This is the accepted manuscript made available via CHORUS. The article has been published as:

**Classical Nucleation Theory Description of Active Colloid Assembly**

Gabriel S. Redner, Caleb G. Wagner, Aparna Baskaran, and Michael F. Hagan


DOI: [10.1103/PhysRevLett.117.148002](http://10.1103/PhysRevLett.117.148002)
A classical nucleation theory description of active colloid assembly

Gabriel S. Redner,† Caleb G. Wagner,* Aparna Baskaran, and Michael F. Hagan†
Martin Fisher School of Physics, Brandeis University, Waltham, MA, USA.

Non-aligning self-propelled particles with purely repulsive excluded volume interactions undergo athermal motility-induced phase separation into a dilute gas and a dense cluster phase. Here, we use enhanced sampling computational methods and analytic theory to examine the kinetics of the dense phase. Despite the intrinsically nonequilibrium nature of the phase transition, we show that the kinetics can be described using an approach analogous to equilibrium classical nucleation theory, governed by an effective free energy of cluster formation with identifiable bulk and surface terms. The theory captures the location of the binodal, nucleation rates as a function of supersaturation, and the cluster size distributions below the binodal, while discrepancies in the metastable region reveal additional physics about the early stages of active crystal formation. The success of the theory shows that a framework similar to equilibrium thermodynamics can be obtained directly from the microdynamics of an active system, and can be used to describe the kinetics of evolution toward nonequilibrium steady states.

Active fluids consisting of self-propelled units are present in many biological systems, including the cell cytoplasm [1–3], bacterial suspensions [4–7], and animal flocks [8–11]. Recently, researchers have also developed synthetic active fluids, consisting of chemically [12–20] or electrically [21] propelled colloids, or monolayers of vibrated granular particles [22–25]. Being intrinsically nonequilibrium, active fluids cannot be described by equilibrium statistical mechanics [26, 27] and exhibit behaviors not possible in equilibrium systems, such as spontaneous flow [28–41] and athermal phase separation [42–52]. Yet, active systems frequently evolve to well-defined time-independent distributions, and in some cases are characterized by equilibrium-like state variables such as temperature or pressure [50, 53–60]. While significant progress has been made toward understanding these stationary distributions, the kinetics of evolution toward steady-state remain poorly understood.

As in equilibrium physics, progress in active matter often stems from simplified model systems. One such system is composed of active Brownian particles (ABPs): self-propelled particles which interact solely by short-range isotropic repulsion. Despite lacking interparticle attractions or alignment interactions, ABPs form macroscopic, crystalline clusters [18–20, 42–47, 49, 61, 62]. (This is an example of a generic instability toward density inhomogeneity, motility-induced phase separation (MIPS), which can arise when particle velocities decrease with increasing local particle density [44, 49, 63]). ABP phase separation is strikingly reminiscent of equilibrium vapor-liquid systems, with the densities of the coexisting phases falling along a binodal, and critical-like behavior near its apex. As a minimal model system possessing nontrivial phase behavior, ABPs are ideal for studying evolution toward steady-state in generic active systems. However, while the coarsening of deeply quenched ABP clusters has been studied numerically [43, 44, 47], there is no theory for the complete kinetics of phase separation. Moreover, while existing phenomenological descriptions of ABPs have led to important insights about MIPS [44, 64–66], there is currently no approach to directly calculate phase behavior from the microdynamics of a particle-based model.

To overcome these limitations, we describe ABP clustering dynamics and steady-state phase behavior with a theory analogous to classical nucleation theory (CNT) for equilibrium phase separation [67–69]. Beginning with a geometric picture of ABP interactions, we construct an effective free energy of cluster formation which resembles that of droplet nucleation in an equilibrium liquid-vapor system. We then apply the framework of CNT to calculate nucleation rates and determine phase behavior as functions of particle density and propulsion velocity.

While previous descriptions of ABP phase separation are based on a functional ansatz at the level of pair correlations or the dependence of particle velocities on local density [44, 64–66], here we show how such frameworks emerge from the kinetics of the microscopic model. In particular, our theory leads to a simple relationship between the microscopic parameters of an ABP model and the driving force for ABP phase separation (analogous to the chemical potential difference between dense and dilute phases in equilibrium phase separation).

We test the theory against simulations of ABPs, employing enhanced sampling techniques to make systematic measurements of nucleation rates. Despite approximations in our microscopic model, the predicted phase boundary matches simulation results almost quantitatively, with no adjustable parameters. Moreover, the predicted and measured cluster size distributions match well below the binodal, though we discuss interesting effects of nontrivial cluster geometry that lead to power law scaling in the metastable regime. The theory qualitatively predicts the dependence of nucleation rates on super-saturation, although significant quantitative differences are seen near the binodal.

In the last century, CNT drove tremendous advances in
the fields of equilibrium crystallization and self-assembly by relating particle-scale interactions to their macroscale assembly dynamics. The framework described here moves toward a similar capability in active materials, showing how the breaking of time-reversal symmetry at the level of individual particles controls emergent nonequilibrium assembly.

Model. ABPs in two dimensions obey the overdamped Langevin equations:

\[
\begin{align*}
\dot{r}_i &= F((r_i))/\xi + v_p \mathbf{v}_i + \sqrt{2D\eta_i^T} \\
\dot{\theta}_i &= \sqrt{2D_i}\eta_i^R.
\end{align*}
\]

Here \( F \) represents the interparticle repulsion force, \( \xi \) is the drag, \( v_p \) is the magnitude of the self-propulsion velocity, and \( \mathbf{v}_i = (\cos \theta_i, \sin \theta_i) \). The \( \eta \) variables introduce Gaussian noise, with \( \langle \eta_i(t) \rangle = 0 \) and \( \langle \eta_i(t)\eta_j(t') \rangle = \delta_{ij} \delta(t - t') \). Although the noise may be non-thermal, we set \( D_i = 3\eta_i^2/\pi \) (with \( \sigma \) the particle diameter) as would apply to a sphere in the low-Reynolds-number regime.

Due to the self-replenishing velocity \( v_p \), collisions between particles are rendered persistent, which leads to cluster formation from the dilute phase. To model such cluster formation in equilibrium systems, CNT assumes a free energy of the form \( G(n) = \Delta \mu V(n) + \gamma A(n) \), with \( n \) the number of molecules in the droplet, \( V(n) \) and \( A(n) \) the droplet volume and surface area, \( \Delta \mu \) the chemical potential difference between the dense and dilute phases, and \( \gamma \) the surface tension. This assumption, together with the Becker-Döring description of cluster growth kinetics [70], allows predicting nucleation rates as a function of material constants and concentrations.

Since ABP phase separation is intrinsically nonequilibrium, the same prescription cannot be directly applied. Therefore, we take the opposite approach, starting from the Becker-Döring kinetics and inferring an effective free energy landscape. To this end, we assume that the state of an ABP system may be represented at the mesoscopic energy landscape. To this end, we assume that the state

\[
\rho = \rho_1 \prod_{m=1}^{n-1} \frac{j_{in}(m)S(m)}{j_{out}(m+1)S(m+1)} \equiv \rho_1 P(n),
\]

Within the phase separation regime, our simulations identify a parameter range within which nucleation is slow in comparison to the settling time of the CSD (analogous to the metastable region between the binodal and the spinodal in an equilibrium system). In these cases the CSD may still be taken to be (quasi) stationary, \( \partial_t \rho_n = 0 \), but to access the slow nucleation dynamics we must acknowledge a small nonzero flux \( J(n) \). Under these conditions the fluxes \( J \) are equal and given by [73]:

\[
J = \rho_1 \left( \sum_{n=1}^{\infty} \frac{1}{j_{in}(n)S(n)P(n)} \right)^{-1}.
\]

The mean nucleation time in a system with volume \( V \) is then 

\[
\tau_{nucl} = (JV)^{-1} [67].
\]

To proceed further, we construct a minimal microscopic model that enables estimating the adsorption and evaporation fluxes. We model each cluster as circular, with volume \( V(n) = \pi r^2 n \) and surface area \( S(n) = 2\pi r n \rho_c \), with \( \rho_c \) near the close-packing density for spheres. Using basic arguments for how particles adsorb on and depart from clusters [43, 71, 73], we obtain

\[
\begin{align*}
\frac{\partial \rho_n}{\partial t} &= J(n-1) - J(n) \\
J(n) &= j_{in}(n)S(n)\rho_n - j_{out}(n+1)S(n+1)\rho_{n+1}
\end{align*}
\]

where the fluxes \( j_{in} \) and \( j_{out} \) represent the rates of monomer gain and loss per unit of cluster surface and \( S(n) \) is the surface area (perimeter) of a cluster of size \( n \).

To gain insight into the phase behavior and kinetics of the system, we first consider a steady-state in which the fluxes \( J \) are zero. While the existence of such steady-states in the physical system is not guaranteed, they are consistently observed in simulations [42–47, 62, 71]. In the absence of phase separation, the steady-state corresponds to free monomers coexisting with small, transient clusters. The steady-state cluster size distribution (CSD) \( P(n) \) can be calculated by iterating Eq. 4:

\[
\rho_n = \rho_1 \prod_{m=1}^{n-1} \frac{j_{in}(m)S(m)}{j_{out}(m+1)S(m+1)} \equiv \rho_1 P(n),
\]

The mean nucleation time in a system with volume \( V \) is then

\[
\tau_{nucl} = (JV)^{-1} [67].
\]

To proceed further, we construct a minimal microscopic model that enables estimating the adsorption and evaporation fluxes. We model each cluster as circular, with volume \( V(n) = \pi r^2 n \) and surface area \( S(n) = 2\pi r n \rho_c \), with \( \rho_c \) near the close-packing density for spheres. Using basic arguments for how particles adsorb on and depart from clusters [43, 71, 73], we obtain

\[
\begin{align*}
\frac{\partial \rho_n}{\partial t} &= J(n-1) - J(n) \\
J(n) &= j_{in}(n)S(n)\rho_n - j_{out}(n+1)S(n+1)\rho_{n+1}
\end{align*}
\]

where the fluxes \( j_{in} \) and \( j_{out} \) represent the rates of monomer gain and loss per unit of cluster surface and \( S(n) \) is the surface area (perimeter) of a cluster of size \( n \).

To gain insight into the phase behavior and kinetics of the system, we first consider a steady-state in which the fluxes \( J \) are zero. While the existence of such steady-states in the physical system is not guaranteed, they are consistently observed in simulations [42–47, 62, 71]. In the absence of phase separation, the steady-state corresponds to free monomers coexisting with small, transient clusters. The steady-state cluster size distribution (CSD) \( P(n) \) can be calculated by iterating Eq. 4:

\[
\rho_n = \rho_1 \prod_{m=1}^{n-1} \frac{j_{in}(m)S(m)}{j_{out}(m+1)S(m+1)} \equiv \rho_1 P(n),
\]

Within the phase separation regime, our simulations identify a parameter range within which nucleation is slow in comparison to the settling time of the CSD (analogous to the metastable region between the binodal and the spinodal in an equilibrium system). In these cases the CSD may still be taken to be (quasi) stationary, \( \partial_t \rho_n = 0 \), but to access the slow nucleation dynamics we must acknowledge a small nonzero flux \( J(n) \). Under these conditions the fluxes \( J \) are equal and given by [73]:

\[
J = \rho_1 \left( \sum_{n=1}^{\infty} \frac{1}{j_{in}(n)S(n)P(n)} \right)^{-1}.
\]

The mean nucleation time in a system with volume \( V \) is then

\[
\tau_{nucl} = (JV)^{-1} [67].
\]

To proceed further, we construct a minimal microscopic model that enables estimating the adsorption and evaporation fluxes. We model each cluster as circular, with volume \( V(n) = \pi r^2 n \) and surface area \( S(n) = 2\pi r n \rho_c \), with \( \rho_c \) near the close-packing density for spheres. Using basic arguments for how particles adsorb on and depart from clusters [43, 71, 73], we obtain

\[
\begin{align*}
\frac{\partial \rho_n}{\partial t} &= J(n-1) - J(n) \\
J(n) &= j_{in}(n)S(n)\rho_n - j_{out}(n+1)S(n+1)\rho_{n+1}
\end{align*}
\]

where \( \rho_k \) is the monomer density in the “free volume” not occupied by clusters (different from \( \rho_1 \)), and

\[
\alpha(n) = \frac{1}{2} \left( \pi - \sin^{-1} \left( \frac{\sqrt{2\pi} n}{2\rho_c} \right) \right)
\]

is the “horizon angle” taking into account cluster curvature (Fig. 1) [74].

Plugging in to Eq. 5, we have:

\[
P(n) = \frac{(z\rho_k)^{n-1}}{\sqrt{n}} P_0(n)
\]

where \( z = \frac{v_p \sigma}{2\nu} \) is analogous to the Péclet number (Pe), and

\[
P_0(n) = \left[ \prod_{m=1}^{n-1} (2\alpha(m+1)/\pi) \right]^2
\]

accounts for the geometric effects of cluster size.

At this stage the kinetic theory is complete, and we can use the above formulæ directly to compute quantities of interest. On the other hand, we may continue the analogy with CNT by considering an effective free energy \( G(n) \) (analogous to the grand potential in equilibrium statistical mechanics). Since in equilibrium fluctuation theory we would have \( \rho_n \propto \exp \left[ -G(n)/(k_B T) \right] \), we write

\[
G(n) = -k_B T \ln \left( \sigma^2 \rho_n \right).
\]
Defined as such, \( G(n) \) serves as a natural formulation of the kinetic theory in terms of a functional landscape. In particular, we note that the existence of such an effective free energy landscape depends only on the presence of a (quasi)-steady state (Eq. 5), and is not contingent on a thermodynamic interpretation.

**Results.** Working within the effective free energy picture, we use Eqs. 7 and 8 to obtain

\[
G(n) = -k_B T \left[ n \ln(z \rho_g) - \frac{1}{2} \ln(n) + \ln(P_0(n)) \right] \tag{9}
\]

where terms not depending on \( n \) have been dropped. It is evident that the quantity \( z \rho_g \) controls the phase behavior of the system, and is analogous to the supersaturation of the fluid phase. As shown in Fig. 1, \( G(n) \) is monotonically increasing when \( z \rho_g < 1 \), corresponding to a homogeneous fluid, while for \( 1 < z \rho_g < \sqrt{2} \left( \frac{2 \alpha_2(n)}{\pi} \right)^2 \approx 2.42 \) it exhibits a barrier followed by a monotonic decrease, corresponding to a supersaturated fluid which is metastable to cluster nucleation. At higher values (beyond the ‘spinodal’), \( G(n) \) is monotonically decreasing and the system is unstable towards cluster formation.

For large clusters, \( P_0 \) can be simplified to (see section D of [73]):

\[
G(n) = -k_B T \left[ \ln(z \rho_g)n - \frac{\sigma_{\rho_c}}{\pi} S(n) \right] + O(\ln n) \tag{10}
\]

From this, we see that a geometric understanding of ABP microwdynamics leads naturally to their equilibrium-like phase behavior. Based on its role in governing the phase behavior of the system, the first term in Eq. 10 controls the relative propensity of a particle to be in the dilute or dense phase. Thus, in analogy with the equilibrium CNT free energy, this term represents the difference in effective chemical potential \( \Delta \mu \) between the two phases. Our expression \( \Delta \mu = \ln(z \rho_g/n) \) has similar structure to the chemical potential \( \Delta \mu = \ln(\rho) + \ln[v_p(\rho)] \) (\( \rho \) being a coarse-grained density field) considered in the continuum theory of Stenhammar, et. al [44]. The crucial difference here is the explicit appearance of the microscopic diffusion constant \( D_r \) in place of the functional ansatz \( v_p(\rho) \). The second term in Eq. 10 is related to the cluster’s surface area, and can be interpreted as an effective line tension that drives coarsening. Note that in a nonequilibrium system, this need not equal the mechanical line tension, and in fact Bialke et al. [66] measured a negative mechanical line tension for a flat interface in an ABP system. Because the excess free energy associated with interface formation must be positive for stability, Bialke et al. suggest the negative line tension is balanced by a positive interfacial stiffness. Since the line tension emerges from our calculation as a consequence of cluster curvature (expressed through the horizon angle \( \alpha(n) \)) there may be a connection to the Bialke et al. measurement, but comparisons at additional values of \( \rho \) are needed to explore this possibility. Finally, solving \( \Delta G(n) = 0 \) gives a prediction for the critical nucleus size as \( n_{\text{crit}} = \frac{\sigma^2 \rho_c}{\pi \ln(z \rho_g)} \), a form familiar from CNT.

Next, we demonstrate the quantitative insights of the theory into ABP systems. Our simulations are performed as in [43] (see [73] for details), though to measure nucleation times much larger than those accessible in brute force simulations, we used a weighted-ensemble dynamics [75–77]. To limit finite-sized effects in the NVT simulations, we consider systems with 15000 particles, thus providing a good estimate of nucleation rates ex-
cept very near the binodal where the critical nucleus size approaches the system size. [78] Also, since the control parameter in our NVT simulations is the overall volume fraction \( \phi \) whereas in the theory it is the theoretical gas density \( \rho_g \), we must construct a coordinate transformation which relates the two [73]:

\[
\phi = \frac{A\rho_g}{\left(4/\pi \sigma^2\right) A - 1 + \rho_c/\rho_g}
\]

where \( A = \frac{2}{3} \sum_{n=1}^{n_{\text{max}}} nP(n) \) with \( n_{\text{max}} \) a cutoff cluster size [73]. Finally, to make our comparison with simulations quantitative, we empirically fit the upper binodal by measuring the density within large clusters, \( \rho_c(n) \), which is found to increase with \( v_p \) due to the imperfectly hard interaction potential [73]. The resulting phase diagram is shown in Fig. 2. The predicted lower binodal is remarkably close to its measured location, although this could be partly fortuitous.

We now employ Eqs. 11 and 6 to compute the mean nucleation time in the metastable regime. Here we find that the theory, thus far constructed without adjustable parameters, lacks quantitative accuracy because the predicted nucleation rate is exquisitely sensitive to small perturbations of \( z\rho_g \). To enable qualitative comparisons, we set \( \rho_g = \chi \rho_g \), with \( \chi \) a fitting parameter that adjusts the monomer chemical potential. We find (by eye) that setting \( \chi = 0.71 \) produces good correspondence with simulation (Fig. 3). Nucleation rates are notoriously difficult to quantitatively predict from first principles even in the equilibrium case (e.g. [79–83]), so the correspondence between theory and simulation with a single small fitting parameter is notable.

Finally, we compare the theoretical and simulated CSDs in Fig. 4. Below and slightly above the binodal, the theory has the correct functional form, but far into the metastable region fails to account for power law scaling \( \rho_n \propto n^{-2} \) below a threshold cluster size, indicative of logarithmic corrections to \( G(n) \) [84]. Similarly, Levis et al. [85], observed power law scaling with exponent \( 1.70 \pm 0.05 \) in a related model system.

In equilibrium CNT, logarithmic corrections in the free energy have been obtained by accounting for degrees of freedom internal to a cluster, representing deviations from a spherical shape (e.g. [86–88]). Our results suggest analogous mechanisms play a role here. Simulations show that clusters have ramified structure, with a fractal surface area scaling \( S(n) \sim n^{0.84 \pm 0.01} \) [73]. Indeed, directly measuring fluxes (Fig. 4) shows deviations from what is expected for large spherical clusters: Simulated rates are larger than predicted by the theory, with the outgoing rates depending on \( v_p \). These attributes are consistent with complex cluster geometry, since on small clusters or near regions of high curvature, particles can escape by “sliding off” each other before completely rotating to the horizon, thus resulting in a higher than predicted, \( v_p \)-dependent outgoing rate. Multi-particle escape events may also contribute [43]. To test whether these mechanisms are responsible for the power law scaling, we fed the measured rates into Eq. 5 to reconstruct the CSDs, which showed good agreement with simulation (Fig. 4). Similar results were obtained elsewhere in the phase diagram (see [73]). Thus, a calculation of the rates which takes these effects into account should recover the scaling...
of simulated CSDs.

In summary, we have shown that an approach analogous to classical nucleation theory can describe the nonequilibrium nucleation of clusters from solutions of ABPs. By linking the microscopic parameters of ABPs to their macroscale kinetics, this framework makes an important step toward developing design principles for applications of motility-induced phase separation [89].

Acknowledgments. This work was supported by the NSF (DMR-1149266 and the Brandeis MRSEC DMR-1420382). Computational resources were provided by the NSF through XSEDE computing resources and the Brandeis HPCC.