

# A Behavioral Multispread Epidemic Model

Carlo Graziani\*

Argonne National Laboratory, Lemont, Illinois, USA

---

## Abstract

We introduce a class of epidemic models that represent multiple spread rates in terms of discrete behavior classes, rather than in terms of discrete compartments comprising individuals. The model is framed in terms of  $D$  behavior classes, each with its own spread rate. The population is represented as a density on the  $D$ -simplex, where each point is a  $D$ -vector  $f$  whose components sum to 1. Each component of  $f$  represents the fraction of time in which an individual spends engaging in the corresponding behavior. The evolution equation is an integro-differential equation on the  $D$ -simplex. The model is capable of describing the “superspread” phenomenon in terms of behavior spread rates, as opposed to terms of individual infectivity. We show the existence of SIR-like separable solutions and discuss their stability. We explore the numeric properties of the model using a  $D = 3$  case featuring a “safe” behavior, a moderate-spread behavior, and a superspread behavior.

**Keywords:** Superspreading, Dynamical System, Integro-Differential Equation

---

## 1. Introduction

The COVID-19 global epidemic has brought renewed focus on heterogeneous spread mechanisms and specifically on the “superspread” phenomenon [see, e.g., 1, 2, 3]. In qualitative epidemiological discussions, the term “superspread” is used in several distinct senses, which can often be found commingled in discussions of the phenomenon. In one sense, a “superspread” is an infected individual whose physiology results in higher infectivity than normal. Another sense of the term refers to events, rather than to individuals. In the context of the COVID-19 epidemic, superspread events such as weddings, funerals, or political rallies have been frequently cited as likely accelerators of the epidemic. “Superspread” may also refer to residential or geographic circumstances such as dense urban areas or assisted-care facilities, where spread is observed to be more rapid than elsewhere. Moreover, the term may refer to behaviors such as frequent attendance at bars, restaurants, and religious services. Obviously, some overlap between these categories exists and explains their frequent conflation in discussions of the phenomenon.

In technical/mathematical epidemiology the first sense of the term appears to dominate modeling. Efforts to model multiple spread rates through systems of ordinary differential equations (ODEs) [4], partial differential equations [5, 6, Chap. 12], delay-differential equations [7], stochastic differential equations [8], or integro-differential equations [9] rely on the first, individual-physiological notion of infectivity, because such models borrow the compartmental structure of classic epidemiological models such as the SIR model of Kermack and McKendrick [10], with discrete compartments representing groupings of susceptible, exposed, infected, and recovered/removed individuals. Dynamically speaking, in these models the compartments are fields over time, or over space-time. Such models differ in compartment selection and in dynamical rules for transferring individuals between compartments, but the role and interpretation of the compartments themselves are invariant.

Efforts have been made to model superspread *events* (as opposed to individuals) that do not change the dynamical stage of compartment fields over time and possibly space and that rely on stochastic source terms to model high-spread events [11, 12]. While these efforts model spread variability by using stochastic processes, they do not appear to explore

---

\*Corresponding Author  
Email address: cgraziani@anl.gov (Carlo Graziani)

behavior explicitly. Herein is a further subtle distinction in the use of “superspread”: we think of the term differently when we refer to “superspread events” and “superespread behavior.” To speak of a set of “events” evokes, in the mind of a modeler, a stochastic process, which is readily superposed on the existing dynamical infrastructure. A set of “behaviors,” on the other hand, calls to mind stable, deterministic relations and leads more naturally to a replacement of the dynamical stage of time, or space-time, with a new stage that we refer to below as “behavior space.”

We are interested in exploring the possible structure of a deterministic epidemic model embodying the second meaning of the term “superspread”: a model in which behaviors—rather than specific categories of individuals—are assigned different infectivities, and individuals become infected to the degree that they participate in such activities. We present one such model here: the *behavioral multispread* (BMS) model.

The plan of this article is as follows. In Section §2 we introduce the BMS model equations and discuss their interpretation. In §3 we derive separable solutions of the BMS equations. In §4 we discuss the stability of the infection-free solutions and of the separable solutions, discuss the nature of herd immunity in the BMS model, and consider late-time asymptotic behavior. In §5 we illustrate the nature of the BMS model in the context of a 3-behavior model featuring a very conservative (i.e., nonpropagating) behavior, a marginally propagating behavior, and a superspread behavior. A discussion and summary follows in §6. Some technical theorems concerning local and global stability of the BMS equations are proven in Appendix A and Appendix B.

## 2. The Model

### 2.1. Framework

The BMS model assumes a population characterized by the degree of participation in each of  $D$  behaviors or activities. An individual is characterized by a vector  $\mathbf{f}$  on the  $D$ -simplex, that is, satisfying

$$\mathbf{1}^T \mathbf{f} = \sum_{l=1}^D f_l = 1, \quad (2.1)$$

and  $f_l \geq 0$ ,  $l = 1, \dots, D$ . Each component  $f_l$  is interpreted as the fraction of time spent by the individual at the  $l$ th behavior. The simplex condition enforces the condition that all of the individual’s time is accounted for. In Equation (2.1) we have introduced the “one” vector  $\mathbf{1}$  whose components are all equal to 1.

Each behavior is characterized by its own infectivity, which we parametrize by a parameter  $a_l$ ,  $l = 1, \dots, D$ . The meaning of these parameters and their relation to epidemiological reproduction numbers will be made clear below. It is convenient to define the diagonal matrix  $\mathbf{A} \equiv \text{Diag}(a_1, \dots, a_D)$ .

We require a normalized population density on the  $D$ -simplex. For this purpose we introduce a  $(D - 1)$ -dimensional measure  $d\eta(\mathbf{f})$ , which we may conveniently define by  $d\eta(\mathbf{f}) \equiv d^D \mathbf{f} \times \delta(\mathbf{1}^T \mathbf{f} - 1) = \prod_{l=1}^{D-1} df_l$ , where  $\delta(x)$  is the Dirac point-mass distribution “density.” We define a population density  $\rho(\mathbf{f})$  and measure  $d\mu(\mathbf{f}) \equiv \rho(\mathbf{f})d\eta(\mathbf{f})$ , satisfying  $\int_{S_D} d\mu(\mathbf{f}) = 1$ , where  $S_D$  is the  $D$ -simplex. With these definitions, we may easily choose a standard form for  $\rho(\mathbf{f})$  such as a member of the family of Dirichlet distributions, or a mixture thereof. Alternative and more general distributions on the simplex with well-understood properties are known from the theory of statistical analysis of compositional data [13]. In the BMS model the measure  $d\mu(\mathbf{f})$  is not dynamical but, rather, is regarded as part of the fixed structure of the model. The kinematic structure of the BMS model is established by fixing the triplet  $(S_D, \mathbf{A}, d\mu(\mathbf{f}))$ .

The dynamical structure of the BMS model is defined in terms of the local susceptible, infected, and removed fractions  $s(t, \mathbf{f})$ ,  $i(t, \mathbf{f})$ , and  $r(t, \mathbf{f})$ , respectively, satisfying  $s + i + r = 1$ . These functions are defined in a manner similar to the corresponding fractions from the well-known SIR compartmental model [10, 6]. Their interpretation is that  $s(t, \mathbf{f})$  is the fraction of the population characterized by the simplex point  $\mathbf{f}$  that is susceptible to infection, and similarly for  $i(t, \mathbf{f})$  and  $r(t, \mathbf{f})$ . The total susceptible and infected compartment fractions in the population are thus  $S(t) = \int d\mu(\mathbf{f}) s(t, \mathbf{f})$  and  $I(t) = \int d\mu(\mathbf{f}) i(t, \mathbf{f})$ , with the removed fraction being simply  $R = 1 - S - I$  as usual.

We assume that individuals are removed at a uniform rate  $v$  from the infected state at each  $\mathbf{f}$ . The parameter  $v$  furnishes a natural time scale, so we will use that scale as a unit of time, defining the dimensionless time variable  $\tau \equiv vt$

## 2.2. Dynamical System

The model dynamics for the fractions  $i$  and  $s$  are given by the following equations:

$$\frac{\partial s}{\partial \tau}(\tau, \mathbf{f}) = -s(\tau, \mathbf{f}) \times \mathbf{f}^T \mathbf{A} \int_{S_D} d\mu(\mathbf{f}') \mathbf{f}' i(\tau, \mathbf{f}') \quad (2.2)$$

$$\frac{\partial i}{\partial \tau}(\tau, \mathbf{f}) = -i(\tau, \mathbf{f}) + s(\tau, \mathbf{f}) \times \mathbf{f}^T \mathbf{A} \int_{S_D} d\mu(\mathbf{f}') \mathbf{f}' i(\tau, \mathbf{f}'). \quad (2.3)$$

These equations are inspired by the SIR model with demography neglected, so that natural births and deaths not considered. Equations (2.2–2.3) are of easy interpretation. Equation (2.2) states that susceptibles at the point  $\mathbf{f} \in S_D$  are infected by contact with infected individuals at other points  $\mathbf{f}' \in S_D$  at a rate equal to  $d\mu(\mathbf{f}') \times \mathbf{f}^T \mathbf{A} \mathbf{f}' = d\mu(\mathbf{f}') \times \sum_{l=1}^D a_l f_l f'_l$ , which is to say at a sum of rates, the  $l$ th of which is in proportion to the respective fractions spent in behavior  $l$  by the susceptible individuals indexed by  $\mathbf{f}$  and by the infected individuals indexed by  $\mathbf{f}'$ . Equation (2.3) states that all infections of susceptibles at  $\mathbf{f}$  result in the creation of infected individuals at  $\mathbf{f}$  and that the fraction of such infected individuals decays at the removal rate, which is 1 with respect to the dimensionless time  $\tau$ .

The parameter  $a_l$  plays a role analogous to a reproduction number for behavior  $l$ . It will be larger for riskier behaviors than for more conservative behaviors. By integrating Equations (2.2–2.3) with respect to  $d\mu(\mathbf{f})$  we may define a time-dependent “effective reproduction number”  $R_E(\tau)$  for the system:

$$\begin{aligned} R_E(\tau) &\equiv \frac{\left[ \int_{S_D} d\mu(\mathbf{f}) s(\tau, \mathbf{f}) \mathbf{f}^T \right] \mathbf{A} \left[ \int_{S_D} d\mu(\mathbf{f}') i(\tau, \mathbf{f}') \mathbf{f}' \right]}{S(t) I(t)} \\ &= \sum_{l=1}^D a_l \left[ \int_{S_D} d\mu(\mathbf{f}) S(\tau)^{-1} s(\tau, \mathbf{f}) f_l \right] \left[ \int_{S_D} d\mu(\mathbf{f}') I(\tau)^{-1} i(\tau, \mathbf{f}') f_l \right], \end{aligned} \quad (2.4)$$

in terms of which we have the SIR-like effective equations

$$\frac{dS}{d\tau} = -R_E(\tau) S(\tau) I(\tau), \quad (2.5)$$

$$\frac{dI}{d\tau} = -I(\tau) + R_E(\tau) S(\tau) I(\tau). \quad (2.6)$$

These equations require the solution of the full system of Equations (2.2–2.3) to obtain  $R_E(\tau)$  and hence are not practical tools for obtaining such solutions. They are nonetheless helpful, illustrating the relationship of the BMS model to the SIR model. To gain intuition concerning the relation of the parameters  $a_l$  and  $R_E(\tau)$ , suppose that both the weighted averages

$$\mathbf{f}_s \equiv S(\tau)^{-1} \int_{S_D} d\mu(\mathbf{f}) s(\tau, \mathbf{f}) \mathbf{f} \quad (2.7)$$

$$\mathbf{f}_i \equiv I(\tau)^{-1} \int_{S_D} d\mu(\mathbf{f}) i(\tau, \mathbf{f}) \mathbf{f} \quad (2.8)$$

are located near the simplex uniform mean point,  $\mathbf{f}_s \approx \mathbf{f}_i \approx D^{-1} \mathbf{1}$ . In this case we would obtain  $R_E(\tau) \approx D^{-2} \text{Tr} \mathbf{A} = D^{-1} \bar{a}$ , where  $\bar{a}$  is the unweighted average of the  $a_l$ .

In §3 we show the existence of true SIR-type separable solutions of the BMS equations, and we establish the relationship between the parameters  $a_l$  and the basic reproduction number of those solutions. This will be convenient since the resulting relation will not depend on specific forms of the functions  $i(\cdot)$ ,  $f(\cdot)$ .

Equations (2.2–2.3) may be solved numerically by discretizing the simplex  $S_D$  and writing the resulting system as a (high-dimensional) ODE. In §5 we demonstrate this method of solution in a  $D = 3$  system using a triangularizing discretization of the simplex.

### 3. Separable Solutions

We assume the existence of separable solutions of the form

$$s(\tau, \mathbf{f}) = S(\tau) \times \tilde{s}(\mathbf{f}) \quad (3.1)$$

$$i(\tau, \mathbf{f}) = I(\tau) \times \tilde{i}(\mathbf{f}), \quad (3.2)$$

where the functions  $\tilde{s}(\mathbf{f})$  and  $\tilde{i}(\mathbf{f})$  are normalized:

$$\int_{S_D} d\mu(\mathbf{f}) \tilde{s}(\mathbf{f}) = \int_{S_D} d\mu(\mathbf{f}) \tilde{i}(\mathbf{f}) = 1. \quad (3.3)$$

Substituting these into Equation (2.2), we obtain

$$\frac{1}{IS} \frac{dS}{d\tau} \tilde{s}(\mathbf{f}) = -\tilde{s}(\mathbf{f}) \mathbf{f}^T \mathbf{A} \int_{S_D} d\mu(\mathbf{f}') \mathbf{f}' \tilde{i}(\mathbf{f}'). \quad (3.4)$$

Since the right-hand side of Equation (3.4) is independent of  $\tau$ , we must have

$$\frac{1}{IS} \frac{dS}{d\tau} = -R_0, \quad (3.5)$$

where  $R_0$  is a constant independent of  $\tau$ . We may use this to rewrite Equation (2.3) as follows:

$$\begin{aligned} \frac{dI}{d\tau} \tilde{i}(\mathbf{f}) &= -I(\tau) \tilde{i}(\mathbf{f}) - \frac{dS}{d\tau} \tilde{s}(\mathbf{f}) \\ &= -I(\tau) \tilde{i}(\mathbf{f}) + R_0 I(\tau) S(\tau) \tilde{s}(\mathbf{f}). \end{aligned} \quad (3.6)$$

This establishes that the functions  $\tilde{s}(\mathbf{f})$  and  $\tilde{i}(\mathbf{f})$  are proportional to each other:  $\tilde{i}(\mathbf{f}) = \chi(\tau) \tilde{s}(\mathbf{f})$ . However, by the normalization condition of Equation (3.3) it follows that  $\chi(\tau) = 1$ , and hence

$$\tilde{i}(\mathbf{f}) = \tilde{s}(\mathbf{f}). \quad (3.7)$$

Equations (3.5) and (3.6) now reduce to

$$\frac{dS}{d\tau} = -R_0 IS \quad (3.8)$$

$$\frac{dI}{d\tau} = -I + R_0 IS, \quad (3.9)$$

which is just the SIR system of equations, with  $R_0$  recognizable as the basic reproduction number.

We now combine Equations (3.4), (3.5), and (3.7) to obtain

$$\left( R_0 - \mathbf{f}^T \mathbf{A} \int_{S_D} d\mu(\mathbf{f}') \mathbf{f}' \tilde{s}(\mathbf{f}') \right) \tilde{s}(\mathbf{f}) = 0. \quad (3.10)$$

Since this equation holds for all  $\mathbf{f} \in S_D$ , it follows that  $\mathbf{f}^T \mathbf{A} \int_{S_D} d\mu(\mathbf{f}') \mathbf{f}' \tilde{s}(\mathbf{f}') = R_0$  for all  $\mathbf{f} \in S_D$ , which is possible only if

$$\int_{S_D} d\mu(\mathbf{f}') \mathbf{f}' \tilde{s}(\mathbf{f}') = \mathbf{f}_s = R_0 \mathbf{A}^{-1} \mathbf{1}. \quad (3.11)$$

Since  $\mathbf{1}^T \mathbf{f} = 1$ , we obtain the relation

$$R_0^{-1} = \mathbf{1}^T \mathbf{A}^{-1} \mathbf{1} = \sum_{l=1}^D a_l^{-1}. \quad (3.12)$$

Using Equations (2.4), (3.1), (3.2), (3.11), and (3.12), one can easily demonstrate that for a separable solution  $R_E(\tau) = R_0$ .

Thus,  $\tilde{s}(\mathbf{f})$  may not be chosen freely but must satisfy an integral constraint on the value of  $f_s$  given by Equations (3.11) and (3.12) in terms of the parameters  $a_l$ . The nature of the constraint is that  $f_s$ —the centroid on  $S_D$  of the measure  $\tilde{s}(\mathbf{f})d\mu(\mathbf{f})$ —has components  $f_{sl}$  that are inversely proportional to the  $l$ -th infectivity parameter  $a_l$ . Thus the more infective the  $l$ th behavior, the smaller the concentration of  $\tilde{s}(\mathbf{f})d\mu(\mathbf{f})$  and of  $\tilde{i}(\mathbf{f})d\mu(\mathbf{f})$  in regions of  $S_D$  where  $f_l \approx 1$ , and hence the greater the concentration of removed individuals  $\tilde{r}(\mathbf{f})d\mu(\mathbf{f}) = (1 - \tilde{s}(\mathbf{f}) - \tilde{i}(\mathbf{f}))d\mu(\mathbf{f}) = (1 - 2\tilde{s}(\mathbf{f}))d\mu(\mathbf{f})$  in such regions. We may interpret this observation if we may regard a separable solution as an asymptotic limit of a solution started at an arbitrary state—say, almost 100% of the population susceptible except for a small infected population at  $\mathbf{f}_0$ . Then we may say that the most infective regions Using Equations (2.4), (3.1), (3.2), (3.11), and (3.12), one can easily demonstrate that for a separable solution  $R_E(\tau) = R_0$ . If  $S_D$  are the ones most thoroughly emptied of susceptibles, who are replaced by removed/recovered individuals.

In order for such an interpretation to be tenable, the stability of asymptotic solutions of the BMS equations must be studied. The required analysis follows in §4. It will show that general solutions of the BMS equations in fact are initially attracted towards separable solutions. However separable solutions are not globally stable, whereas infection-free solutions are globally stable. As a consequence, initial convergence towards a separable solution is cut off by the advent of convergence towards an infection-free solution, a phenomenon we will refer to as *freeze-out* below.

## 4. Stability of Solutions

### 4.1. Stability of the Infection-Free Solution

Recall that the SIR equations (3.8) and (3.9) have an infection-free solution  $S = S_0$ ,  $I = 0$ , which is globally stable if  $S_0 R_0 \leq 1$  (herd immunity) and unstable otherwise [6, Chap. 2].

The BMS equations have an analogous special solution in which  $s(\tau, \mathbf{f}) = S_0 \tilde{s}_0(\mathbf{f})$ ,  $i(\tau, \mathbf{f}) = 0$ , where the function  $\tilde{s}_0(\mathbf{f})$  is normalized as in Equation (3.3). This solution is stable only under certain conditions that we now derive. Linearizing by

$$s(\tau, \mathbf{f}) = S_0 \tilde{s}_0(\mathbf{f}) + e^{\lambda\tau} \hat{s}(\mathbf{f}) \quad (4.1)$$

$$i(\tau, \mathbf{f}) = e^{\lambda\tau} \hat{i}(\mathbf{f}) ; \quad \hat{i}(\mathbf{f}) \geq 0 \quad (4.2)$$

and substituting in Equations (2.2) and (2.3), we find to linear order

$$\lambda \hat{s}(\mathbf{f}) = -S_0 \tilde{s}_0(\mathbf{f}) \mathbf{f}^T \mathbf{A} \int_{S_D} d\mu(\mathbf{f}') \mathbf{f}' \hat{i}(\mathbf{f}') \quad (4.3)$$

$$\lambda \hat{i}(\mathbf{f}) = -\hat{i}(\mathbf{f}) + S_0 \tilde{s}_0(\mathbf{f}) \mathbf{f}^T \mathbf{A} \int_{S_D} d\mu(\mathbf{f}') \mathbf{f}' \hat{i}(\mathbf{f}'). \quad (4.4)$$

Adding these equations, we immediately obtain  $(\lambda + 1)\hat{i}(\mathbf{f}) = -\lambda \hat{s}(\mathbf{f})$ . The condition  $\hat{i}(\mathbf{f}) \geq 0$  rules out nontrivial solutions for both  $\lambda = 0$  and  $\lambda + 1 = 0$ . We may write

$$\hat{i}(\mathbf{f}) = -\frac{\lambda}{\lambda + 1} \hat{s}(\mathbf{f}). \quad (4.5)$$

Combining this with Equation (4.3), we find

$$(\lambda + 1)\hat{s}(\mathbf{f}) = S_0 \tilde{s}_0(\mathbf{f}) \mathbf{f}^T \mathbf{A} \int_{S_D} d\mu(\mathbf{f}') \mathbf{f}' \hat{s}(\mathbf{f}'). \quad (4.6)$$

It is apparent from Equation (4.6) that the dependence on  $\mathbf{f}$  of  $\hat{s}(\mathbf{f})$  requires that there exist an  $\mathbf{f}$ -independent  $D$ -vector  $\mathbf{q}_\lambda$  such that

$$\hat{s}(\mathbf{f}) = \tilde{s}_0(\mathbf{f}) \mathbf{f}^T \mathbf{q}_\lambda. \quad (4.7)$$

Inserting this back in Equation (4.6) results in

$$(\lambda + 1)\mathbf{f}^T \mathbf{q}_\lambda = S_0 \mathbf{f}^T \mathbf{A} \int_{S_D} d\mu(\mathbf{f}') \tilde{s}_0(\mathbf{f}') \mathbf{f}' \mathbf{f}'^T \mathbf{q}_\lambda, \quad (4.8)$$

which can hold for all  $\mathbf{f}$  only if  $\mathbf{q}_\lambda$  is an eigenvector of the operator on the RHS, that is, if

$$\left( \mathbf{A} \int_{S_D} d\mu(\mathbf{f}') \tilde{s}_0(\mathbf{f}') \mathbf{f}' \mathbf{f}'^T \right) \mathbf{q}_\lambda = S_0^{-1}(\lambda + 1) \mathbf{q}_\lambda. \quad (4.9)$$

Since  $d\mu(\mathbf{f})\tilde{s}_0(\mathbf{f})$  is a probability measure, we may appeal to Theorem 2 in Appendix [Appendix A](#). Defining the mean vector  $\mathbf{h} \equiv \int_{S_D} d\mu(\mathbf{f}) \tilde{s}_0(\mathbf{f}) \mathbf{f}$  with components  $h_l$ , and noting that in general  $\mathbf{h} \neq R_0 \mathbf{A}^{-1} \mathbf{1}$  (that is, Equation 3.11 is not required to hold), by Theorem 2 we have that  $\lambda$  is real and that

$$0 \leq S_0^{-1}(\lambda + 1) \leq \bar{\beta}, \quad (4.10)$$

where we have defined the bound  $\bar{\beta}$  as

$$\bar{\beta} \equiv \max_{1 \leq l \leq D} (h_l a_l). \quad (4.11)$$

From Equation (4.10) we see that

$$\lambda \leq \bar{\beta} S_0 - 1, \quad (4.12)$$

and that the herd immunity condition  $\lambda < 0$  can always be achieved by making  $S_0$  sufficiently small that

$$\bar{\beta} S_0 \leq 1, \quad (4.13)$$

similarly to the herd immunity condition of the SIR model, but with  $\bar{\beta}$  in the role of  $R_0$ . This is now a statement not only about the value of  $S_0$  but also (through  $\bar{\beta}$ ) about the distribution density  $\tilde{s}_0(\mathbf{f})$  of susceptibles on  $S_D$ . If there is no epidemic and if susceptibles are arranged so that their mean  $\mathbf{h}$  has components such that  $\bar{\beta} S_0 < 1$ , a new outbreak will not cause a propagating epidemic.

In the context of an ongoing epidemic, we may view  $\bar{\beta} = \bar{\beta}(\tau)$  as a function of time, constructed from the current normalized state  $\tilde{s}(\tau, \mathbf{f})$ , that is,

$$\bar{\beta}(\tau) \equiv \max_{1 \leq l \leq D} (f_{sl}(\tau) a_l), \quad (4.14)$$

where  $f_{sl}(\tau)$  is the  $l$ -th component of the moment  $\mathbf{f}_s(\tau)$  defined in Equation (2.7). By Theorem 3 of Appendix [Appendix B](#) we have the herd immunity condition for an in-progress epidemic:

$$\bar{\beta}(\tau) S(\tau) < 1. \quad (4.15)$$

If Equation (4.15) is satisfied, then by Theorem 3 the solution is guaranteed to tend to an infection-free solution  $i = 0$  as  $\tau \rightarrow \infty$ .

Thus the BMS herd immunity condition has meaning both as a local and a global stability condition, analogously to the herd immunity condition of the SIR model [6].

#### 4.2. Stability of Separable Solutions

The local stability analysis of the separable solutions derived in §3 is similar to that of the infection-free solution. We start with the linearization

$$s(\tau, \mathbf{f}) = S(\tau) \tilde{s}(\mathbf{f}) + e^{\eta\tau} \hat{s}(\mathbf{f}) \quad (4.16)$$

$$i(\tau, \mathbf{f}) = I(\tau) \tilde{i}(\mathbf{f}) + e^{\eta\tau} \hat{i}(\mathbf{f}) \quad (4.17)$$

which perturbs the solutions described by Equations (3.1–3.2) by  $e^{\eta\tau} [\hat{s}(\mathbf{f}), \hat{i}(\mathbf{f})]$ .

Since we are perturbing about a time-dependent solution, the linearized perturbation equations contain time-dependent parameters, and we anticipate obtaining eigenvalues  $\eta(t)$  that are time-dependent. It is therefore necessary to impose a scale-separation condition on the perturbations, to ensure that the results make sense. In particular, the time-dependence of the perturbations is  $e^{\eta(\tau)\tau}$ , which can only be interpreted as an exponential decay or growth if  $d\eta/d\tau$  is in some sense

“small.” To make this precise, we Taylor-expand the exponential at time  $\tau_0$ . At time  $\tau = \tau_0 + \Delta\tau$ , the perturbation time-dependence is

$$\exp \{ \eta(\tau_0 + \Delta\tau)\tau \} \approx \exp \left\{ \left[ \eta(\tau_0) + \frac{d\eta}{d\tau}(\tau_0)\Delta\tau \right] \tau \right\}. \quad (4.18)$$

In order for the exponential decay (or growth) not to be spoiled, the second term in the exponential of Equation (4.18) must be small compared to the first over the course of an exponentiation time  $\Delta\tau = \eta(\tau_0)^{-1}$ . We must therefore require the scale separation condition

$$\left| \frac{1}{\eta(\tau_0)^2} \frac{d\eta}{d\tau}(\tau_0) \right| \ll 1, \quad (4.19)$$

in order for the linear perturbation analysis to make sense. We will check this condition explicitly for each of the perturbation modes turned up by the analysis that follows.

Substituting in Equations (2.2) and (2.3), we find to linear order

$$\begin{aligned} \eta \hat{s}(\mathbf{f}) &= -\hat{s}(\mathbf{f})I(\tau)\mathbf{f}^T \mathbf{A} \int_{S_D} d\mu(\mathbf{f}') \tilde{i}(\mathbf{f}')\mathbf{f}' - S(\tau)\tilde{s}(\mathbf{f})\mathbf{f}^T \mathbf{A} \int_{S_D} d\mu(\mathbf{f}') \hat{i}(\mathbf{f}')\mathbf{f}' \\ &= -R_0 \hat{s}(\mathbf{f})I(\tau) - S(\tau)\tilde{s}(\mathbf{f})\mathbf{f}^T \mathbf{A} \int_{S_D} d\mu(\mathbf{f}') \hat{i}(\mathbf{f}')\mathbf{f}' \end{aligned} \quad (4.20)$$

$$\begin{aligned} \eta \hat{i}(\mathbf{f}) &= -\hat{i}(\mathbf{f}) + \hat{s}(\mathbf{f})I(\tau)\mathbf{f}^T \mathbf{A} \int_{S_D} d\mu(\mathbf{f}') \tilde{i}(\mathbf{f}')\mathbf{f}' + S(\tau)\tilde{s}(\mathbf{f})\mathbf{f}^T \mathbf{A} \int_{S_D} d\mu(\mathbf{f}') \hat{i}(\mathbf{f}')\mathbf{f}' \\ &= -\hat{i}(\mathbf{f}) + R_0 \hat{s}(\mathbf{f})I(\tau) + S(\tau)\tilde{s}(\mathbf{f})\mathbf{f}^T \mathbf{A} \int_{S_D} d\mu(\mathbf{f}') \hat{i}(\mathbf{f}')\mathbf{f}', \end{aligned} \quad (4.21)$$

where we have used Equations (3.7) and (3.11) to produce the terms proportional to  $R_0$ . Summing these equations, we find again that  $\eta \hat{s}(\mathbf{f}) = -(\eta + 1)\hat{i}(\mathbf{f})$ .

#### 4.2.1. Modes Decaying at the Removal-Rate

If  $\eta = 0$ , we have the trivial solution  $\hat{s} = \hat{i} = 0$ . On the other hand, there is a decaying solution  $\eta = -1$ , corresponding to  $\hat{s}(\mathbf{f}) = 0$  and any  $\hat{i}(\mathbf{f})$  satisfying  $\int_{S_D} d\mu(\mathbf{f}') \hat{i}(\mathbf{f}')\mathbf{f} = 0$ . These solutions decay as  $e^{-\tau} = e^{-vt}$ , that is, at the SIR removal rate. Since in this case  $\frac{d\eta}{d\tau} = 0$ , the scale separation assumption of Equation (4.19) holds trivially.

#### 4.2.2. Zero-Moment Modes

Excluding  $\eta = -1$ , we again have that

$$\hat{i}(\mathbf{f}) = -\frac{\eta}{\eta + 1} \hat{s}(\mathbf{f}). \quad (4.22)$$

Inserting this relation into Equation (4.20), we can see that if  $\hat{s}(\mathbf{f})$  satisfies the eigenfunction relation

$$\tilde{s}(\mathbf{f})\mathbf{f}^T \mathbf{A} \int_{S_D} d\mu(\mathbf{f}') \mathbf{f}' \hat{s}(\mathbf{f}') = \beta \hat{s}(\mathbf{f}), \quad (4.23)$$

then  $\eta$  is a solution of the quadratic equation

$$\eta(\eta + 1) + (\eta + 1)R_0 I(\tau) - \eta\beta S(\tau) = 0. \quad (4.24)$$

There are two cases to consider. In the first case, the function  $\hat{s}(\mathbf{f})$  satisfies the zero-moment condition

$$\int_{S_D} d\mu(\mathbf{f}) \mathbf{f} \hat{s}(\mathbf{f}) = 0, \quad (4.25)$$

in which case  $\beta = 0$  and by Equation (4.24) we obtain

$$\eta = -R_0 I(\tau). \quad (4.26)$$

These modes are ostensibly stable. However, at late times, we know from Theorem 3 of Appendix Appendix B that  $I(\tau) \rightarrow 0$  exponentially. Therefore the ostensible exponential decay time of these modes to the separable solution becomes exponentially large at late times. Unsurprisingly, this occurrence is associated with the late-time failure of the scale separation condition, Equation (4.19):

$$\begin{aligned} \left| \frac{1}{\eta(\tau)^2} \frac{d\eta}{d\tau}(\tau) \right| &= \frac{1}{R_0 I(\tau)^2} \left| \frac{dI}{d\tau} \right| \\ &= \frac{|R_0 S(\tau) - 1|}{R_0 I(\tau)}, \end{aligned} \quad (4.27)$$

where we have used Equation (3.9) to obtain the last line. We therefore expect scale separation to break down for these modes when

$$1 - R_0 S(\tau) = R_0 I(\tau). \quad (4.28)$$

This fact is related to the phenomenon of *freeze-out*, which we discuss in more detail in §4.3.

#### 4.2.3. Finite-Moment Modes

The second case follows the solution procedure for Equation (4.6). From Equation (4.23), the  $f$ -dependence of  $\hat{s}(f)$  is necessarily of the form

$$\hat{s}(f) = \tilde{s}(f) f^T w_\beta, \quad (4.29)$$

where the  $D$ -vector  $w_\beta$  is independent of  $f$ . Inserting this into Equation (4.23), and requiring the result to be true for all  $f$ , we find the problem reduced to a  $D$ -dimensional eigenproblem:

$$\left[ A \int_{S_D} d\mu(f') \tilde{s}(f') f' f'^T \right] w_\beta = \beta w_\beta. \quad (4.30)$$

By appealing once again to Theorem 2 in the Appendix, we have that  $\beta$  is real and that

$$0 \leq \beta \leq R_0, \quad (4.31)$$

since now Equation (3.11) holds.

One solution to Equation (4.30) is immediate by Equations (3.11) and (2.1): There is an eigenvalue  $\beta = R_0$  corresponding to the right eigenvector  $w_{R_0} = \mathbf{1}$  and to the left eigenvector  $\mathbf{1}^T A^{-1}$ . Consequently other  $w_\beta$  should be sought in the orthogonal space  $\mathbf{1}^T A^{-1} w_\beta = 0$ . Moreover we learn from this that the upper bound of Inequality (4.31) is sharp, since it holds for at least one eigenvalue. From Equation (4.29), we have that for this eigenmode  $\hat{s}(f) \propto \tilde{s}(f)$ .

For each  $\beta$ , the solutions to Equation (4.24) are

$$\eta_\pm(\beta) = -\frac{1}{2} (1 + R_0 I(\tau) - \beta S(\tau)) \pm \frac{1}{2} \left[ (1 + R_0 I(\tau) - \beta S(\tau))^2 - 4R_0 I(\tau) \right]^{1/2}. \quad (4.32)$$

Clearly stability— $\Re[\eta_\pm] \leq 0$ —requires that for the largest  $\beta$  we have

$$1 + R_0 I(\tau) - \beta S(\tau) \geq 0. \quad (4.33)$$

Since we know that at least one eigenvalue achieves the upper bound in Inequality (4.31), the stability condition becomes

$$R_0 (S(\tau) - I(\tau)) \leq 1. \quad (4.34)$$

When this condition is achieved, we may expect all deviations from an initial-condition-dependent target separable solution to decrease exponentially at rates  $\eta_\pm(\beta)$  given by Equation (4.32). We note, however, that as  $I(\tau) \rightarrow 0$ , the  $\eta_+$  eigenvalues approach zero from below, since

$$\lim_{I(\tau) \rightarrow 0} \eta_+(\beta) = -\frac{R_0 I(\tau)}{1 - \beta S(\tau)} \rightarrow 0. \quad (4.35)$$

Such modes will not decay to zero rapidly late in the evolution of the epidemic but rather will “freeze out,” as discussed in the next subsection.

#### 4.3. Relative Stability of Infection-Free and Separable Solutions

By Part (ii) of Theorem 2, we may rewrite the condition for stability of the infection-free solutions, Equation (4.15), as

$$S(\tau) \leq \frac{1}{\bar{\beta}(\tau)} \leq \frac{1}{R_0}. \quad (4.36)$$

The condition for stability of a separable solution, Equation (4.34), is

$$S(\tau) \leq I(\tau) + \frac{1}{R_0}. \quad (4.37)$$

From Equations (4.36–4.37) we see that if a solution initially satisfies neither of the two stability conditions, then the solution will satisfy Equation (4.37) before it satisfies Equation (4.36). It follows that during the initial part of the solution we will have

$$\bar{\beta}(\tau)^{-1} < S(\tau) \leq R_0^{-1} + I(\tau), \quad (4.38)$$

corresponding to a situation in which attraction toward a separable solution begins before attraction toward the infection-free solution sets in. We may therefore infer that a *freeze-out* phenomenon must occur, because the eigenvalues of the infection-free solution behave as  $|\lambda| = |1 - \bar{\beta}(\tau)S(\tau)|$ , so that their decay rates continue to increase as  $S(\tau)$  declines irrespective of  $I(\tau)$ , whereas the zero-moment and finite-moments modes of the separable solutions have eigenvalues that tend to zero as  $I(\tau) \rightarrow 0$  according to Equations (4.26) and (4.35). These modes will therefore not efficiently decline toward their target solution at large  $\tau$ , and will be overtaken by the infection-free solution modes.

This phenomenon is evidently related to the fact that the infection-free solutions are globally stable, whereas the separable solutions are not. A related phenomenon is the failure of the scale separation condition for the validity of the perturbation analysis of the separable solutions. Scale separation fails for the moment-free solutions when Equation (4.28) first holds. But if the solution has had time to approach a separable solution, so that  $\bar{\beta} \approx R_0$ , then the left-hand side of Equation (4.28) is approximately the threshold eigenvalue for herd immunity, Equation (4.12), whereas the right-hand side is the eigenvalue of the moment-free modes of the separable solution. Thus loss of scale separation occurs coincidentally with the overtaking of separable mode eigenvalues by infection-free mode eigenvalues, that is, at freeze-out. In fact we may use Equation (4.28) as a rough-and-ready measure of when we may expect freeze-out to occur, since as  $I(\tau)$  declines past this point, the eigenvalues  $\lambda$  of the infection-free solutions become more negative than a moiety of the eigenvalues of the separable solutions.

We demonstrate the occurrence of the freeze-out phenomenon numerically in what follows.

## 5. Numerical Study

We performed numerical experiments with the BMS system using a  $D = 3$  model. We used a triangular discretization on  $S_3$  that is illustrated in Figure 5.1. In each case, the number of subdivisions along each axis is 16, so that there are 256 triangles. After discretization the BMS equations (2.2–2.3) take on the character of a 256-dimensional system of ODEs. Each triangle is ascribed its centroid value of  $f$  as its coordinate for the purpose of computing the source term—the right-hand side of the BMS equations. The integration code is written in Python. We perform the ODE integration using the `scipy.integrate.solve_ivp` method from the *SciPy* package [14]. With the above choice of discretization, all integrations take at most a few seconds running on a laptop computer. Our ternary plot visualizations are made by using *Ternary* [15], a third-party module for the *Matplotlib* [16] Python plotting library.

We assume throughout a normalized population density  $\rho(f)$  on the simplex of the Dirichlet type, specifically

$$d\mu(f) = \rho(f) df_1 df_2 \equiv \frac{1}{B(\alpha)} \prod_{l=1}^3 (f_l)^{\alpha_l-1} df_1 df_2, \quad (5.1)$$

where  $B(\alpha) = \left[ \Gamma \left( \sum_{l=1}^3 \alpha_l \right) \right]^{-1} \prod_{l=1}^3 \Gamma(\alpha_l)$  is a normalizing constant. We fix  $\alpha = [1.4, 1.2, 0.9]$ . The resulting density is shown in the right panel of Figure 5.1. The population-weighted centroid  $f_{pop}$ , shown as a red dot in the figure,

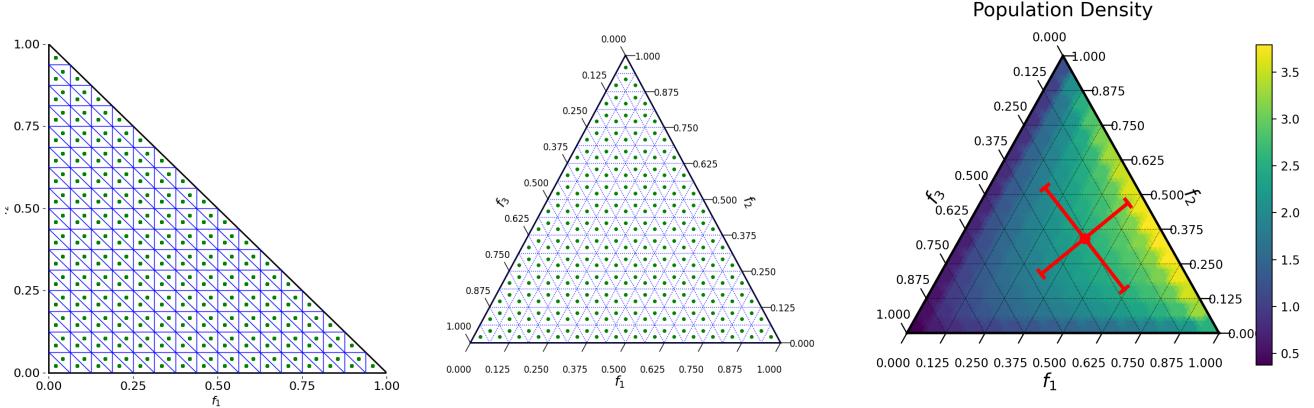


Figure 5.1: Left panel: Triangularizing discretization used for integration of BMS equations. The figure shows the right triangles in the  $f_1 - f_2$  plane (thin lines) and their centroids (green dots). Middle panel: Ternary plot displaying discretization on the simplex face. The lower-left to upper-right lines are contours of constant  $f_1$ , with values shown on the bottom axis. Similarly,  $f_2$  values are displayed on the right axis with left-right contours of constant  $f_2$ , and  $f_3$  values are displayed on the left axis with upper-left to lower-right contours of constant  $f_3$ . Right panel: Population density on the 3-simplex (Equation 5.1) used for the simulations. The red dot is the distribution centroid. The cross arms represent  $\pm 1$  standard deviations along the principal covariance directions.

while the cross shows the principal directions and the root-eigenvalues of the distribution covariance – that is, the cross arms represent  $\pm 1$  standard deviations along the principal covariance directions. The centroid has the value  $\mathbf{f}_{pop} = [0.398, 0.341, 0.260]^T$ . Clearly, this density is somewhat biased toward high values of  $f_1$ , moderate values of  $f_2$ , and low values of  $f_3$ . Below we will specify infectivity parameters  $a_l$  such that  $l = 1$  corresponds to a low-risk behavior,  $l = 2$  corresponds to moderate-low risk, and  $l = 3$  corresponds to high-risk superspread behavior. With these choices, the population is more heavily weighted to lower-risk behaviors 1 and 2 but is by no means free of high-risk behavior 3.

We restore dimensionality to time, choosing a recovery rate  $v = 0.1 \text{ days}^{-1}$ , corresponding to an infective period of 10 days.

### 5.1. Epidemic Propagation in Behavior Space

To illustrate the dynamics of epidemic propagation across behavior space—that is, across  $S_3$ —we consider an epidemic with parameters  $a_1 = 25/6$ ,  $a_2 = 25/3$ ,  $a_3 = 25$ , corresponding to a separable solution with  $R_0 = 2.5$  and a centroid  $\mathbf{f}_s = R_0 \mathbf{A}^{-1} \mathbf{1} = [0.6, 0.3, 0.1]^T$ . We may interpret these parameters using Equation (2.4), assuming  $\mathbf{f}_s = \mathbf{f}_i = \mathbf{f}_{pop}$ , the population centroid on  $S_3$ . The result is that  $R_E = R_{E,1} + R_{E,2} + R_{E,3} = 0.661 + 0.971 + 1.69 = 3.32$ , corresponding to a high-risk average 3-behavior, a medium-low-risk average 2-behavior, and a very low-risk average 1-behavior.

We start with a nearly infection-free population, namely,  $s(0, \mathbf{f}) = 1$ ,  $i(0, \mathbf{f}) = 0$  except at location  $\mathbf{f}_{i,init} = [0.95, 0.04, 0.01]$  (i.e., near the  $f_1 = 1$  vertex of  $S_3$ ), where we add to  $i(0, \cdot)$  (and subtract from  $s(0, \cdot)$ ) an amount  $\delta i$  corresponding to  $10^{-4}$  of the population, that is,  $\delta i = 10^{-4} / [\rho(\mathbf{f}_{i,init}) \times T]$ , where  $T$  is the area of each discretization right triangle in the  $f_1 - f_2$  plane (left panel of Figure 5.1). With this initial choice, the centroids of the initial susceptible and infected distributions are

$$\mathbf{f}_s(\tau = 0) = [0.398, 0.341, 0.260]^T \quad (5.2)$$

$$\mathbf{f}_i(\tau = 0) = [0.958, 0.021, 0.021]^T. \quad (5.3)$$

By Equation (2.4) we have  $R_E(\tau = 0) = R_{E,1} + R_{E,2} + R_{E,3} = 1.591 + 0.05928 + 0.1355 = 1.79$ . We can see that the population bias and initial condition have changed the infectivity order, with more initial infectivity ascribed to behavior 1 than to behavior 3.

The left panel of Figure 5.2 displays the progress of the epidemic integrated over  $S_3$ . The curves show the profiles of susceptible, infected, and removed populations as a function of time. For comparison we also show the progress of a classic SIR model, started with the same initial infected fraction ( $10^{-4}$ ), with the same  $v$ , and with the same  $R_0$

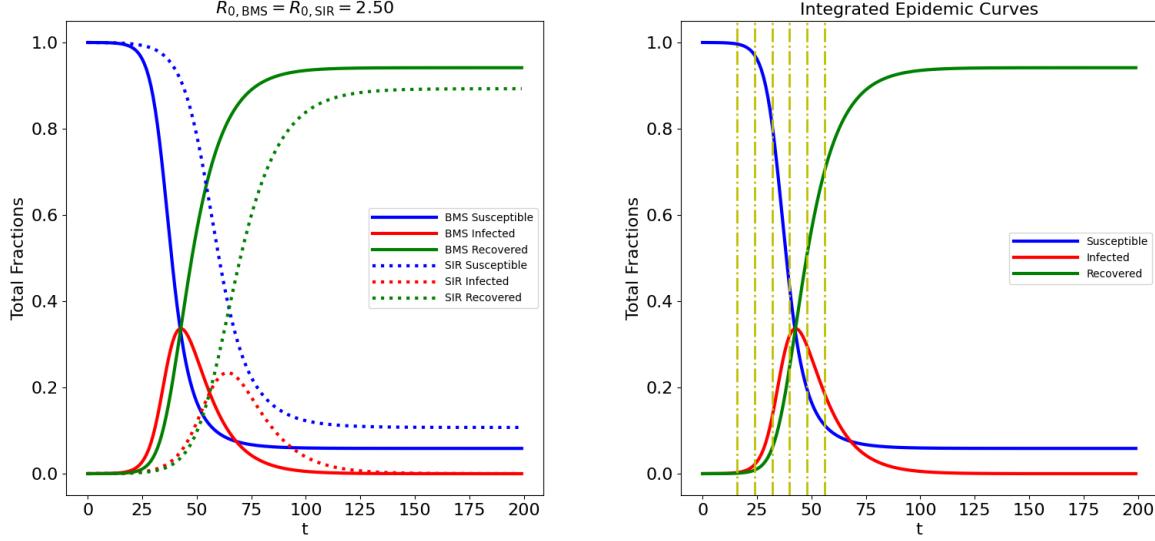


Figure 5.2: Integrated epidemic progress. Left panel: The solid lines show the susceptible, infected, and removed fractions for the integrated BMS equations initialized as described in §5.1. The dotted lines show the classic SIR model integrated from the same initial state, with the same parameters  $R_0$  and  $\nu$  as the separable-solution parameters of the BMS model. The comparison shows that the BMS model progresses more rapidly than the corresponding SIR model. Right panel: The same BMS model, showing the times at which the snapshots in Figure 5.3 were produced..

parameter as that corresponding to the separable solution of the BMS model. This figure shows that the BMS model progresses more rapidly than the SIR model. This may seem surprising since the initial  $R_{Eff}$  of the BMS model is 1.79, which is less than the  $R_0 = 2.5$  value used for the SIR model. However, we will see shortly that  $R_{Eff}$  evolves strongly over time and that there is no discrepancy here. The right panel of Figure 5.2 shows the same integrated BMS model, together with vertical dotted lines at the times corresponding to the snapshots of infected fraction distribution on  $S_D$  shown in Figure 5.3.

The snapshots in Figure 5.3 show infected fraction as a colormap. They also show the centroid and  $\pm 1$  standard deviation along principal covariance directions of the infectious distribution  $d\mu(f)\tilde{i}(\tau, f)$  as a red dot and cross. The first of these snapshots shows the initial infection site at the lower-right corner “reaching over” to infect the high-risk behavior individuals in the lower-left corner. This situation happens rapidly despite the small amount of communication between the two corners in the source term of the BMS equations. Thereafter, we see a wave of infection proceeding from high-risk to low-risk portions of  $S_3$ . Animated visualisations of the progress of all BMS variables across behavior space are available in the supplementary materials to this article, as well as at <https://mcs.anl.gov/~carlo/Articles/Multispread/>.

Figure 5.4 shows the final state of the susceptible population in behavior space. The epidemic has had the greatest effect on the lower-left corner of the space, where the highest-risk behavior is most prevalent. This corner has few remaining susceptibles, compared with the upper-right edge of the diagram, where less risky behavior is prevalent.

## 5.2. Effective Reproduction Numbers and Approach to Separable Solution

Figure 5.5 shows some further properties of the evolution of the BMS equations. The left panel of the figure shows the run of  $R_E$ , computed as per Equation (2.4), together with the runs of the three summands of  $R_E$ . We see here the resolution of question raised earlier: Why does the BMS solution rise so much more rapidly than the SIR solution, despite starting off with a lower effective reproduction number? As is apparent from the curves, the value of  $R_E$  climbs rapidly in the initial phase of the epidemic, driven largely by the contribution to  $R_E$  from the high-risk behavior contribution (the red dot-dash line). Thereafter the value of  $R_E$  settles down after the epidemic empties the high-risk behavior population of susceptibles. The predicted freeze-out time from Equation (4.28) is shown as a vertical dot-dashed line.

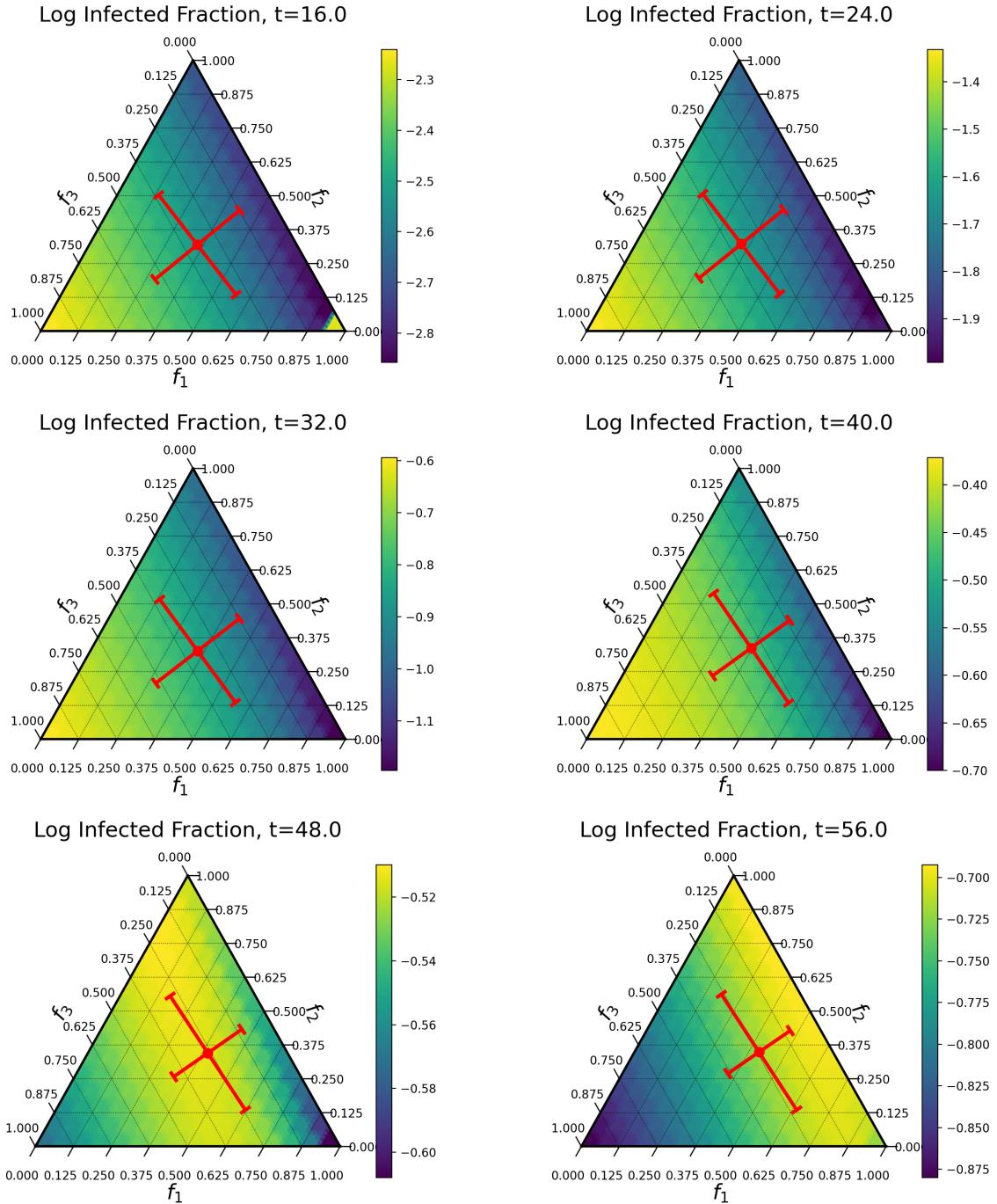


Figure 5.3: Progress of the epidemic in behavior space, at the times shown in the right panel of Figure 5.2. The panels show infected fraction as a colormap. They also show the centroid and  $\pm 1$  standard deviation along principal covariance directions of the infectious distribution  $d\mu(\mathbf{f}) i(\tau, \mathbf{f})$  as a red dot and cross.

Log Susceptible Fraction,  $t=80.0$

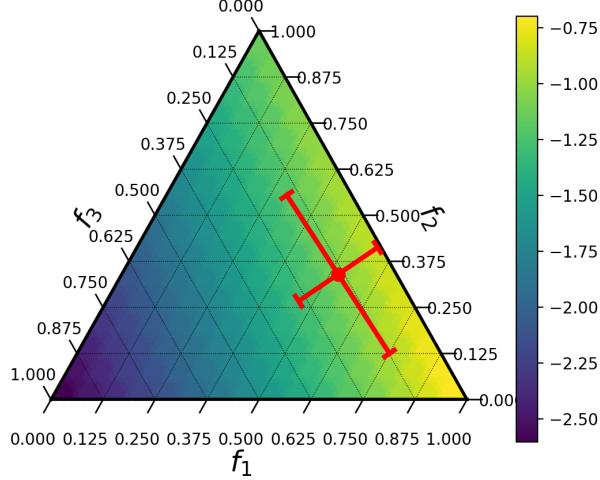


Figure 5.4: Final state of the susceptible population. The colormap shows the susceptible fraction, while the centroid and  $\pm 1$  standard deviation along principal covariance directions of the susceptible distribution  $d\mu(\mathbf{f}) \tilde{s}(\tau, \mathbf{f})$  are shown as a red dot and cross. Note that the lower-left corner—the high-risk portion of behavior space—is essentially emptied of susceptibles, whereas susceptibles remain in the lower-risk portion of the space.

The right panel of Figure 5.5 illustrates the freeze-out phenomenon. Here we see the average  $L_1$  residual of the mean vector  $f_s$  from the value predicted for the separable solution by Equation (3.11). We can see that the residual drops rapidly until the predicted freeze-out time, after which it settles toward a steady value.

### 5.3. Other Values of $R_0$

Figure 5.6 displays the integrated epidemic progress, effective reproduction numbers, and approach to separable solution for different  $R_0$  parameters— $R_0 = 1.0$  in the left column and  $R_0 = 1.5$  in the right column. The  $a_l$  parameters are simply scaled from their versions for  $R_0 = 2.5$ , according to Equation (3.12), preserving the risk ordering of the three behaviors.

In the  $R_0 = 1$  case, we observe that there is an epidemic notwithstanding the fact that the corresponding separable case—SIR with  $R_0 = 1$ —is dormant, as shown in the top figure. In fact, the system state never progresses far enough toward its separable solution to see an appreciable decline in effective reproduction number from the initial value, as is apparent in the middle panel. The lower panel again shows anemic progress toward the separable solution. In fact, in this case the freeze-out condition of Equation (4.28) is never attained.

The  $R_0 = 1.5$  case resembles the  $R_0 = 2.5$  case, although less progress occurs toward the separable solution.

### 5.4. An Actual Separable Solution

Given the importance of separable solutions to the theory presented in this work, an interesting verification is to obtain an actual separable solution numerically. We may obtain such a solution by the following simple expedient. Run a series of integrations of the BMS equations. At each stage, initialize the normalized shape functions  $\tilde{s}(t = 0, \mathbf{f})$ ,  $\tilde{i}(t = 0, \mathbf{f})$  to the final corresponding shape functions from the previous iteration, and reset the integrated variables  $S(t = 0) = 1 - 10^{-4}$ ,  $I(t = 0) = 10^{-4}$ . This procedure allows the convergence to a separable solution to restart and proceed further before freeze-out sets in again. The results of four iterations of this process using the same parameters and initial condition as in §5.1 are displayed in Figure 5.7. The left panel shows that the integrated quantities  $S(t)$ ,  $I(t)$ ,  $R(t)$  are now indistinguishable from the corresponding quantities in the corresponding SIR model, as expected. The middle panel displays the effective reproduction numbers, which are nearly constant and which add up to  $R_0 = 2.5$ , as expected. The right panel displays the final log normalized population density of the susceptibles,  $\tilde{s}(\mathbf{f})\rho(\mathbf{f})$ , where  $\rho(\mathbf{f})$  is the population density with respect to the measure  $df_1 df_2$  defined in §2.1. This distribution shows a pronounced tilt toward the low-risk-behavior sector that accords well with intuition about the progress of an epidemic.

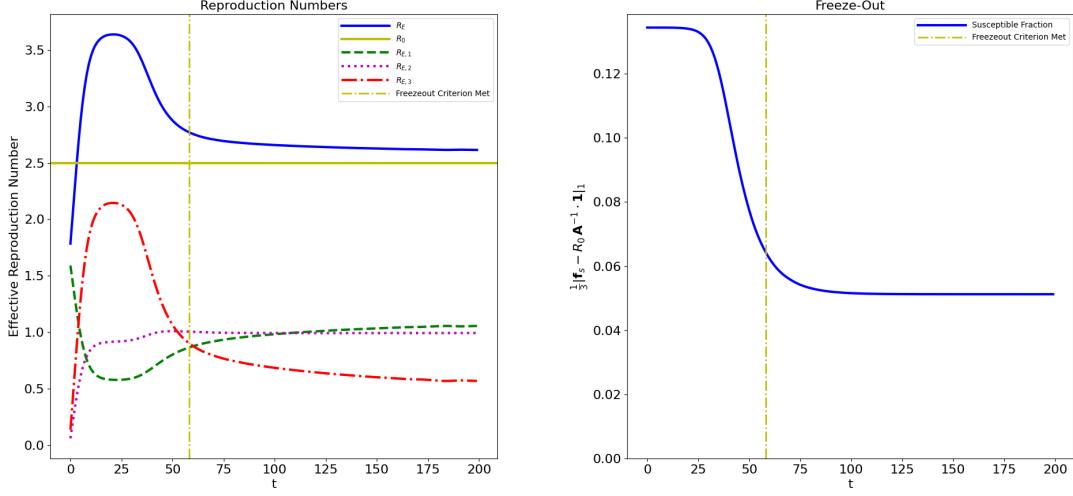


Figure 5.5: Left panel:  $R_E$  and its components. The vertical dot-dashed line shows the time at which the freeze-out criterion (Equation 4.28) is satisfied. The effective reproduction number  $R_E$  initially climbs to a high value, driven mostly by the high-risk behavior 3. Eventually  $R_E$  settles down, as the epidemic empties the high-risk-behavior region of  $S_3$  of susceptibles. Right panel: Approach to separable solution, measured by the deviation of the mean vector  $f_s$  from the predicted value for a separable solution (Equation 3.11). The vertical dot-dashed line again shows the predicted freeze-out time, which agrees with the flattening out of the curve.

## 6. Discussion

We have presented a new dynamical system, the behavioral multispread model, motivated by the desirability of extending compartmental models to behavior space so as to create a context where the behavioral usage of the term “superspread” has a technical-mathematical meaning, analogous to that corresponding to the individual-physiology and event usages of the term. In this work, we have largely focused on the mathematical properties of the BMS system, with particular regard to solution stability and long-term behavior, as well as comparisons with the ancestor of BMS, the SIR system. We demonstrated that the BMS model is, unsurprisingly, capable of exhibiting richer dynamical behavior than is SIR, with dynamically evolving effective reproduction numbers and a more complex description of herd immunity. It is particularly interesting to observe the evolution of  $R_E$ , since this represents transitions in the susceptible and infected distributions in behavior space. As these distributions change, the relative importance of the infectivities corresponding to different behaviors also change.

We now discuss the status of BMS in the field of mathematical epidemic models and sketch some possible directions for future research.

An unexplored mathematical topic in connection with the BMS model is to answer the questions “What happens when  $D$  gets very large, or even infinite?” Does there exist some useful limit, and if so, what are its properties?

At this stage, BMS is useful principally as an illustrative model of epidemic propagation in behavior space, rather than as a detailed model capable of comparison with epidemic data, let alone detailed prediction. In order to make BMS into an epidemiological model capable of realistic description and interpretation, some obvious research steps need to take place. In the first instance, it may be that  $D = 3$  is too simplistic a representation of behavior space. In fact there is likely an infinitude of risk behaviors, but it is not implausible that they could be grouped into a few coarse categories. After all, a behavior such as bar-hopping may well have a parameter  $a_l$  that is close to that of another behavior such as theater-going, so that the two activities look indistinguishable from the model’s point of view, and the only difference between our hypothetical populations of bar-hoppers and theater-goers is that their respective distributions in behavior space differ, because they spend different amounts of their time at these (and other) activities.

Once a reasonable set of risk-activity categories is determined, a second problem is determining the population density  $\rho(f)$  over the corresponding behavior space. This is evidently a sociological challenge, possibly to be addressed using data from surveys, censuses, and the like.

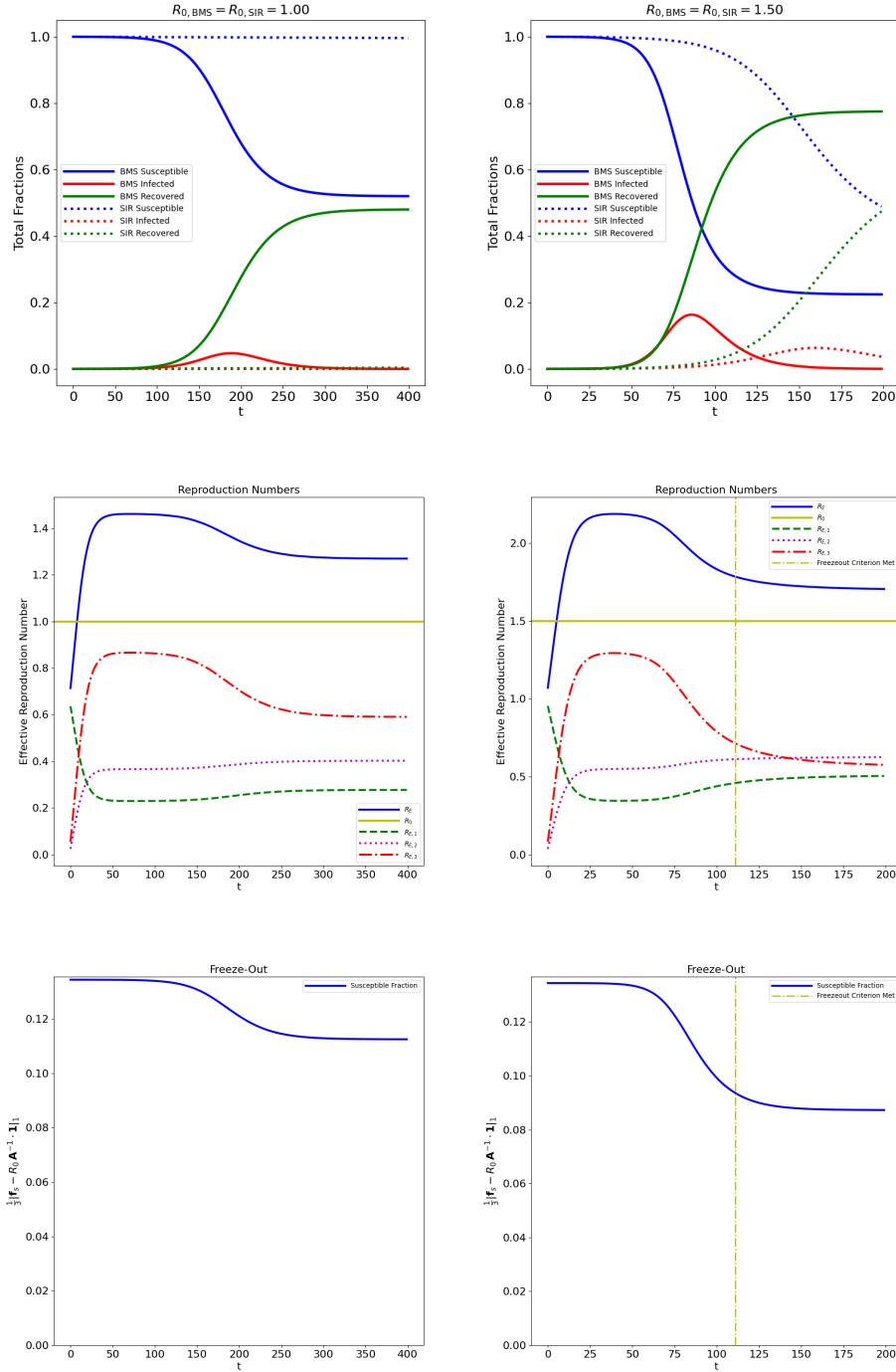


Figure 5.6: Behavior of BMS model for other values of  $R_0$ . Left column:  $R_0 = 1.0$ . Right column:  $R_0 = 1.5$ . The top row displays epidemic progress integrated over  $S_3$ . The middle column shows effective reproduction numbers. The bottom column shows approach to separable solution. The yellow dot-dashed line showing the advent of freeze-out is not displayed for  $R_0 = 1.0$  since the freeze-out condition of Equation (4.28) is never attained in this case.

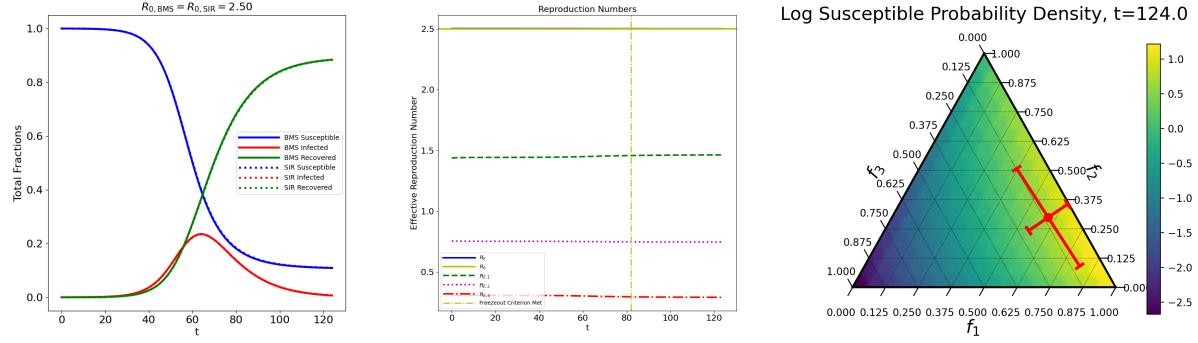


Figure 5.7: Separable solution, obtained by iterating 4 times the process of initializing each integration using the final normalized shape functions  $\tilde{s}(t_{final}, \mathbf{f})$ ,  $\tilde{i}(t_{final}, \mathbf{f})$  from the previous iteration, and resetting the integrated parameters  $S(t = 0) = 1 - 10^{-4}$ ,  $I(t = 0) = 10^{-4}$ . The left panel shows the integrated variables, which are now indistinguishable from their SIR counterparts. The middle panel shows effective reproduction numbers, which are essentially constant and which sum to  $R_0 = 2.5$ . The right panel shows the final population distribution density of susceptibles in behavior space, which shows a pronounced tilt in favor of the low-risk-behavior sector.

A technical issue that can arise is one of visualization. For this work, we chose  $D = 3$  because visualization of 3-component compositions on the 3-simplex is elegantly accomplished via ternary diagrams. Higher-dimensional visualizations would be more challenging and might be better accomplished by other types of diagrams.

Another issue is discretization. The 3-simplex  $S_3$  lends itself to a straightforward triangular decomposition, which can be understood and coded based on simple Euclidean geometry. Higher-dimensional “triangularizations” require input from more advanced computational geometry. A separate but related issue is stability of numerical integration. As carried out in this work, the integro-differential equations of the BMS model were summarily discretized, thus trading the MBS system in for a (large) system of ODEs. Fortunately for us, this worked flawlessly, and in the cases we examined we ran into no troublesome issues such as frequently attend naive discretizations. However, in this work we can offer no guarantees that this should be so irrespective of discretization resolution or system dimensionality. A study of the numerical properties of BMS would thus be of some value.

We point out that despite the aforementioned difficulties in using BMS as a realistic, interpretable model, there is no difficulty in its use as an empirical fitting model, and potentially some advantage to using it in this way. Given the rich dynamical behavior of BMS compared, say, with SIR, the ease with which the system can be numerically discretized, and the efficiency and speed of integration, there could be value in fitting parameters such as  $\alpha_l$ ,  $a_l$  to the progress of a real epidemic, assuming sufficient data is available, to investigate whether, for example, the BMS characteristic of a rapidly increasing and then settling-down effective reproduction number can provide a better description of certain types of epidemics.

## Acknowledgments

This material is based upon work supported by the U.S. Department of Energy, Office of Science, under contract number DE-AC02-06CH11357.

## Declaration of Interests

The author declares that he has no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Appendix A. Eigenvalues of Stability Matrix

We now give proof of the assertions made in §4 concerning the eigenvalues of the  $D$ -dimensional stability matrix  $M \equiv A \int_{S_D} d\mu(f) s(f) f f^T$ . We first establish a necessary lemma.

**Lemma 1.** Fix  $\mathbf{h} \in S_D$  and  $\mathbf{w} \in \mathbb{R}^D$ . Let  $\mathcal{M}(\mathbf{h})$  be the space of probability measures  $d\eta(f)$  on  $S_D$  satisfying  $\int_{S_D} d\eta(f) f = \mathbf{h}$ . Define the objective function  $O : \mathcal{M}(\mathbf{h}) \rightarrow \mathbb{R}^+$  by

$$O[d\eta] \equiv \int_{S_D} d\eta(f) (\mathbf{w}^T f)^2. \quad (\text{A.1})$$

Then the solution of the maximization problem

$$d\zeta_h(f) = \arg \max_{d\eta \in \mathcal{M}(\mathbf{h})} \{O[d\eta]\} \quad (\text{A.2})$$

is the ( $\mathbf{w}$ -independent) additive superposition of Dirac- $\delta$  measures supported at the simplex vertices given by

$$\int_{\Omega \subset S_D} d\zeta_h(f) g(f) \equiv \sum_{l=1}^D h_l g(\mathbf{e}_l) I[\mathbf{e}_l \in \Omega], \quad (\text{A.3})$$

where the  $\mathbf{e}_l$  are unit vectors in the  $l$ -direction, in other words,  $[\mathbf{e}_l]_m = \delta_{lm}$ , and  $I[]$  is the indicator function.

It is straightforward to establish that  $d\zeta_h(f)$  is a probability measure that satisfies the mean constraint, since  $d\zeta_h(f)$  is manifestly non-negative definite,  $\int_{S_D} d\zeta_h(f) = \sum_{l=1}^D h_l = 1$  since  $\mathbf{h} \in S_D$ , and  $\int_{S_D} d\zeta_h(f) f = \sum_{l=1}^D h_l \mathbf{e}_l = \mathbf{h}$ . For the measure  $d\zeta_h(f)$  the objective of Equation (A.1) becomes  $\sum_{l=1}^D h_l w_l^2$ . We will establish the lemma by proving that any finite perturbation of  $d\zeta_h(f)$  that preserves the mean constraint leads to a diminution of the objective function from this value.

We define the perturbed measure  $d\eta(f) = d\zeta_h(f) + d\gamma(f)$ , where the measure  $d\gamma(f)$  may be an additive mixture of a measure that is absolutely continuous with respect to the reference measure  $d\mu(f)$  (i.e., admitting a Radon-Nikodym density function) and some singular measures such as Dirac- $\delta$  components, and including Dirac- $\delta$  components at the simplex vertices that adjust the corresponding components of  $d\zeta_h(f)$ . In other words,

$$\int_{\Omega \subset S_D} d\gamma(f) g(f) = \sum_{l=1}^D b_l g(\mathbf{e}_l) I[\mathbf{e}_l \in \Omega] + \int_{\Omega \subset S_D} d\tilde{\gamma}(f) g(f), \quad (\text{A.4})$$

where the non-negative-definite measure  $d\tilde{\gamma}(f)$  has support excluding the simplex vertices and may itself be an additive superposition of absolutely continuous and singular measures.

Imposing the mean constraint, we have

$$\begin{aligned} 0 &= \int_{S_D} d\gamma(f) f \\ &= \sum_{l=1}^D b_l \mathbf{e}_l + \int_{S_D} d\tilde{\gamma}(f) f, \end{aligned}$$

whence

$$b_l = - \int_{S_D} d\tilde{\gamma}(f) f_l. \quad (\text{A.5})$$

The perturbation to the objective function of Equation (A.2) is therefore

$$\begin{aligned} \delta O &= \int_{S_D} d\gamma(f) (\mathbf{w}^T f)^2 \\ &= \int_{S_D} d\tilde{\gamma}(f) \left[ (\mathbf{w}^T f)^2 - \sum_{l=1}^D f_l w_l^2 \right] \\ &\equiv \int_{S_D} d\tilde{\gamma}(f) \mathbf{w}^T \mathbf{N} \mathbf{w}, \end{aligned} \quad (\text{A.6})$$

where we have defined the symmetric matrix  $\mathbf{N}$  with components  $[\mathbf{N}]_{lm} = f_l f_m - f_l \delta_{lm}$ . Since the measure  $d\tilde{\gamma}(\mathbf{f})$  is non-negative definite, we may establish the lemma by proving that the eigenvalues of  $\mathbf{N}$  are all negative or zero.

So consider the eigenvector equation  $\mathbf{Nr} = \alpha r$ , which is to say

$$\sum_{m=1}^D f_l f_m r_m - f_l r_l = \alpha r_l. \quad (\text{A.7})$$

Since  $\mathbf{f} \in S_D$ , it follows that  $r = \mathbf{1}$  is an eigenvector with eigenvalue  $\alpha = 0$ . One also can easily show that  $\mathbf{f}$  is not an eigenvector of  $\mathbf{N}$ . We may therefore choose the normalization  $\mathbf{f}^T \mathbf{r} = 1$  for  $\mathbf{r}$ . Equation (A.7) then becomes

$$f_l - f_l r_l = \alpha r_l, \quad (\text{A.8})$$

from which we obtain

$$r_l = \frac{f_l}{f_l + \alpha}. \quad (\text{A.9})$$

Imposing the normalization condition  $\mathbf{f}^T \mathbf{r} = 1$ , we obtain the eigenvalue equation

$$1 = \sum_{l=1}^D \frac{f_l^2}{f_l + \alpha} = \sum_{l=1}^D f_l \frac{f_l}{f_l + \alpha}. \quad (\text{A.10})$$

If  $\alpha > 0$ , however, we have that  $f_l/[f_l + \alpha] < 1$  for all  $l$ , so

$$\sum_{l=1}^D f_l \frac{f_l}{f_l + \alpha} < \sum_{l=1}^D f_l = 1. \quad (\text{A.11})$$

This would contradict Equation (A.10). It follows that the eigenvalues of  $\mathbf{N}$  satisfy  $\alpha \leq 0$ . Therefore  $\delta O \leq 0$  in Equation (A.6), and the lemma is proven.  $\square$

A corollary of the lemma that we do not require but that is nonetheless interesting is that the covariance matrix  $\mathbf{C} \equiv \left( \int_{S_D} d\eta(\mathbf{f}) \mathbf{f} \mathbf{f}^T \right) - \mathbf{h} \mathbf{h}^T$  has variances  $\mathbf{w}^T \mathbf{C} \mathbf{w}$  that are maximal when  $d\eta(\mathbf{f}) = d\zeta_h(\mathbf{f})$ .

We may now use the lemma to demonstrate the required statement about the eigenvalues of the stability matrix of the BMS equations.

**Theorem 2.** Suppose that  $\mathbf{A}$  is a positive-definite diagonal  $D$ -dimensional matrix  $\text{Diag}(a_1, \dots, a_D)$ , and that  $\tilde{s}(\mathbf{f})d\mu(\mathbf{f})$  is a probability measure on  $S_D$  with first moment  $\int_{S_D} d\mu(\mathbf{f}) \tilde{s}(\mathbf{f}) \mathbf{f} = \mathbf{f}_s$ . Then:

(i) The eigenvalues  $\beta$  of the stability matrix  $\mathbf{M} \equiv \mathbf{A} \int_{S_D} d\mu(\mathbf{f}) \tilde{s}(\mathbf{f}) \mathbf{f} \mathbf{f}^T$  are real and satisfy  $0 \leq \beta \leq \bar{\beta}$ , where

$$\bar{\beta} \equiv \max_{1 \leq l \leq D} (f_{sl} a_l), \quad (\text{A.12})$$

and where  $f_{sl}$  is the  $l$ -th component of  $\mathbf{f}_s$ .

(ii)  $\bar{\beta} \geq R_0$ .

(iii) If  $\mathbf{A}$ ,  $\tilde{s}(\mathbf{f})d\mu(\mathbf{f})$ , and  $R_0$  are related as in Equation (3.11), so that  $f_{sl} = R_0/a_l$ , then  $0 \leq \beta \leq R_0$ .

To prove reality and non-negativity, we observe that if  $\mathbf{Mx} = \beta \mathbf{x}$ , then

$$\beta = \frac{\int_{S_D} d\mu(\mathbf{f}) \tilde{s}(\mathbf{f}) (\mathbf{x}^T \mathbf{f})^2}{\mathbf{x}^T \mathbf{A}^{-1} \mathbf{x}}. \quad (\text{A.13})$$

This expression is real and non-negative, because of the positive-definiteness of  $\mathbf{A}$ .

Now set  $\mathbf{y} \equiv \mathbf{A}^{-1/2}\mathbf{x}$  and  $\mathbf{P} \equiv \int_{S_D} d\mu(\mathbf{f}) \tilde{s}(\mathbf{f}) \mathbf{f} \mathbf{f}^T$ . Then we have  $\beta = g(\mathbf{y})$ , where

$$g(\mathbf{y}) \equiv \frac{\mathbf{y}^T \mathbf{A}^{1/2} \mathbf{P} \mathbf{A}^{1/2} \mathbf{y}}{\mathbf{y}^T \mathbf{y}}. \quad (\text{A.14})$$

By the Courant-Fisher minimax theorem [17, p.441] the function  $g(\mathbf{y})$  is maximal when  $\mathbf{y}$  is the eigenvector of  $\mathbf{A}^{1/2} \mathbf{P} \mathbf{A}^{1/2}$  corresponding to the largest eigenvalue. But  $\mathbf{A}^{1/2} \mathbf{P} \mathbf{A}^{1/2} = \mathbf{A}^{-1/2} \mathbf{M} \mathbf{A}^{1/2}$ , so that  $\mathbf{M}$  and  $\mathbf{A}^{1/2} \mathbf{P} \mathbf{A}^{1/2}$  are related by a similarity transformation.  $\mathbf{M}$  and  $\mathbf{A}^{1/2} \mathbf{P} \mathbf{A}^{1/2}$  therefore have the same eigenvalues, and their respective eigenvectors  $\mathbf{x}$  and  $\mathbf{y}$  are related by  $\mathbf{A}^{-1/2} \mathbf{x} = \mathbf{y}$ .

We therefore have that if  $\beta_{max}$  is the maximum eigenvalue of  $\mathbf{M}$ , then it is also the maximum eigenvalue of  $\mathbf{A}^{1/2} \mathbf{P} \mathbf{A}^{1/2}$ , so that

$$\begin{aligned} \beta_{max} &= \max_{\mathbf{y} \in \mathbb{R}^D} \left\{ \frac{\mathbf{y}^T \mathbf{A}^{1/2} \left( \int_{S_D} d\mu(\mathbf{f}) \tilde{s}(\mathbf{f}) \mathbf{f} \mathbf{f}^T \right) \mathbf{A}^{1/2} \mathbf{y}}{\mathbf{y}^T \mathbf{y}} \right\} \\ &\leq \max_{\mathbf{y} \in \mathbb{R}^D} \left\{ \max_{d\eta(\mathbf{f}) \in \mathcal{M}(f_s)} \left[ \frac{(\mathbf{A}^{1/2} \mathbf{y})^T \left( \int_{S_D} d\eta(\mathbf{f}) \mathbf{f} \mathbf{f}^T \right) (\mathbf{A}^{1/2} \mathbf{y})}{\mathbf{y}^T \mathbf{y}} \right] \right\} \\ &= \max_{\mathbf{y} \in \mathbb{R}^D} \left\{ \frac{(\mathbf{A}^{1/2} \mathbf{y})^T \left( \int_{S_D} d\zeta_{f_s}(\mathbf{f}) \mathbf{f} \mathbf{f}^T \right) (\mathbf{A}^{1/2} \mathbf{y})}{\mathbf{y}^T \mathbf{y}} \right\}, \end{aligned} \quad (\text{A.15})$$

where we have used Lemma 1 to obtain the last line and  $d\zeta_{f_s}(\mathbf{f})$  is the singular measure in Equation (A.3). We have

$$\begin{aligned} \int_{S_D} d\zeta_{f_s}(\mathbf{f}) \mathbf{f} \mathbf{f}^T &= \sum_{l=1}^D f_{sl} \mathbf{e}_l \mathbf{e}_l^T \\ &\equiv \mathbf{H}, \end{aligned} \quad (\text{A.16})$$

where  $\mathbf{H} = \text{Diag}(f_{s1}, \dots, f_{sD})$ . Combining Equations (A.15) and (A.16), we obtain

$$\begin{aligned} \beta_{max} &\leq \max_{\mathbf{y} \in \mathbb{R}^D} \left\{ \frac{\mathbf{y}^T \mathbf{A}^{1/2} \mathbf{H} \mathbf{A}^{1/2} \mathbf{y}}{\mathbf{y}^T \mathbf{y}} \right\} \\ &= \max_{1 \leq l \leq d} (f_{sl} a_l) = \bar{\beta}, \end{aligned} \quad (\text{A.17})$$

since  $\mathbf{A}$  and  $\mathbf{H}$  are both diagonal matrices. This completes the proof of (i).

To prove (ii), we note that  $\bar{\beta}$  is minimized when all terms  $f_{sl} a_l$  are equal to each other. Indeed, suppose that  $f_{sm} a_m$  is larger than the remaining  $f_{sl} a_l$ ,  $l \neq m$ . Then we may decrease  $\bar{\beta}$  by decreasing  $f_{sm}$  while increasing other components  $f_{sl}$  to maintain the simplex condition  $\mathbf{1}^T \mathbf{f}_s = 1$ . On the other hand, if all the  $f_{sl} a_l$  are equal, then any perturbation of  $\mathbf{f}_s$  preserving the simplex constraint will necessarily increase at least one of the  $f_{sl} a_l$ , resulting in an increase in  $\bar{\beta}$ .

Hence the minimum of  $\bar{\beta}$  occurs when  $f_{sl} a_l = c$ , so that  $f_{sl} = c/a_l$ . By the simplex constraint,  $c^{-1} = \sum_{l=1}^D a_l^{-1} = R_0^{-1}$ . It follows that  $\bar{\beta} \geq R_0$ .

To prove (iii), observe that when Equation (3.11) holds, we have  $f_{sl} = R_0/a_l$ , and

$$\bar{\beta} = R_0. \quad (\text{A.18})$$

□

## Appendix B. Global Stability of Infection-Free Solutions

We now prove a result on the global stability of the infection-free solutions of the BMS equations.

**Theorem 3.** Any solution  $s(\tau, \mathbf{f})$ ,  $i(\tau, \mathbf{f})$  of the BMS equations with a total susceptible fraction  $S(\tau_0) \equiv \int_{S_D} d\mu(\mathbf{f}) s(\tau_0, \mathbf{f})$  satisfying the herd immunity condition at time  $\tau_0$ ,

$$\bar{\beta}(\tau_0)S(\tau_0) \leq 1, \quad (\text{B.1})$$

where  $\bar{\beta}(\tau)$  is given by Equation (4.14), necessarily converges to an infection-free solution as  $\tau \rightarrow \infty$ .

The evolution of the total infected fraction  $I(\tau)$  is given by Equation (2.6),

$$\frac{dI}{d\tau} = -I(\tau) + R_E(\tau)S(\tau)I(\tau), \quad (\text{B.2})$$

with  $R_E(\tau)$  given by

$$R_E(\tau) = \mathbf{f}_s^T \mathbf{A} \mathbf{f}_i. \quad (\text{B.3})$$

The distribution moments  $\mathbf{f}_s$ ,  $\mathbf{f}_i$  are given in Equations (2.7-2.8).

Assume that  $S(\tau_0)$  satisfies the condition of Equation (B.1). The quantity  $S(\tau)\bar{\beta}(\tau)$  may be expanded by combining Equations (4.14) and (2.7):

$$S(\tau)\bar{\beta}(\tau) = \max_{1 \leq l \leq D} \left( a_l \int_{S_D} d\mu(\mathbf{f}) s(\tau, \mathbf{f}) f_l \right). \quad (\text{B.4})$$

Differentiating each term in the argument of the max function on the right-hand side of this equation with respect to  $\tau$ , we may observe that

$$\frac{d}{d\tau} \left( a_l \int_{S_D} d\mu(\mathbf{f}) s(\tau, \mathbf{f}) f_l \right) = a_l \int_{S_D} d\mu(\mathbf{f}) \frac{\partial s}{\partial \tau}(\tau, \mathbf{f}) f_l < 0, \quad (\text{B.5})$$

since by Equation (2.2) the integrand term  $\frac{\partial s}{\partial \tau}$  is strictly negative for non-zero  $s(\tau, \mathbf{f})$  and  $i(\tau, \mathbf{f})$ . It follows that all the terms in the argument of the max function in the right-hand side of Equation (B.4) decline with time, and that after a finite time  $\Delta\tau$  there will exist a finite positive constant  $c$  such that

$$S(\tau)\bar{\beta}(\tau) < (1+c)^{-1} \quad (\text{B.6})$$

for  $\tau > \tau_0 + \Delta\tau$ .

Combining this with Equations (B.6) and (B.2) we find that for  $\tau > \tau_0 + \Delta\tau$ ,

$$\frac{dI}{d\tau} < \left( \frac{R_E(\tau)}{\bar{\beta}(\tau)(1+c)} - 1 \right) I(\tau). \quad (\text{B.7})$$

Substituting for  $R_E(\tau)$  from Equation (B.3) and for  $\bar{\beta}(\tau)$  from Equation (4.11) this becomes

$$\begin{aligned} \frac{dI}{d\tau} &< \left( \frac{\sum_{l=1}^D a_l f_{rl}(\tau) f_{il}(\tau)}{(1+c) \max_{1 \leq l \leq D} (a_l f_{rl}(\tau))} - 1 \right) I(\tau) \\ &\leq \left( \frac{\sum_{l=1}^D f_{il}(\tau)}{1+c} - 1 \right) I(\tau) \\ &= -\frac{c}{1+c} I(\tau). \end{aligned} \quad (\text{B.8})$$

Integrating this inequality with respect to  $\tau$ , we obtain

$$I(\tau) < I(\tau_0 + \Delta\tau) \exp \left( -\frac{c}{1+c} \tau \right), \quad (\text{B.9})$$

which vanishes in the limit  $\tau \rightarrow 0$ . Hence in the asymptotic limit such solutions are always infection-free.  $\square$

## References

- [1] A. Lee, Wuhan novel coronavirus (covid-19): why global control is challenging?, *Public health* 179 (2020) A1.
- [2] R. Bouffanais, S. S. Lim, Cities – try to predict superspreading hotspots for covid-19 (2020).
- [3] M. P. Kain, M. L. Childs, A. D. Becker, E. A. Mordecai, *Chopping the tail: how preventing superspreading can help to maintain covid-19 control*, medRxiv [arXiv:https://www.medrxiv.org/content/early/2020/07/03/2020.06.30.20143115.full.pdf](https://www.medrxiv.org/content/early/2020/07/03/2020.06.30.20143115.full.pdf), doi:[10.1101/2020.06.30.20143115](https://doi.org/10.1101/2020.06.30.20143115).  
URL <https://www.medrxiv.org/content/early/2020/07/03/2020.06.30.20143115>
- [4] J. O. Lloyd-Smith, S. J. Schreiber, P. E. Kopp, W. M. Getz, Superspreading and the effect of individual variation on disease emergence, *Nature* 438 (7066) (2005) 355–359.
- [5] A. G. McKendrick, Applications of mathematics to medical problems, *Proceedings of the Edinburgh Mathematical Society* 44 (1925) 98–130.
- [6] M. Martcheva, *An introduction to mathematical epidemiology*, Vol. 61, Springer, 2015.
- [7] S. Ruan, D. Xiao, J. C. Beier, On the delayed ross–macdonald model for malaria transmission, *Bulletin of mathematical biology* 70 (4) (2008) 1098–1114.
- [8] V. Isham, Stochastic models for epidemics with special reference to aids, *The Annals of Applied Probability* (1993) 1–27.
- [9] J. Medlock, M. Kot, Spreading disease: integro-differential equations old and new, *Mathematical Biosciences* 184 (2) (2003) 201–222.
- [10] W. O. Kermack, A. G. McKendrick, A contribution to the mathematical theory of epidemics, *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character* 115 (772) (1927) 700–721.
- [11] A. James, J. W. Pitchford, M. J. Plank, An event-based model of superspreading in epidemics, *Proceedings of the Royal Society B: Biological Sciences* 274 (1610) (2007) 741–747.
- [12] T. Garske, C. Rhodes, The effect of superspreading on epidemic outbreak size distributions, *Journal of Theoretical Biology* 253 (2) (2008) 228–237.
- [13] J. Aitchison, *The Statistical Analysis of Compositional Data*, Monographs on statistics and applied probability, Blackburn Press, 2003.  
URL [https://books.google.com/books?id=b\\_2YQgAACAAJ](https://books.google.com/books?id=b_2YQgAACAAJ)
- [14] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. Jarrod Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. Carey, İ. Polat, Y. Feng, E. W. Moore, J. Vand erPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, S. . . Contributors, SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, *Nature Methods* 17 (2020) 261–272. doi:<https://doi.org/10.1038/s41592-019-0686-2>.
- [15] M. Harper, et al., python-ternary: Ternary plots in python, Zenodo 10.5281/zenodo.594435 doi:[10.5281/zenodo.594435](https://doi.org/10.5281/zenodo.594435).  
URL <https://github.com/marcharper/python-ternary>
- [16] J. D. Hunter, Matplotlib: A 2d graphics environment, *Computing in Science Engineering* 9 (3) (2007) 90–95.
- [17] G. H. Golub, C. F. van Loan, *Matrix Computations*, 4th Edition, JHU Press, 2013.  
URL <http://www.cs.cornell.edu/cv/GVL4/golubandvanloan.htm>

### Government License

The submitted manuscript has been created by UChicago Argonne, LLC, Operator of Argonne National Laboratory (“Argonne”). Argonne, a U.S. Department of Energy Office of Science laboratory, is operated under Contract No. DE-AC02-06CH11357. The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan. <http://energy.gov/downloads/doe-public-access-plan>.