 Copyright 2021, Sjøkrigsskolen

Peer-reviewed article

Banana Distributions Based on Stochastic Polar Coordinates

Tore Langholm and Harald Totland
Norwegian Defence University College, Naval Academy, Bergen, Norway

A tracking algorithm may employ a probability distribution for object locations at a given time after the last observation. In the two-dimensional case, such a distribution on geographical locations typically takes a curved, oblong shape, a so-called "banana distribution."

One recently proposed version of the banana distribution assigns probabilities to the possible future locations of a moving object for which the original location, speed and direction are known, while subsequent movement is determined by a constant, unknown acceleration and a constant, unknown turn rate with independent, zero-mean Gaussian distributions.

Finding the probabilities of locations in such a distribution can be computationally demanding, since there are no functional, closed forms to recover the acceleration and turn rate that would have taken the object to a particular geographical location. In this paper we propose an approximation to the banana distributions above based on polar coordinates, where the angle out to the next observation is normally distributed with expectation zero, while the distance out to the next observation is also normally distributed, but with parameters given by two functions on the angle. This framework is computationally far more tractable, and we show that remarkably good approximations to the original banana distributions can be achieved by fine-tuning of the two functions on the angle. Finally, we show how to incorporate the additional assumption of an unknown initial direction with a Gaussian distribution.

We believe that the framework proposed here, which is closely related to the shape of the distribution, may serve as a suitable framework in which to compare the various other banana-like distributions obtained from different types of assumptions.

Keywords: tracking, probability distribution

This article is a modified and extended version of the paper Approximating Banana Distributions Based on Constant, Unknown Acceleration and Turn Rate, presented at the IEEE conference Global Oceans 2020 (DOI: 10.1109/ieeeconf38699.2020.9389326).

1 Introduction

Association algorithms employed in tracking of multiple moving objects will in some manner or other utilize a probability distribution for the location of an object at a given time after the last observation. The resulting two-dimensional probability distribution on geographical coordinates will typically assume a curved, oblong shape, giving rise to the informal concept of a "banana distribution." The precise nature of the mathematical distribution used to model the situation will depend on the original speed, elapsed time, turning characteristics, and more generally on the type of assumptions being made.

Long et al. considered a robot moving initially at a certain speed in a certain direction and with a certain curvature (or turn radius) but with its two wheels governed by continuous, stochastic processes introducing uncertainty on the subsequent movement [1]. The probability density function for its location after a certain time was approximated with a probability model where the object is assumed to be moving at constant but unknown speed and with a constant but unknown curvature. A banana distribution of the resulting type, based on constant, unknown speed and curvature with given distribution, is computationally straightforward, since the speed and curvature resulting in a given position at a particular time are easily recovered by functional, closed-form expressions.

Gade et al. proposed a somewhat different probability model in which the initial speed and direction are known, while the subsequent movement is determined by a constant, unknown acceleration and a constant, unknown turn rate (rather than turn radius), producing a so-called CAT-distribution [2]. Such assumptions can be appropriate in connection with frequent observations of moving objects, where it is reasonable to assume near constant acceleration between observations, and with objects that typically turn a constant angle per unit of time rather than per unit of distance. When the acceleration and turn rate are taken to be independent and normally distributed variables with expectation zero, this also produces a "banana shaped" distribution, but in this case the probability density values at geographical coordinates require more computation power to obtain, since there are no functional, closed forms to recover the acceleration and turn rate from the geographical position.

In situations of association and tracking, which is the primary motivation for our work, it is often crucial to have quick, real time recourse to estimates for the probability densities at a number of different locations, simultaneously for a large number of different, hypothetical scenarios. Such applications pose very high demands for the efficiency of the individual estimates. When every such estimate involves an iterative search algorithm to obtain approximate values for the generating, stochastical variables, the situation is less than ideal, and one is led to ponder if it might be possible to replace the whole model with one that is built up in a totally different way, offering functional lookup for probability density values directly from geographical coordinates, but with a net behavior closely approximating the original model.

In the present article we consider how to approximate the CAT-distribution with a model which assumes that the angle out to the next observation, relative to the original direction of movement, is normally distributed with expectation zero, and that the distance out to the next observation is also normally distributed, but with the expected value and standard deviation given by two functions on the angle. We believe that such a model, which is more directly related to the shape of the distribution, may serve as a suitable framework in which to compare the various banana-like distributions obtained from different types of assumptions, but presently we employ this only to formulate an approximation to the CAT-distribution which behaves remarkably well for a wide range of parameter settings.

2 Basic concepts

In the CAT-model introduced by Gade et al. [2], an object is moving at a known, initial speed v_0 with an unknown, constant tangential acceleration a, in a known, original direction and with an unknown, constant turn rate ω. A time parameter t represents the time lapse from an initial observation to the next, and has a fixed value in any given model. When the original position and direction are represented by the origin and the positive y-axis, the position of the object at time t will have the coordinates

$$[x, y] = \frac{v_0}{\omega}\left[1 - \cos(\omega t),\ \sin(\omega t)\right] + \frac{a}{\omega^2}\left[\sin(\omega t) - \omega t\cos(\omega t),\ \cos(\omega t) + \omega t\sin(\omega t) - 1\right]. \tag{1}$$

Now when the two unknowns a and ω are taken to be independent, normally distributed random variables with expectation 0 and standard deviations σ_a and σ_ω, respectively, that is,

$$(a, \omega) \sim \mathcal{N}(0, \sigma_a) \times \mathcal{N}(0, \sigma_\omega)\,,$$

one obtains a probability model with the four parameters v_0, t, σ_a and σ_ω. A contour plot for the resulting probability density function on geographical coordinates is shown in Figure 1.

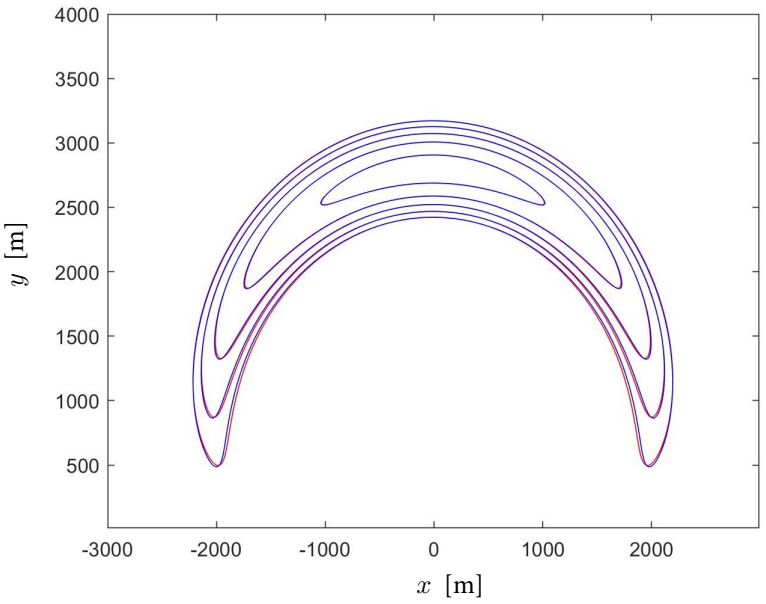

Figure 1: Contour plot shown in blue for the CAT distribution with parameter values $v_0 = 280\,\mathrm{m/s}$, $t = 10\,\mathrm{s}$, $\sigma_a = \frac{5}{3}\mathrm{m/s^2}$ and $\sigma_\omega = \frac{\pi}{54}\mathrm{rad/s} = \frac{10}{3}\,\mathrm{deg/s}$, which are those used by [2] in most of the examples. Blue contour lines are drawn for density values between $10^{-6}\mathrm{m}^{-2}$ and $10^{-10}\mathrm{m}^{-2}$. The y-axis represents the original direction of movement. A rough, first approximation is shown in red.

Other parameter settings than those from Figure 1 yield contour plots with differently shaped bananas, but, with parameter settings going beyond certain bounds, the ba-

nana resemblance goes away. With a significant chance that the object has, by the time of the next observation, come to a complete stop and begun to reverse, the curves will lose their oblong shape, and, similarly, with a significant chance that the object has come through a very large turn of direction, say 270 degrees, the bananas start to merge at the ends. In both cases, the proposed approximations start to fail. Exactly *when* they fail is an empirical question explored in Section 6, but we identify the crucial figures here. The object has come to a stop if $at \leq -v_0$, which happens with significant chance if $\frac{\sigma_a t}{v_0}$ is larger than some small number of choice around 0.25. Similarly, the change of direction ωt has a significant chance of exceeding 270° when $\sigma_\omega t$ is larger than some number of choice around $0.25 \cdot 270^\circ$. In Section 6 one will see that these bounds correspond roughly with the intervals considered in Figure 7.

For any geographical position with Cartesian coordinates (x, y) let (φ, r) be the corresponding polar coordinates, i.e., the angle atan2 (x, y) relative to the positive y-axis and the distance $\sqrt{x^2 + y^2}$ from the origin. It can be shown that the first component vector on the right side in (1) has an angle of exactly $\frac{\omega t}{2}$ while the second component will have an angle of approximately $\frac{2}{3}\omega t$ when $|\omega t| < \pi$. When $|\frac{at}{2}|$ is small in comparison to v_0, the first component will dominate, and φ will be close to $\frac{\omega t}{2}$. A first approximation to the CAT-model using polar coordinates as the basic, random variables, will therefore assume that φ is normally distributed with zero mean and standard deviation $\frac{\sigma_\omega t}{2}$. Next it can be seen that the length of the first component vector on the right side in (1) is exactly[1] $v_0 t \operatorname{sinc}\left(\frac{\omega t}{2}\right)$, while the second can be shown to have a length of approximately $\left|\frac{1}{2}at^2\right|$ for moderate values of ωt, and will add or retract to the total length. A first approximation will therefore assume that the length r has an expected value of $v_0 t \operatorname{sinc}\varphi$ when the angle is φ. As for the standard deviation of r given φ, note that $r = v_0 t + \frac{1}{2}at^2$ when the turn rate is zero, and hence the standard deviation is $\frac{1}{2}\sigma_a t^2$ in this case. Assuming for simplicity that r given φ is normally distributed with the same standard deviation for all φ, one obtains the following rough approximation.

ROUGH APPROXIMATION

1. $\varphi \sim \mathcal{N}\left(0, \frac{1}{2}\sigma_\omega t\right)$
2. $r|\varphi \sim \mathcal{N}\left(v_0 t \operatorname{sinc}\varphi, \frac{1}{2}\sigma_a t^2\right)$

To give an impression of the goodness of this first approximation, the contour lines in the resulting probability density function were drawn in red in a layer underneath the proper CAT-distribution in Figure 1, using the same parameter values. The innermost, red contours are invisible in the figure, as they are completely covered by the blue.

Table 1 compares the probability density values and the approximations at geographical locations within the five blue contour curves, and lists the maximal error rate found within each curve together with the proportion of the probability mass within that curve.

Table 1: Probability mass and maximal error rate within each of the contours of Figure 1.

Contour at density level [m^{-2}]	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}
Total probability mass inside	58%	96%	99.6%	99.96%	99.997%
Maximal error within contour	1.4%	8%	26%	57%	108%

At this level the approximation does not look all bad, but there is considerable room for improvement. We approach this matter with a detailed look at the individual parts of

[1] $\operatorname{sinc} x$ denotes the nonnormalized function $\frac{\sin x}{x}$.

the rough approximation, first the approximate distribution of the angle φ and then the approximate distribution of the distance r.

3 Approximate distribution of angle

Under the CAT-distribution, the angle φ is a non-linear function of the acceleration a and turn-rate ω, and therefore not normally distributed. We show in the appendix that when ωt stays within $\pm\pi$ and $|at|$ is small compared to v_0, φ is roughly equal to the product of the two independent variables $\frac{\omega t}{2}$ and $1 + \frac{at}{6v_0}$. This product has zero expectation and standard deviation $\frac{\sigma_\omega t}{2}\sqrt{1 + (\frac{\sigma_a t}{6v_0})^2}$. Since the factor variables are normally distributed, the product, under the assumption of a small $\frac{\sigma_a t}{v_0}$ ratio, behaves approximately as a normally distributed variable [3]. When $\frac{\sigma_a t}{v_0} < 0.2$, the last factor in the standard deviation is approximately $1 + \frac{1}{72}(\frac{\sigma_a t}{v_0})^2 < 1.0006$, and can be ignored when a very high level of precision is not needed. In the next, improved approximation, part 1 therefore remains unchanged.

4 Approximate distribution of distance

The conditional distribution of r when φ is given, is better understood in the light of some observations made in the appendix, where we show that for "moderate values",

$$\varphi \approx \frac{\omega t}{2}\left(1 + \frac{at}{6v_0}\right), \tag{2}$$

$$r \approx v_0 t \operatorname{sinc}\varphi + \frac{at^2}{6}\left(4\operatorname{sinc}\varphi - \cos\varphi\right). \tag{3}$$

Thus, φ behaves approximately as the product of the two independent, normal variables $\frac{\omega t}{2}$ and $1 + \frac{at}{6v_0}$. In [4] we consider the conditional distribution of the factor variables when the product of two independent, normal variables is known, and present results about the distribution of $X|XY = z$ when X, Y are independent, normal variables, Y has zero mean, X has a low coefficient of variation $\frac{\sigma_X}{\mu_X}$, and z is not too large compared to $\mu_X\sigma_Y$.

Now, there are three components to the approximation of r's probability distribution; the use of the normal distribution and the individual approximations of the mean value and standard deviation. The three are discussed individually.

4.1 Normal distribution

Of course, r is not in fact normally distributed, since it does not take negative values, but the examples examined so far have been remarkably consistent with such an assumption. Figure 2 shows the numerically computed probability density of r for $\varphi = 0.4\pi$ drawn in blue. The normal distribution (density) with matching mean and standard deviation was drawn in red in the same figure, but is not visible behind the blue. In the interval between 1800 and 2400 m in the figure, the two computed probability densities differ by at most a quarter of a percentage of their values.

In fact, the normal distribution is no surprise in the light of [4], where, under the assumptions above, we show that the conditional distribution of X when the product XY is given, is approximately normal. Thus, (2) shows that the conditional distribution

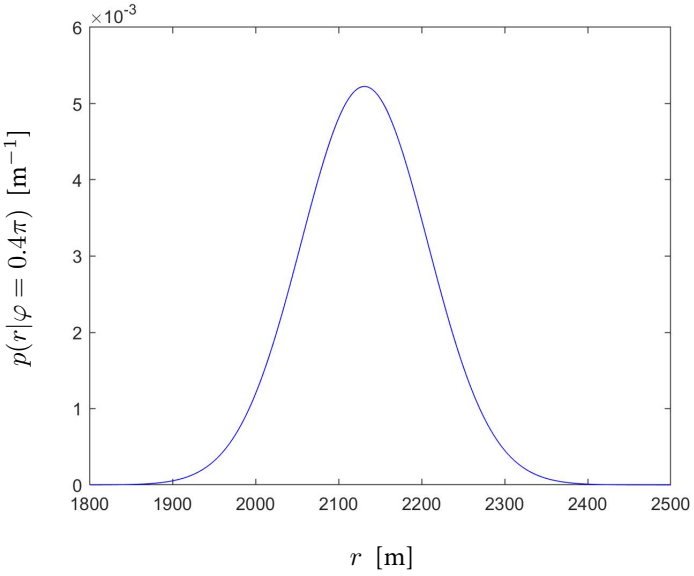

Figure 2: Numerically computed conditional probability density of r for $\varphi = 0.4\pi$, with the parameter values from Figure 1.

of the random variable $1 + \frac{at}{6v_0}$ when φ is given, is approximately normal. Therefore, so is $a|\varphi$ and, from (3), $r|\varphi$ as well.

4.2 Expected value

The approximation $E[r|\varphi] \approx v_0 t \operatorname{sinc} \varphi$ used above, stipulates a decrease in the expected value of r as φ increases in absolute value. This is because a large $|\varphi|$ implies a curved trajectory and hence a net displacement significantly shorter than the trajectory itself. There is, however, another factor in play; large absolute values of φ may in part be the result of a large acceleration, which again implies a longer trajectory. Conversely, a small $|\varphi|$ may in part be the result of a negative acceleration, with the opposite effect. So while the expected value of r does in fact decrease with higher absolute values of φ, the effect is smaller than implied by the rough approximation. Figure 3 shows the actual, numerically computed[2] expected values of r in blue and the function $v_0 t \operatorname{sinc} \varphi$ in red, in the running example with $v_0 = 280$, $t = 10$, $a \sim \mathcal{N}(0, \frac{5}{3})$ and $\omega \sim \mathcal{N}(0, \frac{\pi}{54})$. The difference between the curves is about 24 meters at the end of the diagram.

A better estimate for the blue curve in Figure 3 can be achieved from (2) and (3) in combination with a result in [4], where we show that under the above assumptions on

[2] For each angle, the probability density values in the CAT-distribution were computed at evenly distanced points on a ray with this angle. Such density values relative to Cartesian coordinates $[x, y] = r \cdot [\sin\varphi, \cos\varphi]$ were converted to density values relative to polar coordinates $[r, \varphi]$ by multiplication with the Jacobi determinant r, described in Section 5. The expected value of r given φ was then computed from these values in a straightforward way.

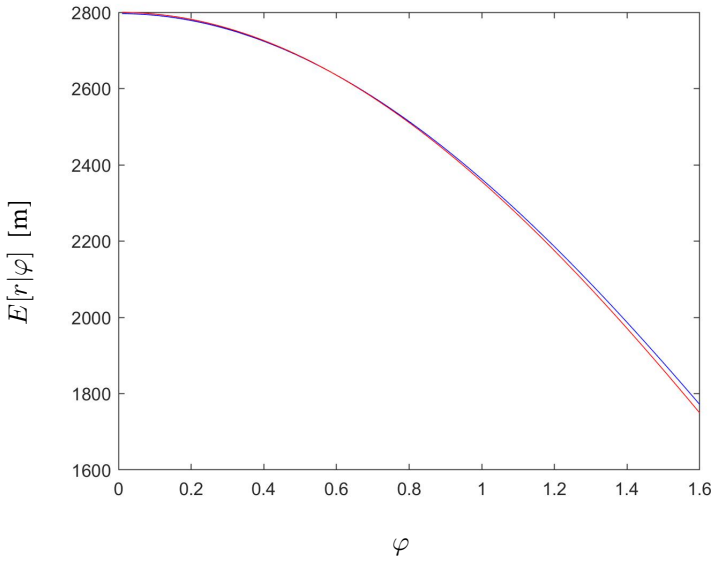

Figure 3: Numerically computed conditional expectation of r in blue compared to the function $v_0 t\,\mathrm{sinc}\,\varphi$ in red, using the parameter values from Figure 1.

the variables X and Y,

$$E\left[X|XY=z\right] \approx \mu_X\left(1-\frac{\sigma_X^2}{\mu_X^2}\right)+\frac{\sigma_X^2 z^2}{\sigma_Y^2\mu_X^3}. \tag{4}$$

In the present case, this implies (5) and hence also (6).

$$E\left[1+\frac{at}{6v_0}\,\middle|\,\varphi\right] \approx 1-\left(\frac{\sigma_a t}{6v_0}\right)^2+\frac{\left(\frac{\sigma_a t}{6v_0}\right)^2\varphi^2}{\left(\frac{1}{2}\sigma_\omega t\right)^2} \tag{5}$$

$$E\left[a\mid\varphi\right] \approx \frac{\sigma_a^2 t}{6v_0}\left(\frac{\varphi^2}{\left(\frac{1}{2}\sigma_\omega t\right)^2}-1\right) \tag{6}$$

Here it is worth noticing that the conditional mean of a is positive for $|\varphi| > \frac{\sigma_\omega t}{2}$ and negative for $|\varphi| < \frac{\sigma_\omega t}{2}$.

Applying (6) to (3) and introducing the notation

$$\mathrm{scc}\,\varphi = \frac{1}{3}\left(4\,\mathrm{sinc}\,\varphi-\cos\varphi\right) \tag{7}$$

yields the approximation $P_\mu(\varphi)$ for the mean of r when φ is given.

$$P_\mu(\varphi) = v_0 t\,\mathrm{sinc}\,\varphi+\frac{\sigma_a^2 t^3}{12v_0}\left(\frac{\varphi^2}{\left(\frac{1}{2}\sigma_\omega t\right)^2}-1\right)\mathrm{scc}\,\varphi \tag{8}$$

The curve for this expression compares much better with the computed curve for the mean, cf. Figure 4.

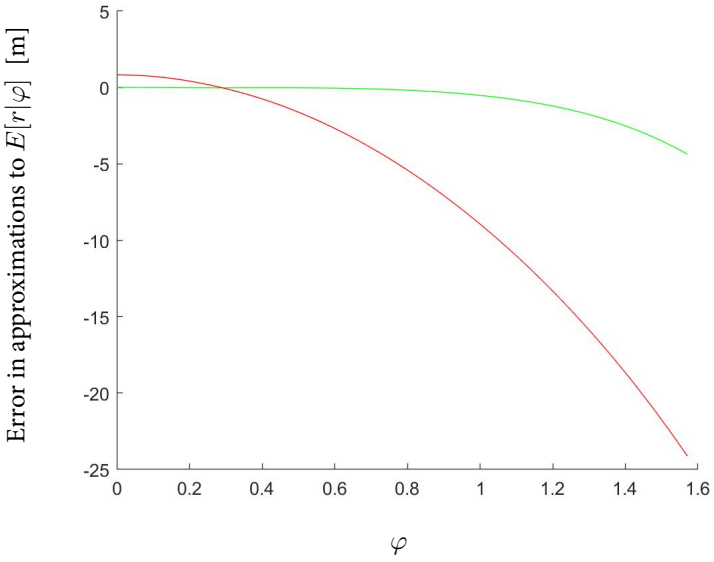

Figure 4: Errors in meters when approximating $E[r|\varphi]$ with the functions $v_0 t \operatorname{sinc} \varphi$ (red) and $P_\mu(\varphi)$ (green). The parameter values are those from Figure 1.

4.3 Standard deviation

In the first, rough approximation, $r|\varphi$ has a standard deviation that does not depend on φ. Numerical computations show that this is not accurate; Figure 5 graphs the standard deviations of these various probability distributions along the vertical axis against the given values of φ along the horizontal axis, with the values decreasing significantly when φ runs from 0 to $\frac{\pi}{2}$.

In [4] we show that under the above assumptions on the variables X and Y,

$$SD\left[X|z\right] \approx \sigma_X \left(1 + \frac{1}{2}\frac{\sigma_X^2}{\mu_X^2}\right) - \frac{3}{2}\frac{\sigma_X^3 z^2}{\sigma_Y^2 \mu_X^4}. \tag{9}$$

In the present case, it follows that the conditional standard deviation of $1 + \frac{at}{6v_0}$ when φ is given, is

$$SD\left[1 + \frac{at}{6v_0}\middle|\varphi\right] \approx \frac{\sigma_a t}{6v_0}\left(1 + \frac{1}{2}\left(\frac{\sigma_a t}{6v_0}\right)^2\right) - \frac{3}{2}\frac{\left(\frac{\sigma_a t}{6v_0}\right)^3 \varphi^2}{\left(\frac{1}{2}\sigma_\omega t\right)^2},$$

yielding a conditional standard deviation for a of

$$SD\left[a|\varphi\right] \approx \sigma_a \left(1 + \frac{\sigma_a^2}{6v_0^2}\left(\frac{t^2}{12} - \frac{\varphi^2}{\sigma_\omega^2}\right)\right), \tag{10}$$

and – using (3) – a conditional standard deviation for r of approximately

$$P_\sigma(\varphi) = \frac{\sigma_a t^2}{2}\left(1 + \frac{\sigma_a^2}{6v_0^2}\left(\frac{t^2}{12} - \frac{\varphi^2}{\sigma_\omega^2}\right)\right) \operatorname{scc} \varphi. \tag{11}$$

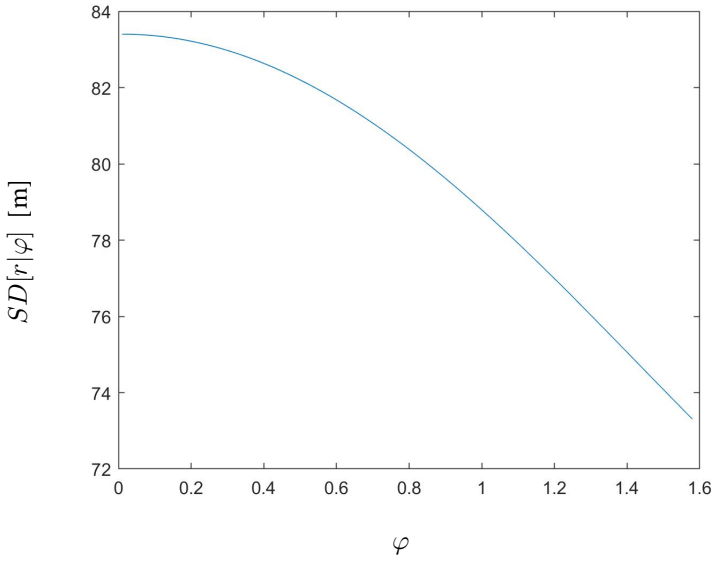

Figure 5: The function $SD[r \mid \varphi]$, numerically computed using the parameter values in Fig. 1.

We keep the term with σ_a^2 in (11), as well as in (8), although a similar term was neglected in the approximation for σ_φ; the reason lies in a difference in relative importance.

5 Use of improved approximation

Using the expressions (8) and (11) for the mean and standard deviation of r given φ, we obtain an improved approximation for the CAT distribution.

IMPROVED APPROXIMATION

1. $\varphi \sim \mathcal{N}\big(0, \frac{1}{2}\sigma_\omega t\big)$
2. $r|\varphi \sim \mathcal{N}(P_\mu(\varphi), P_\sigma(\varphi))$

Cartesian coordinates $[x, y]$ can be found from polar coordinates $[\varphi, r]$ as $[x, y] = r \cdot [\sin\varphi, \cos\varphi]$, with Jacobi determinant $\begin{vmatrix} r\cos\varphi & \sin\varphi \\ -r\sin\varphi & \cos\varphi \end{vmatrix} = r$. If the rough approximation had been accurate, the probability density function on Cartesian coordinates would be

$$p_0(x, y) = \frac{1}{r}\, g(\varphi; 0, \tfrac{1}{2}\sigma_\omega t) \cdot g(r; v_0 t\,\mathrm{sinc}\,\varphi, \tfrac{1}{2}\sigma_a t^2)$$

where $g(x; \mu, \sigma)$ is the normal probability density function $\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ and φ, r are obtained from x, y as above. This is how the red contours in Figure 1 were produced. Using the functions P_μ, P_σ instead, one obtains the probability density function $p_1(x, y)$

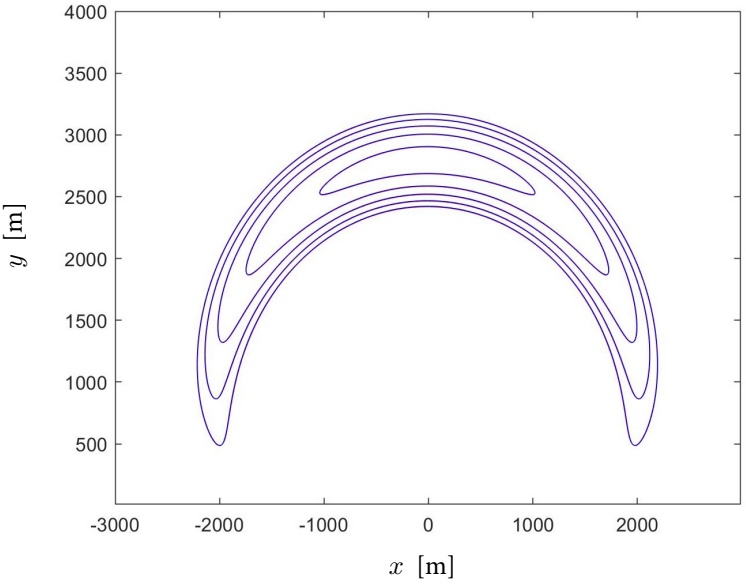

Figure 6: Contour plot shown in blue for the CAT distribution with the parameter values from Figure 1. The contour lines are drawn for density values between $10^{-6}\mathrm{m}^{-2}$ and $10^{-10}\mathrm{m}^{-2}$. Corresponding contours for the improved approximation were drawn in red, but are not visible from under the blue.

of the improved approximation,

$$p_1(x,y) = \frac{1}{r}\, g(\varphi; 0, \tfrac{1}{2}\sigma_\omega t) \cdot g(r; P_\mu(\varphi), P_\sigma(\varphi)).$$

Figure 6 shows contour lines in the density functions for two banana distributions; for the improved approximation in red, and for the genuine CAT-distribution in blue. This time, the red is completely covered by the blue. Table 2 lists the computed percentages of the total probability mass inside each of the contours, and the maximal error rate of the approximations relative to exact density values within each of the contours. Since the values on neighbouring contour lines differ by 900%, the maximal difference of a mere 6.6% in the table is not visible at this level.

Table 2: Probability mass and maximal error rate within each of the contours of Figure 6.

Contour at density level $[\mathrm{m}^{-2}]$	10^{-6}	10^{-7}	10^{-8}	10^{-9}	10^{-10}
Total probability mass inside	58%	96%	99.6%	99.96%	99.997%
Maximal error within contour	0.06%	0.22%	0.64%	2.6%	6.6%

6 Goodness of approximation with different parameter settings

6.1 Choice of measure

The rough and improved approximations presented above were both formulated and presented for use with different settings of the four parameters v_0, t, σ_ω and σ_a, while Table 2 only reflects the behaviour with respect to one particular parameter setting. To see how the improved approximation fares for the wider range of parameter settings, it is convenient to focus on a single metric for the goodness of the approximations. The Bhattacharyya coefficient is such a single measure that is often used to compare probability distributions, but for present purposes we have designed a measure that indicates in a more direct way in which cases the error of approximation is acceptable. The measure we use asks how far out in the distribution it is possible to go before a 5% discrepancy between the actual and the approximate probability density value is encountered. The question is answered by inspecting discrepancies between approximations and exact values inside ever wider contour lines in the density function until a 5% discrepancy is found, and then computing the probability mass inside this particular contour line. From Table 2, it can be seen that the figure must be somewhere between 0.9996 and 0.99997 in the case of the default settings from Figure 1. The true value is approximately 0.99992; this means that inside a geographical area that has a probability of 0.99992 of containing the object, and whose boundary is a contour line for the probability density $p(x, y)$, the estimated probability density value nowhere deviates from the true figure by more than 5% of its value.[3] Association algorithms will typically disregard any observations falling outside of such an area, and a 5% error in the estimated probability density value is well within the acceptable in most cases. The next task is to determine under which other possible settings of the model parameters does the approximation offer a similarly acceptable behavior.

6.2 Variation of parameters

To find how the approximations fare for different parameter values it is, fortunately, not necessary to vary all four parameters v_0, t, σ_a, σ_ω independently. To see this, note that the position vector in (1) can be written on the form $v_0 t A$, where A only contains v_0, t, a and ω inside the subexpressions ωt and $\frac{at}{v_0}$, and similarly that $P_\mu(\varphi)$ and $P_\sigma(\varphi)$ can both be written on the form $v_0 t B$, where B only contains v_0, t, σ_a and σ_ω inside the subexpressions $\sigma_\omega t$ and $\frac{\sigma_a t}{v_0}$. The values of $\sigma_\omega t$ and $\frac{\sigma_a t}{v_0}$ therefore determine the behaviour of both the CAT-distribution and the approximation, while the product $v_0 t$ only gives the scale and has no bearing on the measure outlined above. To evaluate the approximations for various parameter values, it is thus sufficient to vary $\sigma_\omega t$ and $\frac{\sigma_a t}{v_0}$.

[3] Less formally; the largest "banana" inside of which the probability density is guaranteed to be estimated with an error not exceeding 5%, has a likelihood of 0.99992 of containing the object. The whole (not necessarily banana-shaped) region in which the estimation error stays within this bound, will in general be larger and in this case contains 0.99999 of the probability mass. It is possible to define an alternative similarity measure based on the existence of such arbitrarily shaped regions (i.e., non-bananas) with some density values on the inside falling below corresponding values on the outside. Such an alternative measure will always yield figures as least as good as those achieved by our chosen measure, but we consider this original measure more intuitive.

Note also the connection between these measures and the Bhattacharyya coefficient, which in the present example will be at least $\iint_D \sqrt{p(x,y) \cdot 0.95\, p(x,y)} dA = \sqrt{0.95} \iint_D p(x,y) dA = \sqrt{0.95} \cdot 0.99999$, where D is a region containing 0.99999 of the probability mass.

6.3 Results

The goodness measure outlined above was computed for 40000 different parameter settings; the combinations of 200 values of $\sigma_\omega t$ evenly distributed between 0 and 75 degrees, and 200 values of $\frac{\sigma_a t}{v_0}$ evenly distributed between 0 and 0.2. The results are shown in Figure 7, with values of $\frac{\sigma_a t}{v_0}$ along the horizontal axis and values of $\sigma_\omega t$ along the vertical axis, while the labelled curves show the contour lines in our measure; how much of the probability mass to include before an approximation error of 5% is found. The results

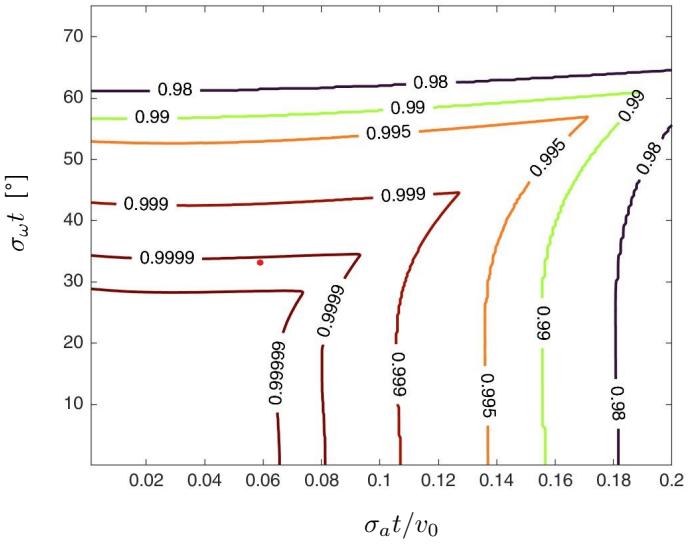

Figure 7: Contour plot showing, for varying parameter settings, how much of the probability mass has to be included to encounter an error of at least 5%. The red dot indicates the location of the model used in Figure 1.

show that the approximations are better for smaller values of the parameters, and that as long as $\frac{\sigma_a t}{v_0}$ is kept below 0.1 (or $\sigma_a < \frac{v_0}{10t}$) and $\sigma_\omega t$ is kept below approximately $42°$, the error of approximation stays below 5% for the innermost 99.9% of the probability distribution.

These results were obtained using the relatively simple expressions $P_\mu(\varphi)$ and $P_\sigma(\varphi)$ for the mean and standard deviation of $r|\varphi$. Using more complex variants of equations (2), (3), (4), and (9), more complex functions $P_\mu(\varphi)$ and $P_\sigma(\varphi)$ can be obtained that allow the contours to move somewhat further to the right and significantly higher up.

7 Stochastic initial direction

At this point we introduce a third element of randomness, namely, in the initial direction. This direction is given by an angle θ_0, while the final angle and distance from the origin are now denoted by θ and r. Then

$$\theta = \theta_0 + \varphi,$$

where, as before, φ is the deviation from the initial direction (but no longer denotes a polar coordinate). We assume that φ and r are independent of θ_0 and are distributed as in the previous sections. Furthermore, we assume that θ_0 is normally distributed with mean 0 and standard deviation σ_0.

The extended model presents some additional challenges for the numerical computation of the probability density function $p(r, \theta)$, but building upon the approximations above, we will show that our framework is well suited for handling the added complication. Indeed, the approximation

$$\varphi \sim \mathcal{N}(0, \sigma_\varphi)\,, \quad \sigma_\varphi = \tfrac{1}{2}\sigma_\omega t,$$

together with the assumption for θ_0, immediately gives

$$\theta \sim \mathcal{N}(0, \sigma_\theta)\,, \quad \sigma_\theta^2 = \sigma_0^2 + \sigma_\varphi^2.$$

The distribution of $r|\theta$ remains to be found. It follows from Theorem 1.4.2 in [5] that if $Z = X + Y$ is the sum of two independent, normally distributed random variables with mean 0, then $X|Z = z$ is normally distributed with mean $z\sigma_X^2/\sigma_Z^2$ and standard deviation $\sigma_X\sigma_Y/\sigma_Z$. Thus, to a good approximation,

$$\varphi|\theta \sim \mathcal{N}\big(\bar{\theta}, \bar{\sigma}\big)\,, \tag{12}$$

where

$$\bar{\theta} = \frac{\theta\sigma_\varphi^2}{\sigma_\theta^2} = \frac{\theta}{1 + 4\sigma_0^2/\sigma_\omega^2 t^2}, \qquad \bar{\sigma} = \frac{\sigma_0\sigma_\varphi}{\sigma_\theta} = \left(\frac{1}{\sigma_0^2} + \frac{4}{\sigma_\omega^2 t^2}\right)^{-\frac{1}{2}}.$$

Note that $\bar{\sigma} \leq \min(\sigma_0, \sigma_\varphi)$.

By the law of total probability,

$$p(r|\theta) = \int d\varphi\, p(r|\theta, \varphi)\, p(\varphi|\theta) = \int d\varphi\, p(r|\varphi)\, p(\varphi|\theta), \tag{13}$$

since r is independent of θ when φ is given. As discussed in Section 4, $r|\varphi$ is very close to being normally distributed. With this in mind, (12) and (13) show that the distribution of $r|\theta$ is approximately normal if $\bar{\sigma}$ is small.[4]

By the law of total expectation,

$$E[r|\theta] = \int d\varphi\, E[r|\theta, \varphi]\, p(\varphi|\theta) = \int d\varphi\, E[r|\varphi]\, p(\varphi|\theta).$$

With the approximation from Section 4 and (12) we have

$$E[r|\theta] \approx \int d\varphi\, P_\mu(\varphi)\, g\left(\varphi; \bar{\theta}, \bar{\sigma}\right). \tag{14}$$

By the same argument,

$$E[r^2|\theta] \approx \int d\varphi\, \left(P_\mu(\varphi)^2 + P_\sigma(\varphi)^2\right) g\left(\varphi; \bar{\theta}, \bar{\sigma}\right).$$

[4] To see this in detail, note that the integral in (13) can be evaluated approximately by assuming that the mean of the near-Gaussian function $p(r|\varphi)$ varies linearly with φ around $\bar{\theta}$ while the standard deviation is constant. With $E[r|\varphi] = k\varphi + b$ and $SD[r|\varphi] = c$, the integrand can be rewritten as $g(r - k\varphi; b, c)|k|g(k\varphi; k\bar{\theta}, |k|\bar{\sigma})$. The integral is thus converted into a convolution of two Gaussians, yielding another Gaussian in r.

Replacing $P_\mu(\varphi)$ in (14) by its second-order Taylor polynomial around $\varphi = \bar{\theta}$ leads to

$$E[r|\theta] \approx P_\mu(\bar{\theta}) + \tfrac{1}{2} P''_\mu(\bar{\theta})\bar{\sigma}^2. \tag{15}$$

Similarly (suppressing the argument $\bar{\theta}$),

$$E[r^2|\theta] \approx P_\mu^2 + P_\sigma^2 + \left((P'_\mu)^2 + P_\mu P''_\mu + (P'_\sigma)^2 + P_\sigma P''_\sigma\right)\bar{\sigma}^2.$$

From the previous two equations (and dropping terms of fourth order in $\bar{\sigma}$ and $\frac{\sigma_a t}{v_0}$), an approximation for the conditional variance is given by

$$SD[r|\theta]^2 \approx P_\sigma(\bar{\theta})^2 + P'_\mu(\bar{\theta})^2\bar{\sigma}^2. \tag{16}$$

Figure 8 shows the effect of a normally distributed initial direction, using the parameter values in Figure 1. The blue curve to the left shows the numerically computed conditional probability density $p(r|\theta = \frac{\pi}{4})$ in the case $\sigma_0 = 0$. Underneath, barely visible in red, is a normal distribution with mean and standard deviation given by (15) and (16) (which in this case simplify to the approximations in Section 4). Corresponding curves are shown for $\sigma_0 = \frac{\pi}{30}$ (middle) and $\sigma_0 = \frac{\pi}{15}$ (right). It should be noted that the above formulas for the conditional mean and standard deviation give results whose deviations from the numerically computed values are well below 1 meter in all three cases; the slight visible difference between a blue and a red curve is purely due to deviation from normality in the blue one.

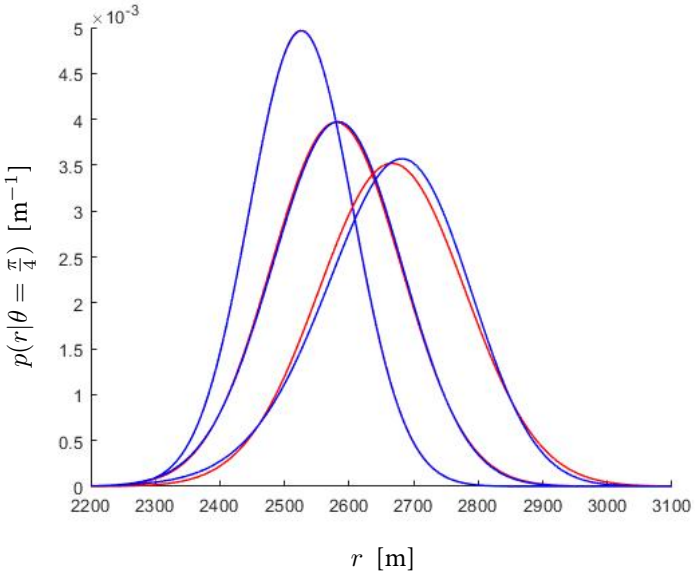

Figure 8: Numerically computed conditional probability density of r for $\theta = \frac{\pi}{4}$ (blue) with the parameter values from Figure 1 and a standard deviation in the initial direction of (from left to right) $\sigma_0 = 0$, $\frac{\pi}{30}$, and $\frac{\pi}{15}$. The three corresponding approximations (red, printed below the blue) are normal distributions based on (15) and (16).

Equation (16), whose two terms on the right-hand side are of second order in $\frac{\sigma_a t}{v_0}$ and $\bar{\sigma}$, respectively, can be improved by including terms of the next order. With the additional term $\frac{1}{8} P_\mu^{(4)}(\bar{\theta})\bar{\sigma}^4$ in (15) and corresponding terms in the expansion for $E[r^2|\theta]$,

the omitted fourth-order terms in (16) are found to be

$$\left((P'_\sigma)^2 + P_\sigma P''_\sigma\right)\bar{\sigma}^2 + \left(\tfrac{1}{2}(P''_\mu)^2 + P'_\mu P'''_\mu\right)\bar{\sigma}^4,$$

where the functions are evaluated at $\bar{\theta}$. These correction terms are insignificant in the case in the middle of Figure 8, while in the rightmost case, where $\bar{\sigma}$ is twice as large, they help to improve the estimate of $SD[r|\theta]$, reducing the deviation from the numerically computed value from about 70 cm to only 5 cm.

Conclusion

We have seen that the use of a simple type of stochastic polar coordinates based on normal distributions to compute probability densities of banana distributions is both computationally tractable, closely related to the geometry of the distribution, and, in a wide range of situations, remarkably accurate.

Acknowledgments

We would like to thank our colleague Knut Meen at the Norwegian Naval Academy, and Brita Gade, Carina Vooren, Morten Kloster, and Ole Halvard Sætran at the Norwegian Defence Research Establishment for inspiring discussions and valuable suggestions. We also thank the anonymous referees for useful comments.

References

[1] A. W. Long, K. C. Wolfe, M. J. Mashner, and G. S. Chirikjian, "The Banana Distribution is Gaussian: A Localization Study with Exponential Coordinates," in N. Roy, P. Newman, and S. Srinivasa (Eds.), *Robotics: Science and Systems VIII*, MIT Press, 2013, pp. 265-272.

[2] B. H. H. Gade, C. N. Vooren, and M. Kloster, "Probability Distribution for Association of Maneuvering Vehicles," 2019 22th International Conference on Information Fusion (FUSION), Ottawa, ON, Canada, 2019, pp. 1-7.

[3] L. A. Aroian, "The Probability Function of the Product of Two Normally Distributed Variables," The Annals of Mathematical Statistics, Vol. 18, No. 2 (Jun., 1947), pp. 265-271.

[4] H. Totland and T. Langholm, "Deformed Normal Distributions," forthcoming.

[5] P. J. Bickel and K. A. Doksum, *Mathematical Statistics*, Prentice Hall, 1977.

Appendix

Proposition 1. *For* $|\omega t| < \pi$ *and* $|\frac{at}{v_0}| \ll 1$,

$$\varphi \approx \frac{\omega t}{2}\left(1 + \frac{at}{6v_0}\right).$$

Proof. With

$$x_1 = 1 - \cos(\omega t), \qquad x_2 = \frac{1}{\omega}\left(\sin(\omega t) - \omega t\cos(\omega t)\right),$$
$$y_1 = \sin(\omega t), \qquad y_2 = \frac{1}{\omega}\left(\cos(\omega t) + \omega t\sin(\omega t) - 1\right),$$

we can write

$$\varphi = \operatorname{atan2}\left(x_1 + \frac{a}{v_0}x_2,\ y_1 + \frac{a}{v_0}y_2\right).$$

The first-order Taylor polynomial in $\frac{a}{v_0}$ of this expression is

$$\operatorname{atan2}(x_1,\ y_1) + \frac{a}{v_0}\cdot\frac{x_2y_1 - x_1y_2}{x_1^2 + y_1^2},$$

where the second term equals

$$\begin{aligned}&\frac{a}{v_0}\cdot\frac{1}{\omega}\cdot\frac{2 - 2\cos(\omega t) - \omega t\sin(\omega t)}{2 - 2\cos(\omega t)}\\ &= \frac{a}{v_0}\cdot\frac{1}{\omega}\cdot\left(1 - \frac{\omega t}{2}\cot\left(\frac{\omega t}{2}\right)\right)\\ &= \frac{at}{2v_0}\cdot\left(\frac{1}{\frac{\omega t}{2}} - \cot\left(\frac{\omega t}{2}\right)\right).\end{aligned}$$

Now $\operatorname{atan2}(x_1,\ y_1) = \frac{\omega t}{2}$, while $\frac{1}{x} - \cot x \approx \frac{x}{3}$ for $|x| < \frac{\pi}{2}$, so

$$\varphi \approx \frac{\omega t}{2} + \frac{at}{2v_0}\cdot\frac{\frac{\omega t}{2}}{3} = \frac{\omega t}{2}\left(1 + \frac{at}{6v_0}\right).$$

□

Proposition 2. *For* $|\varphi| < \frac{\pi}{2}$ *and* $|\frac{at}{v_0}| \ll 1$,

$$r \approx v_0 t\operatorname{sinc}\varphi + \tfrac{at^2}{6}\left(4\operatorname{sinc}\varphi - \cos\varphi\right).$$

Proof. First note that

$$r \approx v_0 t\operatorname{sinc}\left(\frac{\omega t}{2}\right)\cdot\left(1 + \frac{at}{2v_0}\right). \tag{17}$$

This can be shown by obtaining

$$r = v_0 t\,\frac{\sin(\frac{\omega t}{2})}{\frac{\omega t}{2}}\sqrt{\left(1 + \frac{at}{2v_0}\right)^2 + \left(\frac{at}{2v_0}\right)^2\cdot\left(\frac{1}{\frac{\omega t}{2}} - \cot\frac{\omega t}{2}\right)^2}$$

from (1) by standard trigonometric identities, and then expanding in powers of $\frac{at}{2v_0}$. Next we observe that

$$\operatorname{sinc}\left(\frac{\omega t}{2}\right) \approx \operatorname{sinc}\varphi - \frac{at}{6v_0}(\cos\varphi - \operatorname{sinc}\varphi). \tag{18}$$

This follows by noting that Proposition 1 also implies that $\frac{\omega t}{2} \approx \varphi\left(1 - \frac{at}{6v_0}\right)$ when $|\frac{at}{6v_0}| \ll 1$, and that the first-order Taylor polynomial of $\operatorname{sinc}(\varphi + \delta)$ around φ is $\operatorname{sinc}\varphi + \delta\cdot\frac{\cos\varphi - \operatorname{sinc}\varphi}{\varphi}$.

Finally, substituting (18) into (17), multiplying out and deleting second-order terms in $\frac{at}{v_0}$, one obtains the proposition. □