform distribution.

One way in which the wrapped normal distribution can arise is as the result of Brownian diffusion (i.e., accumulation of Gaussian noise) in an angular variable from starting point μ, where the distribution of the variable at time t is wrapped normal with circular standard deviation proportional to $\sqrt{t}$. For a visualization of the wrapped normal distribution and its construction, see Figure 14.

Von Mises Distribution

In practice, the von Mises distribution is often preferred as the circular analog of the normal distribution on the real line. The von Mises distribution $\mathrm{VM}(\mu, \kappa)$ has PDF

$$f_{\mathrm{VM}}(\theta; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp\left(\kappa \cos\left(\theta - \mu\right)\right), \qquad (38)$$

where μ is the circular mean and κ is called the *concentration* of the distribution. Comparing the von Mises with the Gaussian PDF (Equation 36) reveals a very similar functional form: The concentration parameter κ of the von Mises plays an equivalent role to the inverse variance, or precision, of the Gaussian distribution $\frac{1}{\sigma^2}$, and a trigonometric measure of similarity to the mean, $\cos(\theta - \mu)$, replaces the term based on squared error $-\frac{1}{2}(x - \mu)^2$ (Section Measuring Distance).

The function $I_0(\cdot)$ refers to the modified Bessel function of the first kind with order 0. Bessel functions are important mathematically as solutions to certain differential equations, but here $2\pi I_0(\kappa)$ is simply a normalization term that appropriately scales the PDF so it integrates to one over intervals of length 2π.

A von Mises distribution with concentration κ has circular variance $1 - \rho = 1 - A_1(\kappa)$, with $A_1(\kappa) = \frac{I_1(\kappa)}{I_0(\kappa)}$, that is, the ratio of the first- to zeroth-order modified Bessel function of the first kind. Its circular skewness is 0, but its circular kurtosis is greater than 0 for $\kappa > 0$, with a peak value of ~0.834 for $\kappa \approx 2.71$. As $\kappa \to \infty$, the von Mises PDF approaches the Gaussian PDF with mean μ and precision $\frac{1}{\sigma^2}$ equal to κ. For $\kappa = 0$, the von Mises distribution is identical to the circular uniform distribution.

The von Mises is a member of the exponential family of distributions, making it particularly suitable for statistical inference. The maximum likelihood estimator (MLE) of the location parameter μ is the circular mean. In fact, Richard von Mises originally derived the distribution named after him as that distribution for which the sample circular mean is the MLE of location, in the same way as the sample mean is the MLE for the location parameter of the Gaussian (von Mises, 1918).

The von Mises distribution can also be derived from a bivariate Gaussian distribution in the real plane $\mathbb{R}^2$. The von Mises PDF is equal to the conditional density of polar angle θ given unit radial distance $p(\theta | r = 1)$ for a bivariate normal random variable with mean on the unit circle, at $[\cos\mu, \sin\mu]$, equal variances $\frac{1}{\kappa}$ and zero correlation. This approach can be extended to define analogs of the von Mises in higher dimensions, known as von Mises-Fisher distributions: the von Mises-Fisher PDF in $\mathbb{S}^p$ is proportional to the restriction of an isotropic multivariate Gaussian distribution in $\mathbb{R}^{p+1}$ to the unit p-dimensional hypersphere.

Yet another way to obtain the von Mises distribution is as the hitting point on the unit circle of a two-dimensional Brownian diffusion (i.e., an accumulation of isotropic Gaussian noise) with constant drift in direction μ with velocity κ (Gordon & Hudson, 1977). Extending this principle to higher dimensions again leads to the von Mises-Fisher distributions.

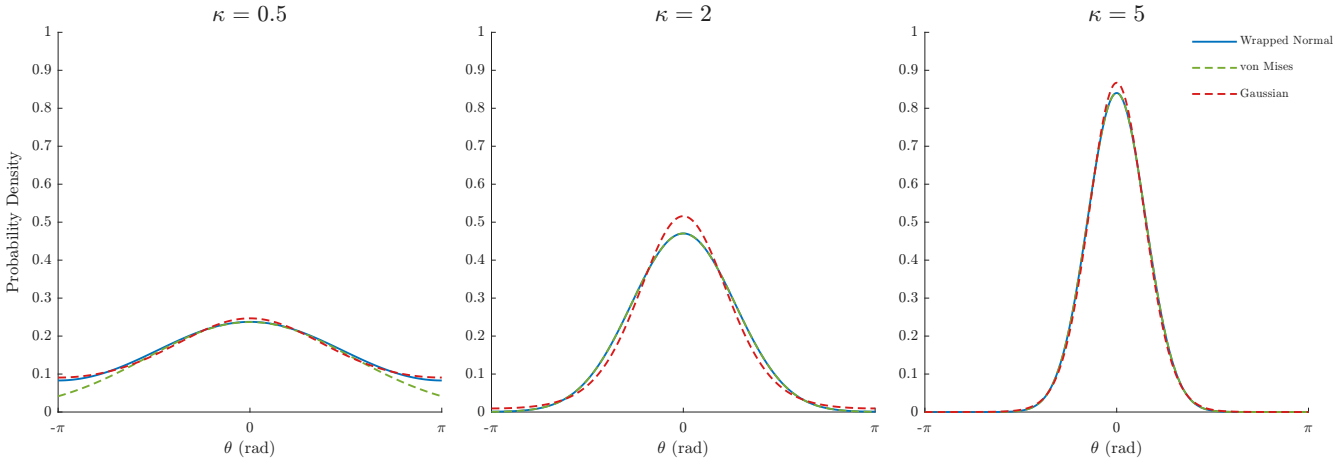
Comparison of Wrapped Normal and von Mises Families

Figure 16 gives a comparison of the Gaussian, wrapped normal, and von Mises distributions with mean 0 and matched (circular) standard deviation (Gumbel et al., 1953). Key properties of these distributions are summarized in Table 2. Unlike the wrapped normal PDF, which involves an infinite sum of Gaussian densities, the PDF of the von Mises distribution has an analytic expression, as do some key quantities such as Fisher information and entropy (most of these involve Bessel functions, but most mathematical software packages have built-in functions for their efficient numerical approximation).

A family of distributions is closed under convolution if the sum of two independent random variables from that family also belongs to the same family (the PDF of the sum is obtained by applying the mathematical operation of convolution to the PDFs of the two variables). The wrapped normal inherits its closure under convolution from its Euclidean counterpart, the Gaussian (Section Sampling Distribution of the Circular Mean below). However, a number of other properties that make the Gaussian distribution special are instead possessed by the von Mises distribution in circular space. Like the Gaussian, the von Mises is a *maximum entropy distribution*. Specifically, among all distributions on the circle with a given circular mean and variance (or equivalently, a given mean resultant vector), the von Mises distribution has the maximum entropy. More similarities and differences between these distributions are described in Section General Properties of Circular Random Variables.

Figure 16

Comparison of Gaussian, Wrapped Normal, and von Mises Distributions



Note. Gaussian, wrapped normal, and von Mises distributions with concentration parameters $\kappa = 0.5, 2.0, 5.0$. Distributions are matched for (circular) standard deviation, using $\sigma = \sqrt{-2 \ln A_1(\kappa)}$ to obtain σ for the wrapped normal and Gaussian distributions. The wrapped normal PDF is approximated using the sum of the first 10 terms of its Fourier series. PDF = probability density function.

Table 2

Properties of the Gaussian, Wrapped Normal, and von Mises Distribution Families

Property	Gaussian	Wrapped normal	von Mises
Support	$x \in \mathbb{R}^1$	$\theta \in \mathbb{S}^1$	$\theta \in \mathbb{S}^1$
Parameters	μ, σ^2	μ, σ	μ, κ
Symmetrical	✓	✓	✓
Unimodal	✓	✓	✓
(Circular) mean/median/mode	μ	μ	μ
circular moments, m_n	(Not applicable)	$e^{in\mu - n^2\sigma^2/2}$	$e^{in\mu} A_n(\kappa)$[a]
Mean resultant length, ρ	(Not applicable)	$e^{-\sigma^2/2}$	$A_1(\kappa)$[a]
(Circular) variance	σ^2	$1 - e^{-\sigma^2/2}$	$1 - A_1(\kappa)$[a]
(Circular) standard deviation	σ	σ	$\sqrt{-2 \ln A_1(\kappa)}$[a]
(Circular) kurtosis	3	0	$\frac{A_2(\kappa) - A_1(\kappa)^4}{(1 - A_1(\kappa))^2}$[a]
Closed under translation[b]	✓	✓	✓
Closed under scaling[b]	✓	✗	✗
Closed under convolution[b]	✓	✓	✗
Fisher information for μ	$\frac{1}{\sigma^2}$	$\sum_{k=-\infty}^{\infty} k^2 e^{-k^2\sigma^2}$	$\kappa A_1(\kappa)$[a]
Entropy (differential)	$\frac{1}{2} \ln (2\pi e \sigma^2)$	(not analytic)	$\ln (2\pi I_0(\kappa)) - \kappa A_1(\kappa)$[a]
Maximum entropy distribution[c]	✓	✗	✓
Exponential family	✓	✗	✓
Sufficient statistics	$(\bar{x}, s^2)$ for (μ, σ^2)	(none-finite)	$(\bar{\theta}, \bar{R})$ for (μ, κ)
Maximum likelihood estimates	$\bar{x}$ (for μ), s^2 (for σ^2)	(not analytic)	$\bar{\theta}$ (for μ), $A_1^{-1}(\bar{R})$ (for κ)[a]
Log-likelihood	Quadratic	(not analytic)	Cosine
Product of PDFs	Scaled Gaussian	(not analytic)	Scaled von Mises

[a] $A_n(\kappa)$ is shorthand for the ratio of Bessel functions $\frac{I_n(\kappa)}{I_0(\kappa)}$. [b] "Closed under X" means that applying operation X to any member of the distribution family produces another member of the family. [c] The Gaussian and von Mises distributions have the maximum entropy for a specified mean and (circular) variance.

Flexible Distribution Families

We next introduce two families of distributions that can serve as flexible models for circular data that are asymmetric or heavy-tailed. Both are four-parameter families of distributions varying in tailedness and skewness as well as location and scale. The *wrapped stable family* generalizes the wrapped normal distribution and retains the property of closure under convolution, meaning that sums of random variables belong to the same family. The sine-skewed Jones-Pewsey family generalizes the von Mises distribution and its PDF has an analytic expression based on trigonometric functions. Particular attention will be given to the wrapped Cauchy distribution, which uniquely belongs to both families.

Wrapped Stable Family

A distribution is called *stable* if linear combinations of random variables have the same distribution up to location and scale. In Euclidean space, stable distributions are important as they generalize the concept of a central limit distribution (Section Central Limit Theorem) to include distributions with nonfinite variance. In a circular space all continuous distributions have finite variance, but wrapping the stable family provides a flexible model for skewed and heavy-tailed data that is also closed under convolution. The wrapped stable PDF can be constructed from its moments, which correspond to the characteristic function of a general stable distribution on the real line,

$$m_k = \varphi(k) = \begin{cases} \exp\left(-(k\gamma)^\alpha + \mathrm{i}k\delta + \mathrm{i}[(k\gamma)^\alpha - k\gamma]\beta\tan\left(\frac{\pi}{2}\alpha\right)\right) & \text{for } \alpha \neq 1, \\ \exp\left(-(k\gamma)^\alpha + \mathrm{i}k\delta - \mathrm{i}k\gamma\log(k\gamma)\beta\frac{2}{\pi}\right) & \text{for } \alpha = 1, \end{cases} \tag{39}$$

$$f_{\mathrm{WS}}(\theta;\alpha,\beta,\gamma,\delta) = \frac{1}{2\pi}\left(1 + 2\sum_{k=1}^{\infty}|m_k|\cos(k\theta - \arg m_k)\right). \tag{40}$$

The above expressions are referred to as the S_0 parametrization by Pewsey (2008). Note that, despite the piecewise definition, necessary because $\tan(\frac{\pi}{2})$ is undefined, the wrapped stable PDF and moments change smoothly through $\alpha = 1$. The mean resultant length is given by $\rho = \exp(-\gamma^\alpha)$.

The stability parameter $\alpha \in (0, 2]$ controls heaviness of the distribution tails, skewness parameter $\beta \in [-1, 1]$ controls the asymmetry, γ is a scale (or dispersion) parameter, and δ is a location parameter. For $\alpha = 2$, the wrapped stable is equivalent to the wrapped normal with $\mu = \delta$ and $\sigma = \sqrt{2}\gamma$ (Figure 17A). As illustrated in Figure 17B–E, larger $|\beta|$ increases skewness (for $\alpha < 2$), smaller α yields heavier tails, and higher γ leads to greater dispersion.

Figure 17A shows some canonical examples of wrapped stable distributions. The wrapped normal ($\alpha = 2, \beta = 0$), the wrapped Holtsmark ($\alpha = 1.5, \beta = 0$) and the wrapped Cauchy ($\alpha = 1, \beta = 0$) are symmetric cases. The standard wrapped Levy distribution ($\alpha = 0.5, \beta = 1$) is right-skewed and exhibits extremely heavy tails, making it suitable for modeling pulse-like behavior on the circle.

Sine-Skewed Jones-Pewsey Family

The Jones-Pewsey family is a generalization of the von Mises distribution on the circle (Jones and Pewsey, 2005), which allows for flexible tailedness and, in its sine-skewed form (Jones and Pewsey, 2012), flexible skewness as well. The sine-skewed Jones-Pewsey PDF is given by

$$f_{\text{SS-JP}}(\theta;\delta,\kappa,\psi,\lambda) = \begin{cases} \dfrac{\cosh^{1/\psi}(\kappa\psi)(1+\tanh(\kappa\psi)\cos(\theta-\delta))^{1/\psi}(1+\lambda\sin(\theta-\delta))}{2\pi P_{1/\psi}(\cosh(\kappa\psi))} & \text{for } \psi \neq 0, \\ \dfrac{\exp(\kappa\cos(\theta-\delta))(1+\lambda\sin(\theta-\delta))}{2\pi I_0(\kappa)} & \text{for } \psi = 0, \end{cases} \tag{41}$$

where $P_\nu(z)$ is the associated Legendre function of the first kind of degree ν and order 1, which can also be expressed in terms of the hypergeometric function ${}_2F_1\left(-\nu, \nu+1; 1; \frac{1-z}{2}\right)$. Despite the piecewise definition, needed to avoid division by 0, the distribution changes smoothly through $\psi = 0$, which corresponds in this case to the sine-skewed von Mises distribution.

The index $\psi \in (-\infty, \infty)$ controls the tailedness of the distribution, $\lambda \in [-1, 1]$ controls skewness, κ is an inverse scale (precision) parameter and μ is a location parameter.

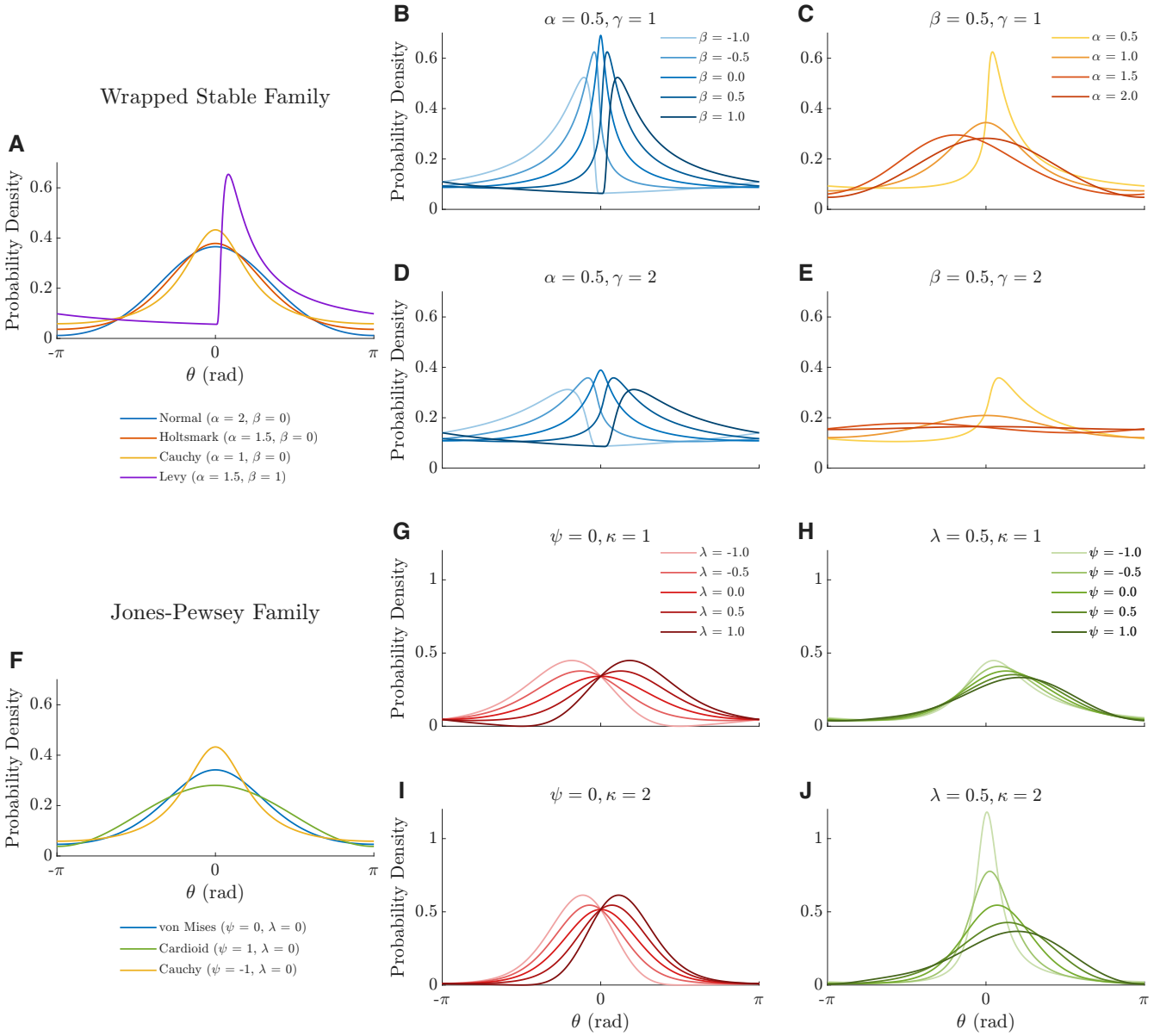
When $\psi < 0$, the distribution has heavier tails than the von Mises distribution, while $\psi > 0$ leads to lighter tails (compare the wrapped stable family, which contains only heavier tailed distributions; Figure 17H and 17J). $\lambda > 0$ leads to right-skew, and conversely for $\lambda < 0$; larger values of $|\lambda|$ lead to stronger skewness (Figure 17G and 17I).

Figure 17F shows some special cases of the symmetric Jones-Pewsey family ($\lambda = 0$), including the von Mises distribution (with $\psi = 0$; mean resultant length is $\rho = A_1(\kappa)$), the Cardioid distribution ($\psi = 1$; $\rho = \tanh(\kappa)$), and the wrapped Cauchy distribution ($\psi = -1$; $\rho = \tanh(\frac{\kappa}{2})$).

In general, the analytic form of the Jones-Pewsey family PDF makes it particularly convenient for modeling and statistical inference. As for the von Mises distribution, the normalization term of the PDF involves a nonelementary function, the associated Legendre, but most mathematical software packages provide efficient methods for its computation as a special case of the hypergeometric function. Although analytically convenient, direct numerical evaluation via the hypergeometric function becomes unreliable when ψ is large and negative. Two practical ways to ensure numerical stability

Figure 17

Examples of Flexible Distribution Families



Note. Example distributions from the wrapped stable family (A–E) and the sine-skewed Jones-Pewsey family (F–J). (A) and (F) show special cases of the wrapped stable and Jones-Pewsey family, respectively. The Cauchy distributions in (A) and (F) are matched in mean resultant length, using $\gamma = -\ln(\tanh(\frac{\kappa}{2}))$ to convert the concentration parameter of the Jones-Pewsey family, κ, to the scale parameter of the Wrapped stable family, γ. Shapes of the wrapped stable family with changing β (B,D) and α (C,E). The PDFs of the wrapped stable family are approximated using the first 100 terms of the Fourier series. Shapes of the Jones-Pewsey family with changing λ (G,I) and ψ (H,J). Location parameters δ are 0 in all cases. PDF = probability density function.

in this regime are computing the Legendre term using more robust algorithms (Gutleb et al., 2022) or bypassing the Legendre representation with a numerical integration approach (see Appendix D).

Wrapped Cauchy Distribution

The wrapped Cauchy distribution belongs to both the wrapped stable family (with $\alpha = 1$, $\beta = 0$; Figure 17A) and the symmetrical Jones-Pewsey family (with $\psi = -1$; Figure 17F). It is closed under convolution like the rest of the wrapped stable family, but its PDF also has a closed-form trigonometric expression as part of the Jones-Pewsey family. Specifically, its PDF can be obtained directly as the wrapping of the Cauchy distribution onto the circle,

$$f_{\mathrm{WC}}(\theta;\mu,\gamma) = \sum_{k=-\infty}^{\infty} \frac{1}{\pi} \frac{\gamma}{(\gamma^2 + (\theta - \mu + 2\pi k)^2}, \tag{42}$$

or equivalently in trigonometric terms as

$$f_{\mathrm{WC}}(\theta;\mu,\gamma) = \frac{1}{2\pi} \frac{\sinh(\gamma)}{\cosh(\gamma) - \cos(\theta - \mu)}. \tag{43}$$

It is parameterized by a location μ (the mean and center of symmetry) and a scale γ. The mean resultant length of the wrapped Cauchy distribution is $\rho = e^{-\gamma}$ and its circular standard deviation is $\sqrt{2\gamma}$. Its circular moments are $m_n = e^{in\mu - |n|\gamma}$. Its differential entropy can be written in closed form as $\log(2\pi(1 - e^{-2\gamma}))$. Note that although the Cauchy distribution on the real line is the classic exemplar of a "pathological" distribution, with mean, variance, and higher moments undefined, its circular counterpart retains some of the heavy-tailedness yet has finite circular moments, making it more tractable for modeling circular data with heavy tails.

General Properties of Circular Random Variables

Additivity of Mean and Variance

Models of Euclidean random variables often make use of properties of the mean and variance, such as additivity, that do not apply or apply differently in circular spaces. The Euclidean mean or expectation is additive in the sense that the mean of the sum of any two random variables is equal to the sum of their means,

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y], \tag{44}$$

and Euclidean variance is additive in the sense that the variance of the sum of two random variables is equal to the sum of their variances, *so long as the variables are uncorrelated*,

$$\mathbb{V}\mathrm{ar}[X + Y] = \mathbb{V}\mathrm{ar}[X] + \mathbb{V}\mathrm{ar}[Y] \quad \text{iff } X \perp\!\!\!\perp Y. \tag{45}$$

These properties are key to understanding how variability from different sources combines and how random noise accumulates (e.g., in drift-diffusion models; Bogacz et al., 2006; Ratcliff, 1978). However, caution is required when extrapolating these properties to circular variables (see Smith, 2016 for a diffusion model of circular data).

Circular equivalents to mean and variance are related to the direction and length of the mean resultant vector, which, for the angular sum of two independent random variables, is the *product* of their individual mean resultant vectors. As a result, the circular mean – although not additive in general – does have the additive property for variables that are uncorrelated,

$$\mu[\theta + \phi] = \mu[\theta] + \mu[\phi] \pmod{2\pi} \quad \text{if } \theta \perp\!\!\!\perp \phi. \tag{46}$$

Circular measures of dispersion are related to the mean resultant length, ρ, for which

$$\rho[\theta + \phi] = \rho[\theta]\rho[\phi] \quad \text{if } \theta \perp\!\!\!\perp \phi, \tag{47}$$

so ρ does not have the additive property, and neither does the circular variance, based on its usual definition $V_\theta = 1 - \rho$. However, the *logarithm* of the mean resultant length, $\log \rho$, *is* additive for independent circular random variables, and (despite not being a usual circular measure of variance) so is the squared circular standard deviation,

$$\sigma^2[\theta + \phi] = \sigma^2[\theta] + \sigma^2[\phi] \quad \text{if } \theta \perp\!\!\!\perp \phi. \tag{48}$$

A more detailed exposition with derivations can be found in Appendix E.

Products of Probability Density Functions

Products of PDFs are important in contexts where information from different sources is combined. These include maximum likelihood estimation and Bayesian inference based on independent samples drawn from the same distribution, as well as sensory fusion and cue combination (Trommershauser et al., 2011), where different probabilistic estimates of a parameter are integrated. These computations take a particularly simple form for Gaussian variables, because the product of Gaussian PDFs is proportional to another Gaussian PDF, with precision $\frac{1}{\sigma^2}$ equal to the sum of the component precisions, and mean equal to the precision-weighted average of the component means.

For circular variables, the von Mises distribution possesses a similar property (Murray and Morgenstern, 2010): the product of von Mises PDFs is proportional to another von Mises PDF with mean and concentration determined by the concentration-weighted resultant vector of the component mean directions (Figure 18A):

$$\begin{gathered} \prod_{i=1}^{n} f_{\mathrm{VM}}(\theta;\mu_i,\kappa_i) = S f_{\mathrm{VM}}(\theta;\arg\bar{r},|\bar{r}|) \\ \bar{r} = \sum_{i=1}^{n} \kappa_i e^{\mathrm{i}\mu_i}, \quad S = \frac{1}{(2\pi)^{n-1}} \frac{I_0(|\bar{r}|)}{\prod_i I_0(\kappa_i)}. \end{gathered} \tag{49}$$

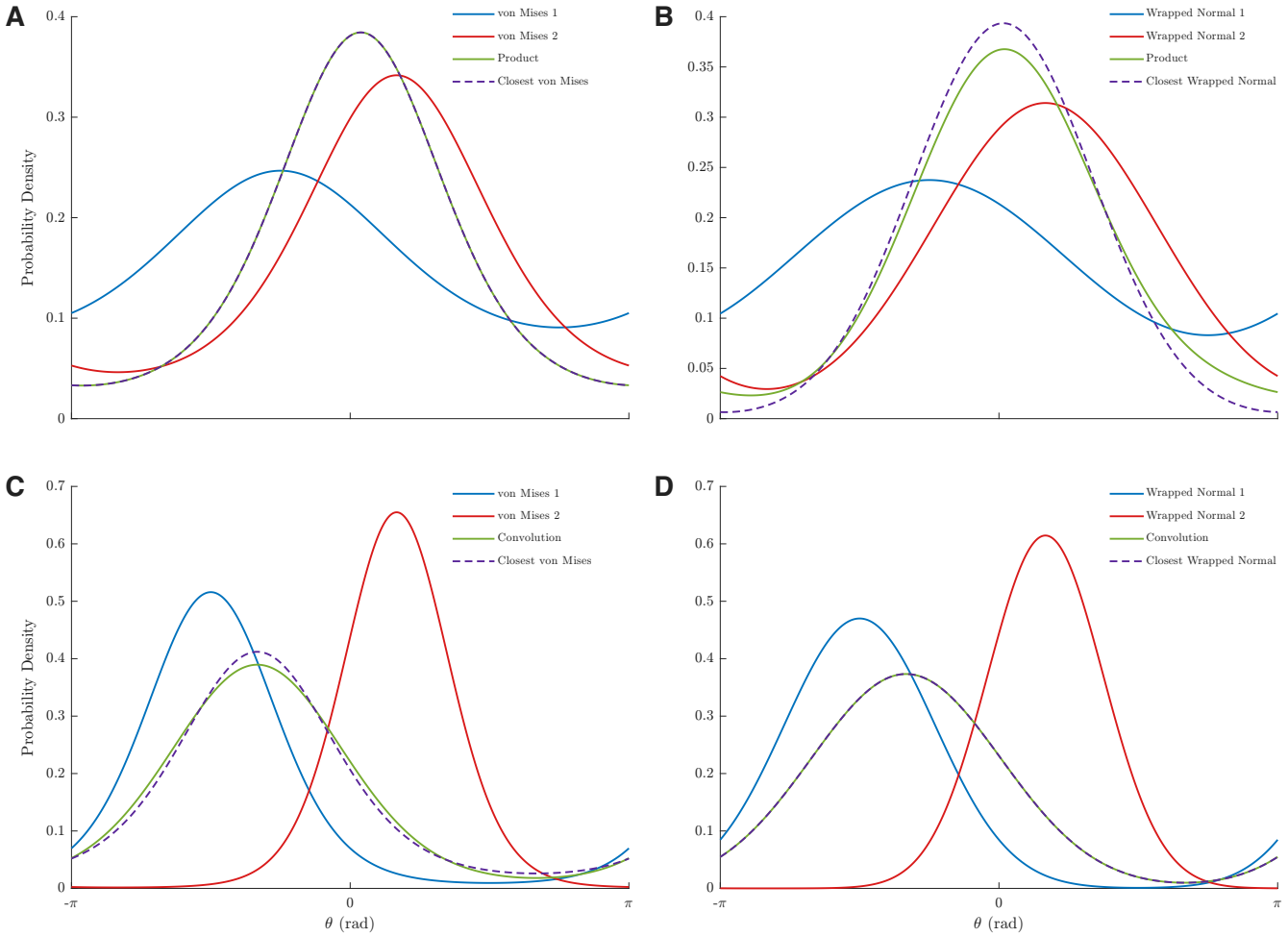
By contrast, the product of two wrapped normal distributions does not yield another wrapped normal distribution, although an approximation based on the Gaussian product formula holds well for small standard deviations, for example, $\sigma < 1$ (Figure 18B).

Sums and Differences of Random Variables

The probability distribution of the sum of two independent random variables $Z = X + Y$ is obtained by convolving their individual PDFs; the distribution of their difference is obtained by reflecting one of the PDFs about 0, then convolving.

Figure 18

Products and Convolutions of von Mises and Wrapped Normal PDFs



Note. (A, B) Products of PDFs of von Mises distributions (A) and wrapped normal distributions (B). Each panel shows two PDFs (red and blue), their normalized product (green), and the closest match to the normalized product by a PDF from the same family (purple, dashed). Note that the product of two von Mises PDFs (A) is proportional to another von Mises PDF, whereas the product of two wrapped normal PDFs (B) does not in general correspond to any wrapped normal PDF. Examples shown are $\mathrm{VM}_1(\mu_1 = -\frac{\pi}{4}, \kappa_1 = 0.5)$ and $\mathrm{VM}_2(\mu_2 = \frac{\pi}{6}, \kappa_2 = 1)$ and wrapped normal PDFs with matching resultant vectors. (C, D) Convolutions of PDFs of von Mises distributions (C) and wrapped normal distributions (D). PDFs shown in green are convolutions of red and blue PDFs, and purple is the closest PDF from the same family. Note that the convolution of the two von Mises distributions (C) is not a von Mises distribution, whereas the convolution of two wrapped normal distributions is another wrapped normal distribution (D). Examples shown are $\mathrm{VM}_1(\mu_1 = -\frac{\pi}{2}, \kappa_1 = 2)$ and $\mathrm{VM}_2(\mu_2 = \frac{\pi}{6}, \kappa_2 = 3)$ and wrapped normal PDFs with matched resultant vectors. Wrapped normal PDFs are approximated using the sum of the first 10 terms of their Fourier series. PDF = probability density function.

These operations are the same for circular variables, where the convolution is of their periodic PDFs over $(-\infty, \infty)$. For Euclidean variables drawn from Gaussian distributions, the sum is always distributed as another Gaussian: We say the Gaussian family is *closed under convolution*. Because convolving two wrapped distributions is identical to wrapping the convolution of the corresponding unwrapped distributions, the wrapped normal distribution inherits this property directly from the normal distribution (Figure 18D),

$$\begin{aligned} &f_{\mathrm{WN}}(\theta; \mu_x, \sigma_x) * f_{\mathrm{WN}}(\theta; \mu_y, \sigma_y) \\ &= f_{\mathrm{WN}}\left(\theta; \mu_x + \mu_y, \sqrt{\sigma_x^2 + \sigma_y^2}\right). \end{aligned} \tag{50}$$

The von Mises distribution is not closed under convolution, which means that the sum (or difference) of von Mises distributed variables is not in general a member of any recognized distribution family (Figure 18C). The PDF for the sum can be calculated as

$$\begin{aligned} &f_{\mathrm{VM}}(\theta; \mu_x, \kappa_x) * f_{\mathrm{VM}}(\theta; \mu_y, \kappa_y) \\ &= \frac{I_0\left(\sqrt{\kappa_x^2 + \kappa_y^2 + 2\kappa_x\kappa_y \cos(\theta - (\mu_x + \mu_y))}\right)}{2\pi I_0(\kappa_x) I_0(\kappa_y).} \end{aligned} \tag{51}$$

Sampling Distribution of the Circular Mean

Because the Gaussian family is closed under both convolution and scaling, the mean of n i.i.d. Gaussian random variables $\bar{X} = \frac{1}{n}\sum_i^n X_i$ is distributed as another Gaussian (with the same mean and variance reduced by a factor of n). This important property is not shared by any family of circular distributions.

The circular equivalent of the mean is the angle of the resultant vector, the vector sum of the unit vectors corresponding to individual random variables. In general, the distribution of the resultant vector angle cannot be obtained independently of its length. The only exception is the circular uniform distribution (Kent et al., 1979), for which the distribution of the resultant angle $\bar{\theta}$ of i.i.d. samples is, trivially, also circular uniform. For two samples from the circular uniform distribution, the distribution of resultant length R is

$$h_2(R) = \frac{2}{\pi\sqrt{4-R^2}} \text{ for } R \in [0,2). \tag{52}$$

For three or more samples, the distribution h_n of resultant length quickly becomes intractable and can only be approximated with numerical methods. An approximation for large n is given in terms of the chi-squared PDF by

$$h_n(R) \approx \frac{4}{n} R f_{\chi^2}\left(\frac{2}{n}R^2; 2\right). \tag{53}$$

For the von Mises distribution, the circular mean of i.i.d. samples is not von Mises distributed, but it can be described as a mixture of von Mises distributions with different concentrations. Specifically, for n samples from $\mathrm{VM}(\mu, \kappa)$, the conditional distribution of resultant angle $\bar{\theta}$ for a given resultant length R is von Mises, with concentration proportional to R,

$$p(\bar{\theta}|R) = f_{\mathrm{VM}}(\bar{\theta}; \mu, \kappa R). \tag{54}$$

Informally, when the sample directions are similar (large R), their mean $\bar{\theta}$ tends to be close to the distribution mean, μ, while when the samples are widely dispersed, the sample mean is too. This contrasts with the Gaussian distribution, for which the sample mean is independent of the sample standard deviation.

The distribution of resultant length is related to the equivalent distribution for the circular uniform by

$$p(R) = \frac{I_0(\kappa R)}{I_0(\kappa)^n} h_n(R). \tag{55}$$

The distribution of the circular mean (the marginal distribution of resultant angle) can then be obtained as the mixture density

$$p(\bar{\theta}) = \int_0^n p(\bar{\theta}|R) p(R)\, dR. \tag{56}$$

For the wrapped normal, the circular mean of i.i.d. samples does not have any recognized distribution, even when conditioned on resultant length.

Central Limit Theorem

The central limit theorem (CLT) states that, under quite general conditions, the sum or mean of a sample of Euclidean random variables converges to a normal (Gaussian) distribution as the number of observations in the sample grows, even if the individual observations do not come from a normal distribution. This fact is often called upon to justify the use of the normal distribution family as a default model for a wide range of behavioral and neural observations, and also underlies many parametric statistical tests. Analogous results to the CLT for circular variables are therefore of considerable interest; however, the topic is complex.

In circular spaces, the limiting distribution depends on precisely what operation is performed on the random variables. It should be plain that the angular sum of increasingly many independent circular random variables, regardless of their individual distributions, will in general grow in dispersion and ultimately converge to the circular uniform distribution. Of more interest is the limiting distribution of the mean resultant vector, the direction of which is the closest circular analog of the Euclidean mean. For i.i.d. circular variables, the mean resultant vector converges to a bivariate Gaussian centered on the mean resultant vector of the distribution they are sampled from, and with covariance that can be calculated from its first two circular moments (Jammalamadaka & SenGupta, 2001, Section 3.4). In particular, if that distribution is symmetrical, the bivariate Gaussian is uncorrelated, so the *tangent* of the resultant direction converges to a Gaussian distribution. If the distribution of resultant direction is narrow enough (which it will be for sufficiently large numbers of samples, as long as they are not uniformly distributed), then it will have approximately the same distribution as its tangent.

Based on this, we can deduce that the limiting distribution of the circular mean of symmetrically distributed circular random variables is approximately Gaussian in shape. Unfortunately, this is not sufficient to identify a specific circular distribution. In particular, both wrapped normal and von Mises PDFs converge to Gaussian as their width approaches 0, so we cannot identify either of these circular analogs of the Gaussian as the unique limiting distribution.

An Equivalent of the Cramér-Rao Bound

The Cramér-Rao Bound (CRB; Rao et al., 1945) is an important result in estimation theory; it describes a limit on the precision of a point estimator $\hat{\theta}(X)$ of an unknown parameter $\theta \in \mathbb{R}^1$ based on data X with likelihood function $p(x|\theta)$. For an unbiased estimator, that is, one with expectation equal to

the true parameter, $\mathrm{E}[\hat{\theta}] = \theta$, the CRB can be expressed as

$$\mathbb{V}\mathrm{ar}[\hat{\theta}] \geq \frac{1}{\mathcal{I}(\theta)}, \tag{57}$$

where the Fisher information $\mathcal{I}(\theta)$ is defined as $\mathcal{I}(\theta) = \mathbb{V}\mathrm{ar}[\frac{\partial}{\partial\theta}\log p(x|\theta)] = -\mathbb{E}[\frac{\partial^2}{\partial\theta^2}\log p(x|\theta)]$. An unbiased estimator that achieves this bound is said to be efficient (an example is the MLE of the mean based on i.i.d. samples drawn from a normal distribution).

An equivalent of the CRB for an estimator $\hat{\theta}(X)$ of a circular parameter $\theta \in \mathbb{S}^1$ that is unbiased in the sense of having a circular mean equal to θ (or equivalently, $\mathbb{E}[\sin(\hat{\theta} - \theta)] = 0$) is given by

$$\frac{\mathbb{V}\mathrm{ar}[\sin(\hat{\theta} - \theta)]}{\rho_{\hat{\theta}}^2} \geq \frac{1}{\mathcal{I}(\theta)}, \tag{58}$$

where $\rho_{\hat{\theta}}$ is the mean resultant length of $\hat{\theta}$ (Mardia, 1975). Note that the left-hand side of this equation differs from the definition of circular variance, $1 - \rho_{\hat{\theta}}$. However, it can equivalently be written in terms of the first and second circular moments of $\hat{\theta}$, as

$$\delta := \frac{1 - |m_2|}{2|m_1|^2}, \tag{59}$$

which is an alternative measure of variability for circular variables, defined as the circular dispersion by Mardia and Jupp (2009, Section 3.4.2). The bound is met in the circular case if and only if $\frac{\partial}{\partial\theta}\log p(x|\theta) \propto \sin(\hat{\theta} - \theta)$, that is, if the likelihood is von Mises. An example is the MLE of the circular mean based on i.i.d. samples from a von Mises distribution with known concentration $\kappa > 0$.

Choosing an Appropriate Circular Distribution

After being introduced to a range of established circular distribution families and their properties, the reader should now be able to choose suitable distributions when building statistical models of circular data. If the data come from a source with no directional bias, the circular uniform distribution should be used. If not, two criteria can be employed for selecting one of the distribution families discussed above: first, whether the data can be expected to have an approximately normal distribution; second, whether the purposes of analysis better suit the properties of wrapped distributions or those defined intrinsically on the circle. If a normal distribution can be assumed (based on inspection or because the conditions described in Section Central Limit Theorem are met), then the von Mises or the wrapped normal is appropriate; if not, and assuming the data are nonetheless approximately unimodal, the Jones-Pewsey or wrapped stable should be chosen for their ability to capture asymmetry and tailedness of data. Wrapped distribution families are more appropriate for data obtained by wrapping Euclidean random variables or when the analysis involves addition or subtraction of circular variables (because these families are closed under convolution). Distributions that are defined intrinsically on the circle are appropriate if the reader needs an analytically and computationally tractable model for statistical inference, such as MLE or Maximum A Posteriori (MAP) estimation, or if they want a closed-form PDF, entropy or Fisher information. If theoretical and practical considerations do not identify a single distribution family, we recommend selecting based on how well the different possible distributions fit the data, noting, however, that differences in fit between wrapped and intrinsically circular distributions are typically small.

Sampling Random Values From Circular Distributions

Sampling from circular distributions is essential for generating synthetic data and experimental stimuli and for use in simulations and model recovery. For wrapped distributions, the usual method is to draw samples from the corresponding distribution in Euclidean space and then wrap the generated random values onto the circle. Distributions that are defined intrinsically on the circle require different sampling methods. For the von Mises distribution, the rejection sampling algorithm introduced by Best and Fisher (1979) has become the standard. This method samples values from an auxiliary distribution on $[0, 2\pi)$ whose density is proportional to $\exp(\kappa\cos\theta)$ (a distribution closely resembling the von Mises), and then uses a rejection criterion to ensure that accepted samples exactly follow the von Mises distribution. This algorithm is implemented, for example, in the function `rvm` in the R package `CircStats` (Lund & Agostinelli, 2002) and in `scipy.stats.vonmises.rvs` in the Python library `SciPy` (Virtanen et al., 2020).

More generally, inverse transform sampling provides a universal framework for generating random variables from any continuous distribution with a known CDF: generate a random value U from a uniform distribution in the interval $[0, 1]$, and return the smallest number X for which $F(X) \geq U$, where F is the CDF of the target distribution. This method can be used for sampling from circular distributions with the CDF defined over any specified interval of length 2π.

Beyond This Tutorial

This tutorial was intended to provide a practical foundation in circular statistics tailored to behavioral and cognitive research. For a comprehensive mathematical introduction to the topic, see the foundational textbooks by Mardia and Jupp (2009) and Jammalamadaka and SenGupta (2001). In this section, we briefly outline topics that we did not cover in the tutorial and point to recommended resources.

This tutorial did not cover circular *inferential statistics*. A wide range of tests assess whether circular data deviate from expected distributions – for an accessible overview,

see Landler et al. (2018). For an introduction to circular-circular correlations, refer to Jammalamadaka and SenGupta (2001, Chapter 8) and Mardia and Jupp (2009, Chapter 3.5.2). There are circular versions of many of the familiar tests for comparing samples. These include the Watson–Williams Test (a circular equivalent to the two-sample two-sided t-test; Watson and Williams, 1956), analyses of variance (ANOVA; Harrison et al., 1986), and multivariate analyses of variance (MANOVA; Landler et al., 2022). Landler et al. (2021) evaluate different statistical tests comparing two independent samples of circular data in terms of Type-I error rate and statistical power. Tasdan and Yeniay (2018) discuss Bonferroni corrections for circular data. For *regression models* of circular variables with circular predictors, there are methodological guidelines for frequentist (p-value) tests (Downs, 2002; Jammalamadaka & SenGupta, 2001, Chapter 8; Cremers et al., 2020), and Bayesian statistics (Jha & Bhuyan, 2021). For a review of methods for principal component analysis of circular and higher-dimensional directional data, we recommend Nodehi et al. (2026).

Clustering algorithms offer a way to identify subgroups or outlier candidates based on angular proximity. There are circular versions of many established clustering algorithms, including circular versions of k-means (Charalampidis, 2005), fuzzy c-means (Kesemen et al., 2016), and the machine-learning algorithm mean-shift (Y. Zhang & Chen, 2021). Machine learning methods have also been adopted to predict (Lang et al., 2020) and simulate (Wessel et al., 2026) circular variables.

Some specific cases of circular data require special considerations to avoid biases and distorted results. While this tutorial focused on continuous circular variables, treating non-continuous data as continuous can substantially bias results (Mardia & Sriram, 2023), even with a fine-grained scale such as rounding to the nearest full degree (Landler et al., 2019). Mardia and Sriram (2023) introduce different families of *circular discrete distributions*, including the circular lattice distribution, which is defined over a finite set of equally spaced points on a circle (Mardia & Jupp, 2009, Chapter 3.5.2). For the classification and time-series modeling of ordinal circular data, refer to Gehlot and Laha (2025) and Jha and Biswas (2018), respectively. Adaptations for grouped or rounded data of tests that assess whether circular data deviate from expected distributions are discussed by Jammalamadaka et al. (2024) for uniformity testing and by Landler et al. (2020) for testing against expected distributions in general.

Incomplete or nonuniform sampling can also bias results; for instance, if a variable is collected only during a constrained time window of a 24-hour cycle (e.g., during daytime). Alsammani et al. (2024) discuss methods to mitigate these measurement biases. For the analysis of circular longitudinal and time-series data, some special considerations are warranted. Refer to Artes and Toloi (2009), Fisher and Lee (1994), Nuñez-Antonio and Gutiérrez-Peña (2014), and Zhan et al. (2025) for guidelines on time-series modeling, as well as to Abuzaid et al. (2014) for outlier-classification methods in circular time-series models.

Other forms of data closely related to circular data are also relevant for specific topics in behavioral and cognitive research but go beyond the scope of this tutorial. Cylindrical data comprise a circular component and a linear component. Cremers and Klugkist (2018) provide an introduction to cylindrical data in behavioral research. For an introduction to circular-linear correlations, refer to Jammalamadaka and SenGupta (2001, Chapter 8) and Mardia and Jupp (2009, Chapter 3.5.2). For regression models of circular variables on linear predictors, there are methodological guidelines for frequentist statistics (Fisher & Lee, 1992; Jammalamadaka & SenGupta, 2001, Chapter 8; Cremers et al., 2020), as well as Bayesian statistics (Cremers et al., 2018; Mulder & Klugkist, 2017). For perspectives on directional data analysis that include spherical and higher-dimensional data, refer to Fisher et al. (1993), Mardia and Jupp (2009), and Pewsey and García-Portugués (2021).

Acknowledgments

S.W. received a PhD scholarship from the Foundation of German Businesses, financed by the German Federal Ministry of Research, Technology, and Space (BMFTR). I.T. was supported by the Croatian Science Foundation (UIP-2025-02-6353). A.T. was funded by the UKRI Biotechnology and Biological Sciences Research Council (BBSRC) through a Doctoral Training Programme. The authors would like to thank Puria Radmard and Julie Lehnert for helpful discussions and comments during preparation of the manuscript and Andrey Chetverikov for valuable feedback on the preprint. The authors declare no conflicts of interest.

References

Abuzaid, A. H., Mohamed, I. B., & Hussin, A. G. (2014). Procedures for outlier detection in circular time series models. *Environmental and Ecological Statistics*, *21*(4), 793–809. https://doi.org/10.1007/s10651-014-0281-8

Abuzaid, A. H., Mohamed, I. B., & Hussin, A. G. (2012). Boxplot for circular variables. *Comput. Stat.*, *27*(3), 381–392. https://doi.org/10.1007/s00180-011-0261-5

Alsammani, A., Stacey, W. C., & Gliske, S. V. (2024). Estimation of circular statistics in the presence of measurement bias. *IEEE Journal of Biomedical and Health Informatics*, *28*(2), 1089–1100. https://doi.org/10.1109/JBHI.2023.3334684

Alzueta, E., Gombert-Labedens, M., Javitz, H., Yuksel, D., Perez-Amparan, E., Camacho, L., Kiss, O., de Zambotti, M., Sattari, N., Alejandro-Pena, A., Zhang, J., Shuster, A., Morehouse, A., Simon, K., Mednick, S., & Baker, F. C. (2024). Menstrual cycle variations in wearable-detected finger temperature and heart rate, but not in sleep metrics, in young and midlife individuals [PMID: 39108015]. *Journal of Biological Rhythms*, *39*(5), 395–412. https://doi.org/10.1177/07487304241265018

Ameijeiras-Alonso, J., Benali, A., Crujeiras, R. M., Rodríguez-Casal, A., & Pereira, J. M. C. (2019). Fire seasonality identification with multimodality tests. *Annals of Applied Statistics*, *13*(4). https://doi.org/10.1214/19-AOAS1273

Artes, R., & Toloi, C. M. C. (2009). An autoregressive model for time series of circular data. *Communications in Statistics - Theory and Methods*, *39*(1), 186–194. https://doi.org/10.1080/03610920802650338

Bays, P. M., Catalao, R. F. G., & Husain, M. (2009). The precision of visual working memory is set by allocation of a shared resource. *Journal of Vision*, *9*(10), 7–7. https://doi.org/10.1167/9.10.7

Best, D. J., & Fisher, N. I. (1979). Efficient simulation of the von Mises distribution. *Applied Statistics*, *28*(2), 152–157. https://doi.org/10.2307/2346732

Blasdel, G. G. (1992). Differential imaging of ocular dominance and orientation selectivity in monkey striate cortex. *Journal of Neuroscience*, *12*(8), 3115–3138. https://doi.org/10.1523/JNEUROSCI.12-08-03115.1992

Blasdel, G. G., & Salama, G. (1986). Voltage-sensitive dyes reveal a modular organization in monkey striate cortex. *Nature*, *321*, 579–585. https://doi.org/10.1038/321579a0

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, *113*(4), 700–765. https://doi.org/10.1037/0033-295X.113.4.700

Bolón, D., Crujeiras, R. M., & Rodríguez-Casal, A. (2025). A likelihood ratio test for circular multimodality. *Environmental and Ecological Statistics*, *32*(1), 57–87. https://doi.org/10.1007/s10651-024-00639-1

Bonhoeffer, T., & Grinvald, A. (1991). Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns [Publisher: Nature Publishing Group]. *Nature*, *353*(6343), 429–431. https://doi.org/10.1038/353429a0

Bonhoeffer, T., Kim, D.-S., Malonek, D., Shoham, D., & Grinvald, A. (1995). Optical imaging of the layout of functional domains in area 17 and across the area 17/18 border in cat visual cortex. *European Journal of Neuroscience*, *7*(9), 1973–1988. https://doi.org/10.1111/j.1460-9568.1995.tb00720.x

Bull, J. R., Rowland, S. P., Scherwitzl, E. B., Scherwitzl, R., Danielsson, K. G., & Harper, J. (2019). Real-world menstrual cycle characteristics of more than 600,000 menstrual cycles. *npj Digital Medicine*, *2*(1), 83. https://doi.org/10.1038/s41746-019-0152-7

Buttarazzi, D., Pandolfo, G., & Porzio, G. C. (2018). A Boxplot for Circular Data. *Biometrics*, *74*(4), 1492–1501. https://doi.org/10.1111/biom.12889

Chang-Chien, S.-J., Hung, W.-L., & Yang, M.-S. (2012). On mean shift-based clustering for circular data. *Soft Computing*, *16*(6), 1043–1060. https://doi.org/10.1007/s00500-012-0802-z

Charalampidis, D. (2005). A modified k-means algorithm for circular invariant clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *27*(12), 1856–1865. https://doi.org/10.1109/TPAMI.2005.230

Chaubey, Y. P. (2022). On nonparametric density estimation for circular data: An overview [Series Title: Forum for Interdisciplinary Mathematics]. In A. SenGupta & B. C. Arnold (Eds.), *Directional Statistics for Innovative Applications* (pp. 351–378). Springer Nature Singapore. https://doi.org/10.1007/978-981-19-1044-9_19

Cremers, J., & Klugkist, I. (2018). One direction? a tutorial for circular data analysis using R with examples in cognitive psychology. *Frontiers in Psychology*, *9*, 2040. https://doi.org/10.3389/fpsyg.2018.02040

Cremers, J., Mainhard, T., & Klugkist, I. (2018). Assessing a Bayesian embedding approach to circular regression models. *Methodology*, *14*(2), 69–81. https://doi.org/10.1027/1614-2241/a000147

Cremers, J., Pennings, H. J. M., & Ley, C. (2020). Regression models for cylindrical data in psychology. *Multivariate Behavioral Research*, *55*(6), 910–925. https://doi.org/10.1080/00273171.2019.1693332

Cremers, J., Pennings, H. J. M., Mainhard, T., & Klugkist, I. (2021). Circular modelling of circumplex measurements for interpersonal behavior. *Assessment*, *28*(2), 585–600. https://doi.org/10.1177/1073191119858407

Cunningham, A. C., Pal, L., Wickham, A. P., Prentice, C., Goddard, F. G. B., Klepchukova, A., & Zhaunova, L. (2024). Chronicling menstrual cycle patterns across the reproductive lifespan with real-world data. *Scientific Reports*, *14*(1), 10172. https://doi.org/10.1038/s41598-024-60373-3

Deverell, M. (2003). Communication, nonhuman. In M. J. Aminoff & R. B. Daroff (Eds.), *Encyclopedia of the neurological sciences* (pp. 746–749). Academic

Press. https://doi.org/https://doi.org/10.1016/B0-12-226870-9/00284-7

Doeller, C. F., Barry, C., & Burgess, N. (2010). Evidence for grid cells in a human memory network. *Nature*, *463*(7281), 657–661. https://doi.org/10.1038/nature08704

Downs, T. D. (2002). Circular regression. *Biometrika*, *89*(3), 683–698. https://doi.org/10.1093/biomet/89.3.683

Drew, T., & Doucet, S. (1991). Application of circular statistics to the study of neuronal discharge during locomotion. *Journal of Neuroscience Methods*, *38*(2-3), 171–181. https://doi.org/10.1016/0165-0270(91)90167-X

Ecochard, R., Leiva, R., Bouchard, T. P., Van Lamsweerde, A., Pearson, J. T., Stanford, J. B., & Gronfier, C. (2024). The menstrual cycle is influenced by weekly and lunar rhythms. *Fertility and Sterility*, *121*(4), 651–659. https://doi.org/https://doi.org/10.1016/j.fertnstert.2023.12.009

Eriksen, C. W., & Colegate, R. L. (1971). Selective attention and serial processing in briefly presented visual displays. *Perception & Psychophysics*, *10*(5), 321–326. https://doi.org/10.3758/BF03207451

Ester, E. F., Ho, T. C., Brown, S. D., & Serences, J. T. (2014). Variability in visual working memory ability limits the efficiency of perceptual decision making. *Journal of Vision*, *14*(4), 2. https://doi.org/10.1167/14.4.2

Fairchild, M. D. (2013). *Color appearance models* (Third edition). Wiley.

Favila, S. E., Kuhl, B. A., & Winawer, J. (2022, October). *Perception and memory have distinct spatial tuning properties in human visual cortex*. https://doi.org/10.1038/s41467-022-33161-8

Fisher, N. I., & Lee, A. J. (1992). Regression models for an angular response. *Biometrics*, *48*(3), 665–677. https://doi.org/10.2307/2532334

Fisher, N. I., & Lee, A. J. (1994). Time series analysis of circular data. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *56*(2), 327–339. https://doi.org/10.1111/j.2517-6161.1994.tb01981.x

Fisher, N. I., Lewis, T., & Embleton, B. J. (1993). *Statistical analysis of spherical data* ([New ed.], 1. paperback ed). Cambridge Univ. Press.
Literaturverz. S. [312] - 320.

Fortenbaugh, F. C., Sanghvi, S., Silver, M. A., & Robertson, L. C. (2012). Exploring the edges of visual space: The influence of visual boundaries on peripheral localization. *Journal of Vision*, *12*(2), 19–19. https://doi.org/10.1167/12.2.19

Foster, J. J., Kirwan, J. D., el Jundi, B., Smolka, J., Khaldy, L., Baird, E., Byrne, M. J., Nilsson, D.-E., Johnsen, S., & Dacke, M. (2019). Orienting to polarized light at night – matching lunar skylight to performance in a nocturnal beetle. *Journal of Experimental Biology*, *222*(2), jeb188532. https://doi.org/10.1242/jeb.188532

Fougnie, D., Suchow, J. W., & Alvarez, G. A. (2012). Variability in the quality of visual working memory. *Nature Communications*, *3*(1), 1229. https://doi.org/10.1038/ncomms2237

Freas, C. A., Narendra, A., Murray, T., & Cheng, K. (2024). Polarised moonlight guides nocturnal bull ants home [Publisher: eLife Sciences Publications, Ltd]. *eLife*. https://doi.org/10.7554/elife.97615.3

Freedman, D., & Diaconis, P. (1981). On the histogram as a density estimator:L 2 theory [Publisher: Springer Science and Business Media LLC]. *Zeitschrift fr Wahrscheinlichkeitstheorie und Verwandte Gebiete*, *57*(4), 453–476. https://doi.org/10.1007/bf01025868

Gehlot, S., & Laha, A. K. (2025). Classification with ordinal circular data. In S. Kumar, B. C. Arnold, K. Shimizu, & A. K. Laha (Eds.), *Directional and Multivariate Statistics* (pp. 87–102). Springer Nature Singapore. https://doi.org/10.1007/978-981-96-2004-3_5

Georgopoulos, A. P., Caminiti, R., Kalaska, J. F., & Massey, J. T. (1983). Spatial Coding of Movement: A Hypothesis Concerning the Coding of Movement Direction by Motor Cortical Populations [Series Title: Experimental Brain Research Supplementum]. In J. Massion, J. Paillard, W. Schultz, & M. Wiesendanger (Eds.), *Neural Coding of Motor Performance* (pp. 327–336, Vol. 7). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-68915-4_34

Georgopoulos, A. P., Kalaska, J. F., Crutcher, M. D., Caminiti, R., & Massey, J. T. (1984). The representation of movement direction in the motor cortex: Single cell and population studies. In G. M. Edelman, W. E. Goil, & W. M. Cowan (Eds.), *Dynamic aspects of neocortical function* (pp. 501–524). Neurosciences Research Foundation Inc.

Georgopoulos, A. P., Schwartz, A. B., & Kettner, R. E. (1986). Neuronal Population Coding of Movement Direction. *Science*, *233*(4771), 1416–1419. https://doi.org/10.1126/science.3749885

Ghosh, P., Chatterjee, D., Banerjee, A., & Das, S. S. (2024). Do magnetic murmurs guide birds? a directional statistical investigation for influence of earth's magnetic field on bird navigation (A. M. Saqr, Ed.). *PLOS ONE*, *19*(6), e0304279. https://doi.org/10.1371/journal.pone.0304279

Giraldo, Y. M., Leitch, K. J., Ros, I. G., Warren, T. L., Weir, P. T., & Dickinson, M. H. (2018). Sun navigation requires compass neurons in drosophila. *Current Bi-*

ology, *28*(17), 2845–2852.e4. https://doi.org/https://doi.org/10.1016/j.cub.2018.07.002

Gombert-Labedens, M., Alzueta, E., Perez-Amparan, E., Yuksel, D., Kiss, O., de Zambotti, M., Simon, K., Zhang, J., Shuster, A., Morehouse, A., Pena, A. A., Mednick, S., & Baker, F. C. (2024). Using wearable skin temperature data to advance tracking and characterization of the menstrual cycle in a real-world setting [PMID: 38767963]. *Journal of Biological Rhythms*, *39*(4), 331–350. https://doi.org/10.1177/07487304241247893

Gordon, L., & Hudson, M. (1977). A characterization of the von Mises distribution. *Annals of Statistics*, *5*(4), 813–814. https://doi.org/10.1214/aos/1176343906

Gumbel, E. J., Greenwood, J. A., & Durand, D. (1953). The circular normal distribution: Theory and tables. *Journal of the American Statistical Association*, *48*(261), 131–152. https://doi.org/10.1080/01621459.1953.10483462

Gurtman, M. B. (2009). Exploring personality with the interpersonalcircumplex. *Social and Personality Psychology Compass*, *3*(4), 601–619. https://doi.org/10.1111/j.1751-9004.2009.00172.x

Gutleb, T., Carrillo, J., & Olver, S. (2022). Computing equilibrium measures with power law kernels. *Mathematics of Computation*, *91*(337), 2247–2281. https://doi.org/10.1090/mcom/3740

Harrison, D., Kanji, G. K., & Gadsden, R. J. (1986). Analysis of variance for circular data. *Journal of Applied Statistics*, *13*(2), 123–138. https://doi.org/10.1080/02664768600000021

Huang, J., Feng, H., Drake, V. A., Reynolds, D. R., Gao, B., Chen, F., Zhang, G., Zhu, J., Gao, Y., Zhai, B., Li, G., Tian, C., Huang, B., Hu, G., & Chapman, J. W. (2024). Massive seasonal high-altitude migrations of nocturnal insects above the agricultural plains of east china. *Proceedings of the National Academy of Sciences*, *121*(18), e2317646121. https://doi.org/10.1073/pnas.2317646121

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, *160*(1), 106–154. https://doi.org/10.1113/jphysiol.1962.sp006837

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, *195*(1), 215–243. https://doi.org/10.1113/jphysiol.1968.sp008455

Hubel, D. H., Wiesel, T. N., & Stryker, M. P. (1977). Orientation columns in macaque monkey visual cortex demonstrated by the 2-deoxyglucose autoradiographic technique. *Nature*, *269*(5626), 328–330. https://doi.org/10.1038/269328a0

Hunter, J. D. (2007). Matplotlib: A 2d graphics environment. *Computing in Science & Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Jammalamadaka, S. R., Ghosh, K., & Akiri, S. (2024). Testing uniformity on the circle using spacings when data are rounded. *Statistics and Probability Letters*, *210*, 110114. https://doi.org/10.1016/j.spl.2024.110114

Jammalamadaka, S. R., & SenGupta, A. (2001, April). *Topics in circular statistics* (Vol. 5). World Scientific. https://doi.org/10.1142/4031

Jha, J., & Biswas, A. (2018). Distribution and time-series modelling of ordinal circular data. *Environmetrics*, *29*(2), e2490. https://doi.org/10.1002/env.2490

Jha, J., & Bhuyan, P. (2021). Two-stage circular-circular regression with zero inflation: Application to medical sciences. *Annals of Applied Statistics*, *15*(3). https://doi.org/10.1214/20-AOAS1429

Jones, M., & Pewsey, A. (2005). A family of symmetric distributions on the circle. *Journal of the American Statistical Association*, *100*(472), 1422–1428. https://doi.org/10.1198/016214505000000286

Jones, M., & Pewsey, A. (2012). Inverse Batschelet distributions for circular data. *Biometrics*, *68*(1), 183–193. https://doi.org/10.1111/j.1541-0420.2011.01651.x

Karoly, P. J., Goldenholz, D. M., Freestone, D. R., Moss, R. E., Grayden, D. B., Theodore, W. H., & Cook, M. J. (2018). Circadian and circaseptan rhythms in human epilepsy: A retrospective cohort study. *The Lancet Neurology*, *17*(11), 977–985. https://doi.org/https://doi.org/10.1016/S1474-4422(18)30274-6

Kempter, R., Leibold, C., Buzsáki, G., Diba, K., & Schmidt, R. (2012). Quantifying circular–linear associations: Hippocampal phase precession. *Journal of Neuroscience Methods*, *207*(1), 113–124. https://doi.org/10.1016/j.jneumeth.2012.03.007

Kent, J. T., Mardia, K. V., & Rao, J. S. (1979). A characterization of the uniform distribution on the circle. *Annals of Statistics*, *7*(4), 882–889. https://doi.org/10.1214/aos/1176344737

Kesemen, O., Tezel, Ö., & Özkul, E. (2016). Fuzzy c-means clustering algorithm for directional data (FCM4DD). *Expert Systems with Applications*, *58*, 76–82. https://doi.org/https://doi.org/10.1016/j.eswa.2016.03.034

Kondo, N., Sekijima, T., Kondo, J., Takamatsu, N., Tohya, K., & Ohtsu, T. (2006). Circannual control of hibernation by HP complex in the brain. *Cell*, *125*(1), 161–172. https://doi.org/10.1016/j.cell.2006.03.017

Koul, A., Ahmar, D., Iannetti, G. D., & Novembre, G. (2023). Interpersonal synchronization of spontaneously generated body movements. *iScience*, *26*(3), 106104. https://doi.org/10.1016/j.isci.2023.106104

Landler, L., Ruxton, G. D., & Malkemper, E. P. (2018). Circular data in biology: Advice for effectively implementing statistical procedures. *Behavioral Ecology and Sociobiology*, *72*(8), 128. https://doi.org/10.1007/s00265-018-2538-y

Landler, L., Ruxton, G. D., & Malkemper, E. P. (2019). Circular statistics meets practical limitations: A simulation-based Rao's spacing test for non-continuous data. *Movement Ecology*, *7*(1), 15. https://doi.org/10.1186/s40462-019-0160-x

Landler, L., Ruxton, G. D., & Malkemper, E. P. (2020). Grouped circular data in biology: Advice for effectively implementing statistical procedures. *Behavioral Ecology and Sociobiology*, *74*(8), 100. https://doi.org/10.1007/s00265-020-02881-6

Landler, L., Ruxton, G. D., & Malkemper, E. P. (2021). Advice on comparing two independent samples of circular data in biology. *Scientific Reports*, *11*(1), 20337. https://doi.org/10.1038/s41598-021-99299-5

Landler, L., Ruxton, G. D., & Malkemper, E. P. (2022). The multivariate analysis of variance as a powerful approach for circular data. *Movement Ecology*, *10*(1), 21. https://doi.org/10.1186/s40462-022-00323-8

Lang, M. N., Schlosser, L., Hothorn, T., Mayr, G. J., Stauffer, R., & Zeileis, A. (2020). Circular regression trees and forests with an application to probabilistic wind direction forecasting. *Journal of the Royal Statistical Society Series C: Applied Statistics*, *69*(5), 1357–1374. https://doi.org/10.1111/rssc.12437

Lautenschlager, L., Souza, Y., Villar, N., Galetti, M., & Feeley, K. J. (2024). Communal tapir latrines are foraging sites for tropical forest vertebrates. *Global Ecology and Conservation*, *52*, e02950. https://doi.org/10.1016/j.gecco.2024.e02950

Leary, T. (1957). *Interpersonal diagnosis of personality: A functional theory and methodology for personality evaluation*. Ronald Press.

Lee Rodgers, J., Beasley, W. H., & Schuelke, M. (2014). Graphical data analysis on the circle: Wrap-around time series plots for (interrupted) time series designs. *Multivariate Behavioral Research*, *49*(6), 571–580. https://doi.org/10.1080/00273171.2014.946589

Lemon, J. (2006). plotrix: A package in the red light district of R. *R-News*, *6*(4), 8–12.

Lindecke, O., Elksne, A., Holland, R. A., Pētersons, G., & Voigt, C. C. (2019). Experienced migratory bats integrate the sun's position at dusk for navigation at night [Publisher: Elsevier BV]. *Current Biology*, *29*(8), 1369–1373.e3. https://doi.org/10.1016/j.cub.2019.03.002

Liu, Y., & West, S. G. (2016). Weekly cycles in daily report data: An overlooked issue. *Journal of Personality*, *84*(5), 560–579. https://doi.org/10.1111/jopy.12182

Lund, U., & Agostinelli, C. (2002, February). CircStats: Circular statistics, from "topics in circular statistics" [Institution: Comprehensive R Archive Network Pages: 0.2-7]. https://doi.org/10.32614/CRAN.package.CircStats

Ma, W. J., Husain, M., & Bays, P. M. (2014). Changing concepts of working memory. *Nature Neuroscience*, *17*(3), 347–356. https://doi.org/10.1038/nn.3655

Macpherson, M. C., Brown, A. J., Kallen, R. W., Richardson, M. J., & Miles, L. K. (2024). Partner gaze shapes the relationship between symptoms of psychopathology and interpersonal coordination. *Scientific Reports*, *14*(1), 14288. https://doi.org/10.1038/s41598-024-65139-5

Maldonado, P. E., Gödecke, I., Gray, C. M., & Bonhoeffer, T. (1997). Orientation Selectivity in Pinwheel Centers in Cat Striate Cortex. *Science*, *276*(5318), 1551–1555. https://doi.org/10.1126/science.276.5318.1551

Mardia, K. V., & Jupp, P. E. (2009). *Directional statistics*. John Wiley & Sons.

Mardia, K. V., & Sriram, K. (2023). Families of discrete circular distributions with some novel applications. *Sankhya A*, *85*(1), 1–42. https://doi.org/10.1007/s13171-022-00298-z

Mardia, K. V. (1975). Statistics of directional data. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *37*(3), 349–371. https://doi.org/10.1111/j.2517-6161.1975.tb01550.x

May, K. A., & Zhaoping, L. (2009). Effects of surrounding frame on visual search for vertical or tilted bars. *Journal of Vision*, *9*(13), 20–20. https://doi.org/10.1167/9.13.20

Meyer, C., Muto, V., Jaspar, M., Kussé, C., Lambot, E., Chellappa, S. L., Degueldre, C., Balteau, E., Luxen, A., Middleton, B., Archer, S. N., Collette, F., Dijk, D.-J., Phillips, C., Maquet, P., & Vandewalle, G. (2016). Seasonality in human cognitive brain responses. *Proceedings of the National Academy of Sciences*, *113*(11), 3066–3071. https://doi.org/10.1073/pnas.1518129113

Mitre-Ortiz, A., & Mitre-Hernandez, H. (2019). Study of spontaneous and acted learn-related emotions through facial expressions and galvanic skin response. *Research in Computing Science*, *148*(5), 97–105. https://doi.org/10.13053/rcs-148-5-11

Moore, J. J., Cushman, J. D., Acharya, L., Popeney, B., & Mehta, M. R. (2021). Linking hippocampal multiplexed tuning, Hebbian plasticity and navigation.

Nature, *599*(7885), 442–448. https://doi.org/10.1038/s41586-021-03989-z

Mulder, K., & Klugkist, I. (2017). Bayesian estimation and hypothesis tests for a circular generalized linear model. *Journal of Mathematical Psychology*, *80*, 4–14. https://doi.org/10.1016/j.jmp.2017.07.001

Murray, R. F., & Morgenstern, Y. (2010). Cue combination on the circle and the sphere. *Journal of Vision*, *10*(11), Article 15.

Nau, M., Navarro Schröder, T., Frey, M., & Doeller, C. F. (2020). Behavior-dependent directional tuning in the human visual-navigation network. *Nature Communications*, *11*(1), 3247. https://doi.org/10.1038/s41467-020-17000-2

Nelson, N., Malhan, D., Hesse, J., Aboumanify, O., Yalçin, M., Lüers, G., & Relógio, A. (2025). Comprehensive integrative analysis of circadian rhythms in human saliva. *npj Biological Timing and Sleep*, *2*(1), 17. https://doi.org/10.1038/s44323-025-00035-3

Nodehi, A., Moghimbeygi, M., & Ley, C. (2026). Recent advances in principal component analysis for directional data. *Journal of Multivariate Analysis*, *212*, 105528. https://doi.org/10.1016/j.jmva.2025.105528

Noorman, M., Hulse, B. K., Jayaraman, V., Romani, S., & Hermundstad, A. M. (2024). Maintaining and updating accurate internal representations of continuous variables with a handful of neurons [Publisher: Nature Publishing Group]. *Nature Neuroscience*, *27*(11), 2207–2217. https://doi.org/10.1038/s41593-024-01766-5

Nuñez-Antonio, G., & Gutiérrez-Peña, E. (2014). A Bayesian model for longitudinal circular data based on the projected normal distribution. *Computational Statistics and Data Analysis*, *71*, 506–519. https://doi.org/10.1016/j.csda.2012.07.025

O'Keefe, J., & Recce, M. L. (1993). Phase relationship between hippocampal place units and the eeg theta rhythm. *Hippocampus*, *3*(3), 317–330. https://doi.org/https://doi.org/10.1002/hipo.450030307

Olivers, C. N. (2025). Selective attention and eccentricity: A comprehensive review. *Neuroscience & Biobehavioral Reviews*, *178*, 106368. https://doi.org/10.1016/j.neubiorev.2025.106368

Olson, S. J., & Grossberg, S. (1998). A neural network model for the development of simple and complex cell receptive fields within cortical maps of orientation and ocular dominance1technical report cas/cns-tr-96-021.1. *Neural Networks*, *11*(2), 189–208. https://doi.org/https://doi.org/10.1016/S0893-6080(98)00003-3

Packmor, F., Kishkinev, D., Zechmeister, T., Mouritsen, H., & Holland, R. A. (2024). Migratory birds can extract positional information from magnetic inclination and magnetic declination alone. *Proceedings of the Royal Society B: Biological Sciences*, *291*(2034), rspb20241363. https://doi.org/10.1098/rspb.2024.1363

Palmeirim, A. F., Araújo-Fernandes, A. C., Castro-Fernandes, A. S., Guedes, P., dos Santos, Y., Alves, J. C., Mata, V., Yoh, N., & Rocha, R. (2025). Habitat use and diel activity of insectivorous bats across land-use types on an afrotropical oceanic island. *bioRxiv*, 2025.04.10.648189. https://doi.org/10.1101/2025.04.10.648189

Panagopoulos, D. J., Karabarbounis, A., & Chrousos, G. P. (2024). Biophysical mechanism of animal magnetoreception, orientation and navigation [Publisher: Springer Science and Business Media LLC]. *Scientific Reports*, *14*(1). https://doi.org/10.1038/s41598-024-77883-9

Patke, A., Young, M. W., & Axelrod, S. (2020). Molecular mechanisms and physiological importance of circadian rhythms. *Nature Reviews Molecular Cell Biology*, *21*(2), 67–84. https://doi.org/10.1038/s41580-019-0179-2

Paul, M. J., Zucker, I., & Schwartz, W. J. (2008). Tracking the seasons: The internal calendars of vertebrates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1490), 341–361. https://doi.org/10.1098/rstb.2007.2143

Perkins, S. M., Amematsro, E. A., Cunningham, J., Wang, Q., & Churchland, M. M. (2025). An emerging view of neural geometry in motor cortex supports high-performance decoding (M. W. Mathis & J. I. Gold, Eds.). *eLife*, *12*, RP89421. https://doi.org/10.7554/eLife.89421

Pewsey, A. (2008). The wrapped stable family of distributions as a flexible model for circular data. *Computational Statistics & Data Analysis*, *52*(3), 1516–1523. https://doi.org/10.1016/j.csda.2007.04.017

Pewsey, A., & García-Portugués, E. (2021). Recent advances in directional statistics. *TEST*, *30*(1), 1–58. https://doi.org/10.1007/s11749-021-00759-x

Prinzmetal, W., Amiri, H., Allen, K., & Edwards, T. (1998). Phenomenology of attention: I. Color, location, orientation, and spatial frequency. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(1), 261.

Purkayastha, S. (1995). An almost sure representation of sample circular median. *Journal of Statistical Planning and Inference*, *46*(1), 77–91. https://doi.org/10.1016/0378-3758(94)00092-A

Rao, C. R., et al. (1945). Information and the accuracy attainable in the estimation of statistical parameters. *Bul-*

letin of the Calcutta Mathematical Society, *37*(3), 81–91.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*(2), 59.

Refinetti, R. (2020). Circadian rhythmicity of body temperature and metabolism. *Temperature*, *7*(4), 321–362. https://doi.org/10.1080/23328940.2020.1743605

Ringach, D. L., Shapley, R. M., & Hawken, M. J. (2002). Orientation selectivity in macaque v1: Diversity and laminar dependence. *Journal of Neuroscience*, *22*(13), 5639–5651. https://doi.org/10.1523/JNEUROSCI.22-13-05639.2002

Sadikaj, G., Rappaport, L. M., Moskowitz, D. S., Zuroff, D. C., Koestner, R., & Powers, T. (2015). Consequences of interpersonal spin on couple-relevant goal progress and relationship satisfaction in romantic relationships. *Journal of Personality and Social Psychology*, *109*(4), 722–737. https://doi.org/10.1037/pspp0000043

Salmela, V. R., Lahde, M., & Saarinen, J. (2012). Visual working memory for amplitude-modulated shapes. *Journal of Vision*, *12*(6), 2–2. https://doi.org/10.1167/12.6.2

Sanderson, D. J., & Peacock, D. C. P. (2020). Making rose diagrams fit-for-purpose. *Earth-Science Reviews*, *201*, 103055. https://doi.org/10.1016/j.earscirev.2019.103055

Saper, C. B., Scammell, T. E., & Lu, J. (2005). Hypothalamic regulation of sleep and circadian rhythms. *Nature*, *437*(7063), 1257–1263. https://doi.org/10.1038/nature04284

Schanda, J., & International Commission on Illumination (Eds.). (2007). *Colorimetry: Understanding the CIE system*. CIE/Commission internationale de l'eclairage ; Wiley-Interscience.

Schneegans, S., & Bays, P. M. (2017). Neural architecture for feature binding in visual working memory. *The Journal of Neuroscience*, *37*(14), 3913–3925. https://doi.org/10.1523/JNEUROSCI.3493-16.2017

Schulz-Hildebrandt, H. (2025). When Purple Perceived Only at Fixation: A Fixation and Distance-Dependent Color Illusion [Version Number: 3]. https://doi.org/10.48550/ARXIV.2509.11582
Other 5 pages, 6 figures, 3 visualizations (Figshare).

Schürch, R., & Couvillon, M. J. (2013). Too much noise on the dance floor [PMID: 23750292]. *Communicative & Integrative Biology*, *6*(1), e22298. https://doi.org/10.4161/cib.22298

Scott, D. W. (1979). On optimal and data-based histograms [Publisher: Oxford University Press (OUP)]. *Biometrika*, *66*(3), 605–610. https://doi.org/10.1093/biomet/66.3.605

Scott, D. W. (2010). Scott's rule [Publisher: Wiley]. *WIREs Computational Statistics*, *2*(4), 497–502. https://doi.org/10.1002/wics.103

Seeber, M., Stangl, M., Vallejo Martelo, M., Topalovic, U., Hiller, S., Halpern, C. H., Langevin, J.-P., Rao, V. R., Fried, I., Eliashiv, D., & Suthana, N. (2025). Human neural dynamics of real-world and imagined navigation. *Nature Human Behaviour*, *9*(4), 781–793. https://doi.org/10.1038/s41562-025-02119-3

Sides, K., Kilungeja, G., Tapia, M., Kreidl, P., Brinkmann, B. H., & Nasseri, M. (2023). Analyzing physiological signals recorded with a wearable sensor across the menstrual cycle using circular statistics. *Frontiers in Network Physiology*, *3*, 1227228. https://doi.org/10.3389/fnetp.2023.1227228

Skaggs, W. E., McNaughton, B. L., Wilson, M. A., & Barnes, C. A. (1996). Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus*, *6*(2), 149–172. https://doi.org/https://doi.org/10.1002/(SICI)1098-1063(1996)6:2<149::AID-HIPO6>3.0.CO;2-K

Smith, P. L. (2016). Diffusion theory of decision making in continuous report. *Psychological Review*, *123*(4), 425–451. https://doi.org/10.1037/rev0000023

Staugaard, C. F., Petersen, A., & Vangkilde, S. (2016). Eccentricity effects in vision and attention. *Neuropsychologia*, *92*, 69–78. https://doi.org/10.1016/j.neuropsychologia.2016.06.020

Stouffer, P. C., Johnson, E. I., & Bierregaard, R. O. (2013). Breeding seasonality in central amazonian rainforest birds. *The Auk*, *130*(3), 529–540. https://doi.org/10.1525/auk.2013.12179

Sturges, H. A. (1926). The Choice of a Class Interval [Publisher: Informa UK Limited]. *Journal of the American Statistical Association*, *21*(153), 65–66. https://doi.org/10.1080/01621459.1926.10502161

Tasdan, F., & Yeniay, O. (2018). A comparative simulation of multiple testing procedures in circular data problems. *Journal of Applied Statistics*, *45*(2), 255–269. https://doi.org/10.1080/02664763.2016.1273886

Tendler, A., Bar, A., Mendelsohn-Cohen, N., Karin, O., Korem Kohanim, Y., Maimon, L., Milo, T., Raz, M., Mayo, A., Tanay, A., & Alon, U. (2021). Hormone seasonality in medical records suggests circannual endocrine circuits. *Proceedings of the National Academy of Sciences*, *118*(7), e2003926118. https://doi.org/10.1073/pnas.2003926118

Thompson, P. R., Harrington, P. D., Mallory, C. D., Lele, S. R., Bayne, E. M., Derocher, A. E., Edwards, M. A., Campbell, M., & Lewis, M. A. (2024). Simultaneous estimation of the temporal and spatial extent of animal migration using step lengths and

turning angles. *Movement Ecology*, *12*(1), 1. https://doi.org/10.1186/s40462-023-00444-8

Trommershauser, J., Kording, K., & Landy, M. S. (2011). *Sensory cue integration*. Oxford University Press.

Ts'o, D. Y., Frostig, R. D., Lieke, E. E., & Grinvald, A. (1990). Functional organization of primate visual cortex revealed by high resolution optical imaging. *Science*, *249*(4967), 417–420. https://doi.org/10.1126/science.2165630

van den Berg, R., Shin, H., Chou, W.-C., George, R., & Ma, W. J. (2012). Variability in encoding precision accounts for visual short-term memory limitations. *Proceedings of the National Academy of Sciences*, *109*(22), 8780–8785. https://doi.org/10.1073/pnas.1117465109

van der Lans, R. M., Cremers, J., Klugkist, I., & Zwart, R. (2020). Teachers' interpersonal relationships and instructional expertise: How are they related? *Studies in Educational Evaluation*, *66*, 100902. https://doi.org/10.1016/j.stueduc.2020.100902

Vijayan, S., & Somanathan, H. (2024). Cog in the migration wheel: Migratory waggle dances are necessary but not sufficient in the nest movement of the asian giant honeybee apis dorsata. *Insectes Sociaux*, *71*(2), 197–209. https://doi.org/10.1007/s00040-024-00968-1

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., ... SciPy 1.0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2

von Mises, R. (1918). Über die "Ganzzahligkeit" der Atomgewicht und verwandte Fragen. *Physikalische Zeitschrift*, *19*, 490–500.

Vowels, M. J., Mark, K. P., Vowels, L. M., & Wood, N. D. (2018). Using spectral and cross-spectral analysis to identify patterns and synchrony in couples' sexual desire (A. R. Dalby, Ed.). *PLOS ONE*, *13*(10), e0205330. https://doi.org/10.1371/journal.pone.0205330

Watson, G. S., & Williams, E. J. (1956). On the construction of significance tests on the circle and the sphere. *Biometrika*, *43*(3/4), 344–352. https://doi.org/10.1093/biomet/43.3-4.344

Wells, N. A. (2000). Are there better alternatives to standard rose diagrams? *Journal of Sedimentary Research*, *70*(1), 37–46. https://doi.org/10.1306/2DC408FC-0E47-11D7-8643000102C1865D

Wessel, J. B., Murphy-Barltrop, C. J. R., & Simpson, E. S. (2026). Generative machine learning for multivariate angular simulation. *Extremes*, *29*(2), 195–243. https://doi.org/10.1007/s10687-025-00522-7

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wiggins, J. S. (1991). Agency and communion as conceptual coordinates for the understanding and measurement of interpersonal behavior. In D. Cicchetti & W. M. Grove (Eds.), *Thinking clearly about psychology: Essays in honor of paul e. meehl* (pp. 89–113). University of Minnesota Press.

Wilken, P., & Ma, W. J. (2004). A detection theory account of change detection. *Journal of Vision*, *4*(12), 11. https://doi.org/10.1167/4.12.11

Xu, D., & Wang, Y. (2020). Area-proportional visualization for circular data. *Journal of Computational and Graphical Statistics*, *29*(2), 351–357. https://doi.org/10.1080/10618600.2019.1654881

Zhan, X., Ma, T., Liu, S., & Shimizu, K. (2025). Circular data diagnostics in regression and time series models. In S. Kumar, B. C. Arnold, K. Shimizu, & A. K. Laha (Eds.), *Directional and Multivariate Statistics* (pp. 3–23). Springer Nature Singapore. https://doi.org/10.1007/978-981-96-2004-3_1

Zhang, L.-Q., Mao, J., Aguirre, G. K., & Stocker, A. A. (2025). The tilt illusion arises from an efficient reallocation of neural coding resources at the contextual boundary. *Proceedings of the National Academy of Sciences*, *122*(17), e2421565122. https://doi.org/10.1073/pnas.2421565122

Zhang, Y., & Chen, Y.-C. (2021). Kernel smoothing, mean shift, and their learning theory with directional data. *Journal of Machine Learning Research*, *22*(154), 1–92. http://jmlr.org/papers/v22/20-1194.html

Zheng, J., Yebra, M., Schjetnan, A. G. P., Patel, K., Katz, C. N., Kyzar, M., Mosher, C. P., Kalia, S. K., Chung, J. M., Reed, C. M., Valiante, T. A., Mamelak, A. N., Kreiman, G., & Rutishauser, U. (2024). Theta phase precession supports memory formation and retrieval of naturalistic experience in humans. *Nature Human Behaviour*, *8*(12), 2423–2436. https://doi.org/10.1038/s41562-024-01983-9

Zisapel, N. (2018). New perspectives on the role of melatonin in human sleep, circadian rhythms and their regulation. *British Journal of Pharmacology*, *175*(16), 3190–3199. https://doi.org/10.1111/bph.14116

Appendix A
List of Symbols

Table A1

List of Symbols

Symbol	Description
$\mathbb{R}^p$	p-dimensional Euclidean space.
$\mathbb{S}^p$	p-dimensional directional space.
$\mathbb{S}^1$	Circular (one-dimensional directional) space.
x, y (Set: X, Y)	Random variables.
θ (Set: Θ)	Polar angle on a circle.
$[x, y]$	Pair of rectangular coordinates. Used to represent circular values in a Cartesian coordinate system.
(r, θ)	Polar vector representation. r is the radial coordinate, indicating the length of the vector; θ is the angular coordinate, specifying the direction of the vector.
O	The origin, i.e., point where all coordinates are 0, e.g., $[0, 0]$ in $\mathbb{R}^2$.
i	The imaginary unit, $\sqrt{-1}$.
z	A unit complex number.
$\oplus$ and $\ominus$	Modular addition and subtraction in a circular space.
$D(\theta, \phi)$	Circular distance between the angles θ and ϕ.
κ	Concentration parameter of a von Mises distribution ($\frac{1}{\kappa}$ is analogous to σ^2).
σ^2	Euclidean population variance.
$\widehat{m}_p$	p-th circular moment of a sample.
$\mathbf{R}$	Resultant vector of a sample.
$\widehat{m}_1$	Sample mean resultant vector, i.e., sample-size–adjusted resultant vector.
R	The resultant length of a sample, i.e., the length of the resultant vector.
$\bar{R}$	Sample mean resultant length, i.e., the length of the mean resultant vector.
$\bar{\theta}$	Sample circular mean.
θ_{med}	Sample circular median.
θ_{mode}	Sample circular mode.
$\widehat{V}_\theta$	Sample circular variance.
s_θ	Sample circular standard deviation.
g_1	Sample Euclidean skewness.
$\mathrm{g}_{1\theta}$	Sample circular skewness.
b_2	Sample Euclidean kurtosis.
g_2	Sample Euclidean excess kurtosis.
$\mathrm{g}_{2\theta}$	Sample circular kurtosis.
m_p	p-th circular moment of a population.
m_1	Population mean resultant vector.
μ	Population circular mean.
μ_{med}	Population circular median.
μ_{mode}	Population circular mode.
V_θ	Population circular variance.
σ_θ	Population circular standard deviation.
$\gamma_{1\theta}$	Population circular skewness.
$\gamma_{2\theta}$	Population circular kurtosis.
ρ	Population mean resultant vector length.
φ_k	Characteristic function of a distribution.
$f_{\text{N}}(x; \mu, \sigma)$ or $\mathcal{N}(\mu, \sigma^2)$	The Gaussian (normal) distribution.
$f_{\text{U}}(\theta)$	The circular uniform distribution.
$f_{\text{VM}}(\theta; \mu, \kappa)$	The von Mises distribution (circular normal).
$f_{\text{WN}}(\theta; \mu, \sigma)$	The wrapped normal distribution.

Symbol	Description
$f_{\text{WC}}(\theta; \mu, \gamma)$	The wrapped Cauchy distribution with scale parameter γ.
$f_{\text{WS}}(\theta; \alpha, \beta, \gamma, \delta)$	The wrapped stable family with stability parameter α, skewness parameter β, scale parameter γ, and location parameter δ.
$f_{\text{SS-JP}}(\theta; \delta, \kappa, \psi, \lambda)$	The sine-skewed Jones-Pewsey family with location parameter δ, concentration κ, shape parameter ψ, and skewness parameter λ.
$I_n(\kappa)$	The modified Bessel function of the first kind of order n.
$A_n(\kappa)$	The ratio of Bessel functions $\frac{I_n(\kappa)}{I_0(\kappa)}$.
$\mathbb{E}[X]$	The expectation of random variable X.
$\mathbb{V}\text{ar}[X]$	The variance of random variable X.
$\mathcal{I}(\theta)$	Fisher information of a circular variable θ.

Appendix B
List of Abbreviations

Table B1

List of Abbreviations

Symbol	Description
ANOVA	Analysis of variance
a.u.	Arbitrary units
CDF	Cumulative distribution function
CLT	Central limit theorem
CRB	Cramér-Rao Bound
eCDF	Empirical cumulative distribution function
EEG	Electroencephalography
ePMF	Empirical probability mass function
Hz	Hertz; unit for frequency, measuring cycles per second
i.i.d.	Independent and identically distributed
IPC	Interpersonal Circumplex
LFP	Local field potential
MANOVA	Multivariate analyses of variance
MAP	Maximum A Posteriori (estimation)
MLE	Maximum likelihood estimator
M1	Primary motor cortex
PDF	Propability density function
PMF	Probability mass function
PSTH	Peristimulus time histograms
Q1–4	Quartile one to four
rad	Radians
V1	Primary visual cortex

Appendix C
Sampling Circular Values With Minimum Separation

Sampling values from a circular uniform distribution while enforcing a minimum angular separation between samples can be achieved using the following algorithm. The function takes as input the number of samples N, the minimum angular separation minDistance, and the lower and upper bound of the circular space.

Algorithm 1 RandMinDistance(N, minDistance, lower, upper)

Require: Number of samples N, minimum separation minDistance

if lower, upper is undefined **then**
  lower $\leftarrow -\pi$, upper $\leftarrow \pi$ ▶ default circular range
end if
range $\leftarrow$ upper $-$ lower
$l \leftarrow$ range $- N \times$ minDistance
if $l < 0$ **then**
  error: "Range too small to draw N samples with given minimum distance."
end if
▶ Generate N random values within the remaining free length
for $i = 1$ **to** N **do**
  $r[i] \leftarrow$ Uniform(0, l)
end for
$(s, \text{idx}) \leftarrow$ Sort(r) ▶ Sort and save original indices
▶ Enforce minimum distance
for $i = 1$ **to** N **do**
  $s[i] \leftarrow s[i] + (i - 1) \times$ minDistance
end for
$r \leftarrow$ Order(s, idx) ▶ Restore original order (optional)
▶ Apply global random rotation and wrap onto the circle
$u \leftarrow$ Uniform(0, range)
for $i = 1$ **to** N **do**
  $r[i] \leftarrow$ lower + $\big((r[i] + u - \text{lower}) \bmod \text{rangeLength}\big)$
end for

return r

First, this algorithm computes the total free length available after enforcing the minimum spacing requirement. It then samples N random values within this remaining space. After sorting the samples, a minimum separation of at least minDistance of adjacent values is ensured by applying a cumulative offset of $(i - 1) \times$ minDistance. Optionally, the original order of the values can be restored – for instance, when experimental stimuli are presented sequentially and should not appear in a sorted order. Finally, the list of values is rotated by a random offset and wrapped back onto the circular range, yielding a uniform random arrangement of points while ensuring the specified minimum separation.

The separation constraints and the maximum number of samples are inversely related: for a minimum separation distance Δ_{min}, the number of samples n is constrained by $n \leq \left\lfloor \frac{R}{\Delta_{\min}} \right\rfloor$. Once the separation constraints are specified, angular distance (Section Measuring Distance) can be computed pairwise to all data points to verify that every pair of circular observations meets the requirements.

Appendix D
Jones-Pewsey PDF via Numerical Integration

In Equation 41, the normalizing constant of the Jones–Pewsey distribution is expressed as an associated Legendre function, which becomes numerically unstable when ψ is large and negative. The algorithm below avoids this by computing the density using numerical integration of the kernel and working in log space to prevent floating-point overflow. When $\psi = 0$, the distribution reduces to the von Mises and its analytic normalizing constant is used instead.

Algorithm 2 JonesPewseyPDF($\theta, \mu, \kappa, \psi$)

Require: Angles θ, location parameter μ, concentration κ, shape parameter ψ
if $|\psi| < 10^{-6}$ **then**
  $f \leftarrow \exp(\kappa \times \cos(\theta - \mu)) \;/\; 2 \times \pi \times I_0(\kappa)$ ▶ Special case: von Mises
else
  core $\leftarrow \cosh(\kappa \times \psi) + \sinh(\kappa \times \psi) \times \cos(\theta - \mu)$ ▶ Compute kernel argument at each angle
  if core ≤ 0 for any θ **then**
    error: "Non-positive JP kernel; check parameter regime."
  end if
  $logg \leftarrow \frac{1}{\psi} \times \log(\text{core})$ ▶ Compute log-kernel
  $logg \leftarrow logg - \max(logg)$ ▶ Subtract maximum for numerical stability
  $g \leftarrow \exp(logg)$ ▶ Transform to linear scale
  $f \leftarrow g \;/\; \text{trapezoidalIntegral}(\theta,\, g)$ ▶ Normalize numerically
end if
return f

To calculate the sine-skewed Jones-Pewsey distribution, using the same approach, the following wrapper function can be used:

Algorithm 3 SineSkewedJonesPewseyPDF($\theta, \mu, \kappa, \psi, \lambda$)

Require: Angles θ, location parameter μ, concentration κ, shape parameter ψ, skewness $\lambda \in [-1, 1]$
$f_{\text{sym}} \leftarrow \text{JonesPewseyPDF}(\theta, \mu, \kappa, \psi)$ ▶ Evaluate Jones–Pewsey base density
$f \leftarrow (1 + \lambda \times \sin(\theta - \mu)) \times f_{\text{sym}}$ ▶ Apply sine-skewing
$f \leftarrow f \;/\; \text{trapezoidalIntegral}(\theta,\, f)$ ▶ Renormalize numerically (safeguard)
return f

Appendix E
Additivity of Mean and Variance

The Euclidean mean or expectation is additive in the sense that the mean of the sum of any two random variables is equal to the sum of their means,

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y], \tag{E1}$$

and Euclidean variance is additive in the sense that the variance of the sum of two random variables is equal to the sum of their variances, *so long as the variables are uncorrelated*,

$$\begin{aligned} \mathbb{V}\mathrm{ar}[X + Y] &= \mathbb{V}\mathrm{ar}[X] + \mathbb{V}\mathrm{ar}[Y] + 2\,\mathbb{C}\mathrm{ov}[X, Y] \\ &= \mathbb{V}\mathrm{ar}[X] + \mathbb{V}\mathrm{ar}[Y] \quad \text{iff } X \perp\!\!\!\perp Y \end{aligned} \tag{E2}$$

Circular equivalents to mean and variance are related to the direction and length of the mean resultant vector, which for the angular sum of two variables θ and ϕ is given by

$$\begin{aligned} m_{\theta+\phi} &= \mathbb{E}[e^{\mathrm{i}(\theta+\phi)}] = \mathbb{E}[e^{\mathrm{i}\theta} e^{\mathrm{i}\phi}] \\ &= \mathbb{E}[e^{\mathrm{i}\theta}]\mathbb{E}[e^{\mathrm{i}\phi}] + \mathbb{C}\mathrm{ov}[e^{\mathrm{i}\theta}, e^{\mathrm{i}\phi}] \\ &= m_\theta m_\phi \quad \text{iff } \theta \perp\!\!\!\perp \phi. \end{aligned} \tag{E3}$$

that is, the mean resultant vector of the sum of two independent random angles is the *product* of their individual mean resultant vectors.

We confine ourselves now to the case of independent θ and ϕ. The circular mean is the direction or argument of the mean resultant vector,

$$\begin{aligned} \mu[\theta + \phi] &= \arg m_{\theta+\phi} = \arg(m_\theta m_\phi) \\ &= \arg m_\theta + \arg m_\phi \pmod{2\pi} \\ &= \mu[\theta] + \mu[\phi] \pmod{2\pi} \end{aligned} \tag{E4}$$

using $\arg(z_1 z_2) = \arg z_1 + \arg z_2$ for any complex z_1, z_2. In other words, the circular mean is additive for independent variables. Circular measures of dispersion are related to the mean resultant length,

$$\rho[\theta + \phi] = |m_{\theta+\phi}| = |m_\theta m_\phi| = |m_\theta||m_\phi| = \rho[\theta]\rho[\phi] \tag{E5}$$

using $|z_1 z_2| = |z_1||z_2|$ for any complex z_1, z_2. In other words, the mean resultant length does not have the additive property, and by extension neither does the circular variance, based on its usual definition $V_\theta = 1 - \rho$. However,

$$\log \rho[\theta + \phi] = \log \rho[\theta] + \log \rho[\phi] \tag{E6}$$

that is, the logarithm of mean resultant length is additive for independent circular random variables. Based on the definition of circular standard deviation,

$$\sigma_\theta = \sqrt{-2 \log \rho} \implies \sigma_\theta^2 = -2 \log \rho \tag{E7}$$

it follows that squared circular standard deviation is also additive for independent variables,

$$\sigma_\theta^2[\theta + \phi] = \sigma_\theta^2[\theta] + \sigma_\theta^2[\phi] \quad \text{if } \theta \perp\!\!\!\perp \phi. \tag{E8}$$