The Study of Plant Disease Epidemics
Laurence V. Madden, Gareth Hughes, and Frank van den Bosch

ISBN: 978-0-89054-505-8

8

Spatial Aspects of Epidemics—II: A Theory of Spatio-Temporal Disease Dynamics

Although this may seem a paradox, all exact science is dominated by the idea of approximation.

Bertrand Russell

8.1 Introduction

The previous chapter dealt with the basic concepts of dispersal and disease gradients and how gradients change with time. Some simple ways of characterizing the simultaneous expansion of disease intensity in space and increase in time were discussed based on descriptive models. In this chapter, we study the spatio-temporal dynamics of plant disease epidemics using a more mechanistic modeling approach.

The chapter starts with two examples of the expansion of plant diseases, and based on these examples, some key questions will be posed. These questions will be answered using a model for the spatial expansion of plant diseases. Just as we expanded the simple logistic model in Chapter 4 to the H-L-I-R or similar models in Chapter 5, we here expand on the models in Chapter 5 for temporal increase and Chapter 7 for disease gradients to obtain a more satisfactory understanding of spatial expansion of epidemics. Through the developed models, we expand on some of the finer points of the previous chapters to develop a better overall understanding of epidemics. Finally, we then show an example of how to apply the derived models to evaluate a disease control strategy.

8.1.1 Large scale spread: the case of potato late blight

Plant diseases newly emerging in an area can have devastating consequences for the crop and the human population (Schumann, 1991). Potato late blight, caused by *Phytophthora infestans,* is one of the best documented examples. Information on the epidemiology of this disease has been used in several parts of this book. Here we provide a synopsis on the emergence and long-range spread of this disease based on published material (Bourke, 1964; Stevens, 1933; Schumann, 1991).

The disease, previously unknown, was first reported in 1843 in the United States around the large harbors on the east coast. Diseased specimens were possibly brought in by shipping from South America (the Andes region) where potato is an endemic plant species. Between 1843 and 1845 the disease spread in a *wave-like* fashion, from the east coast inwards, at an average rate of 295 km/year (Stevens, 1933). Fig. 8.1, which is redrawn from Bourke's map, shows the spatial expansion of the disease in these years. In 1845, the disease reached the geographical limits of potato production. Yield losses of up to 40% were reported for many places in the country. The social consequences were, however, limited due to the fact that several other crops were grown in sufficient quantities. Up to the present day, the disease continues to be a constraint to potato production.

In Europe the situation was much different. The disease was first found at the end of June 1845, near the west coast in Belgium (Bourke, 1964), presumably brought in by a ship from America. Within one season, the disease spread in a wave-like fashion all over Europe. Its rate of expansion was estimated by Bourke (1964) to be around 13 km *per day*. Yield losses of up to 60% were reported, and losses continued to be severe during the following years (Bourke, 1964). Fig. 8.2, which is redrawn from Bourke (1964), shows the expansion of the epidemic.

Because of the reliance on a single crop for local food consumption, coupled with a very unstable political situation, the losses were devastating to the country, both economically, politically, and socially. Many farmers in Ireland who could not pay their rents to the English landlords were evicted from their land and homes, often resulting in death due to starvation. From the eight million people in Ireland, approximately one million died due to starvation and malnutrition related diseases. Another 1.5 million people managed to emigrate, most to the United States and Canada (Schumann, 1991).

This continental expansion of an epidemic is just one well documented example of disease spread. Many more such examples are described in the literature and we refer the interested reader to Heesterbeek and Zadoks (1987, and references therein), Zadoks (2001), and Hengveld (1989, and references therein).

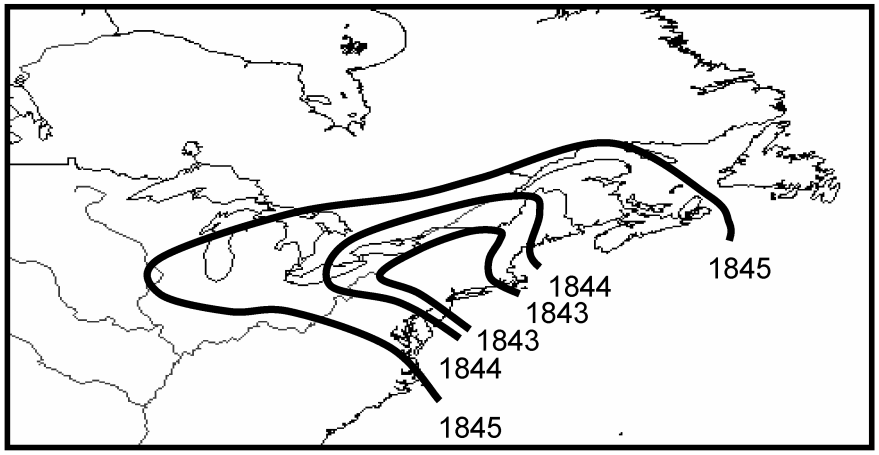

FIG. 8.1. Spatial expansion of potato late blight in the United Sates and Canada. Lines represent boundaries of disease observations. Figure is drawn from information in Bourke (1964).

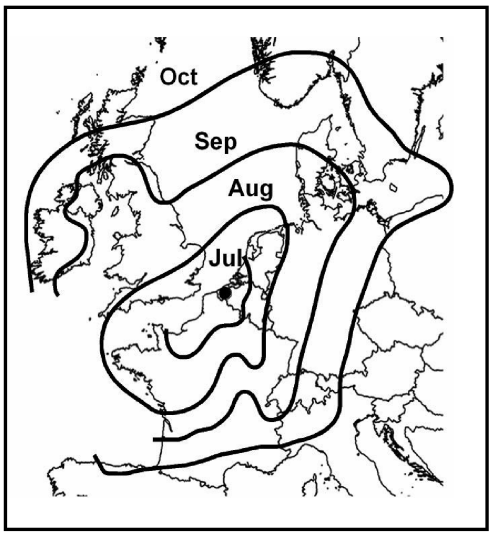

FIG. 8.2. Spatial expansion of potato late blight in Europe. Dot represents first observed infected field. Lines represent boundaries of disease observation. (Redrawn from Heesterbeek and Zadoks, 1987, and based on information in Bourke, 1964.)

8.1.2 Small scale, focus expansion

On a different spatial scale are the disease foci developing and expanding in single fields during epidemics. Such foci originate at the point of infection, where disease intensity subsequently expands radially (but not necessarily symmetrically) due to production of infectious units (e.g., spores), dispersal of the spores (or any infectious units), and new infections caused by some of the dispersed spores. Disease foci of this type were first described by Duhamel in 1728 (Zadoks, 1981)! Without knowing it at that time, he was dealing with an infectious disease of saffron crocus caused by the soil-borne fungus, *Rhizoctonia violacea* (telemorph: *Helicobasidium* sp.). His work led him to recommend the containment of such foci by means of a circular trench because he realized that the disease was expanding from the center outward. Duhamel was far ahead of his time with his research and it took centuries before the research on focal disease expansion was taken further.

One of the best studied examples of focal disease expansion in crop fields is stripe rust disease of wheat, caused by the fungal pathogen *Puccinia striiformis*. Research on foci of this plant pathogen system was done by Zadoks and co-workers (Zadoks and Schein, 1979; van den Bosch et al., 1988c, 1990b), as well as by Emge and Shrum (1976). Early work showed that the rate of focus expansion was as low as 8 cm/day (Zadoks and Schein, 1979). More recent research, based on large-size field experiments, shows that the rate of expansion for stripe rust actually increases in time under many circumstances (Cowger et al., 2005; Sackett and Mundt, 2005a, b) (Fig. 8.3).

Many other diseases are known to spread through crop fields as expanding foci. Rates of expansion vary from just a few centimeters per day, such as for downy mildew on spinach, caused by *Peronospora farinosa* (van den Bosch et al., 1988c), to close to a meter per day for rust on oats caused by *Puccinia coronata* (Berger and Luke, 1979). In Chapter 7 we covered the basic ways of quantifying the focus expansion through the use of observed disease (dispersal) gradients at multiple times, as well as with the use of simple descriptive spatio-temporal models for epidemics. The reader should review the material in section 7.6 before proceeding here. Those with more interest in the subject can read Zadoks and Schein (1979), Zadoks and van den Bosch (1994), Heesterbeek and Zadoks (1987), Jeger (1989), and references therein.

8.1.3 Common features of spatial disease expansion

Although the examples given above discuss spatial expansion of disease on widely different spatial scales, these epidemics have some basic features in common. First of all, the epidemics, whether expanding on a continental scale

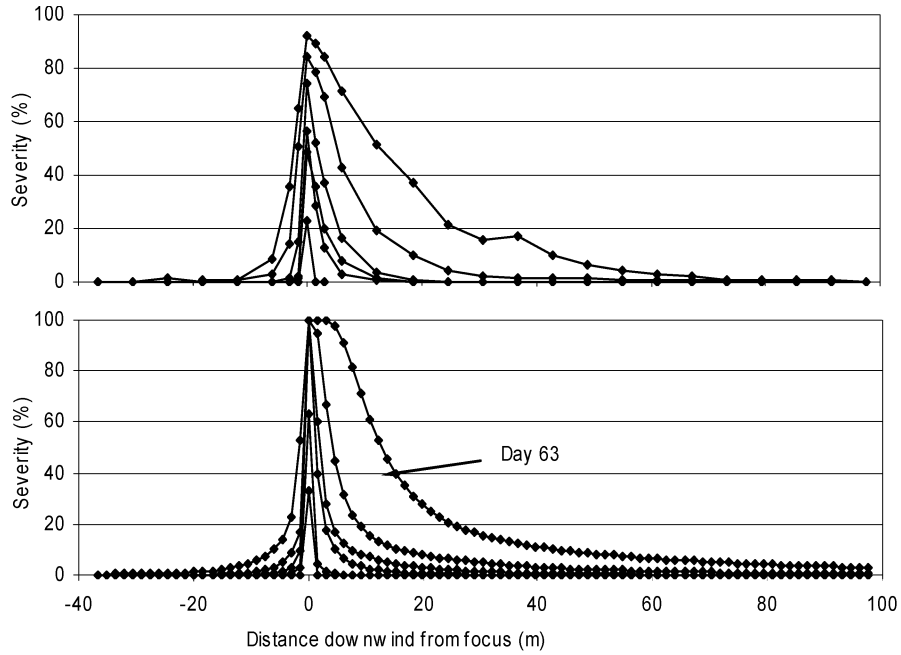

FIG. 8.3. Gradients of the severity of stripe rust of wheat at multiple times for (**A**) observed epidemics and (**B**) computer simulations. Differences in curves for different assessment times at a fixed value of severity increase with time, typical for increasing velocity of expansion. (Reprinted from Sackett and Mundt, 2005b.)

or on a field scale, showed this common wave-like expansion characteristic of many biological invasions. Often the rate of expansion of such epidemic waves is constant and can thus be expressed as single rate value (change in distance divided by change in time). In other cases (Ferrandino, 1993; Aylor, 1999; Frantzen and van den Bosch, 2000), it seems as if the wave increases its rate of expansion continuously, till it runs into the boundaries, either physical or climatological, of the area where disease can occur. Fig. 7.16 demonstrates both of these scenarios.

Underlying epidemic expansion both on a continental scale and on a field scale, and all scales in between, are two basic epidemiological processes we have studied in previous chapters. In Chapters 4, 5, and 6 we studied how epidemics develop in time. In particular, we have studied how epidemic increase depends on mean latent period ($\wp$) and the basic reproduction number (R_0) of the disease, the latter being determined by the per-capita transmission rate (βH_0) and the mean infectious period ($\iota = 1/\mu$) (see equation 5.33). As will be shown, *reproduction* [in the general sense, including transmission (from production of infectious units and probability of infectious units causing infection)] is one underlying processes of epidemic expansion. The disease can spread in space (i.e., occupy new area in the crop, region, and so on) because of the production of daughter infections over time. In Chapter 7 we discussed the dispersal of infectious units, such as spores, around sources of infected individuals, and the description of the resulting disease gradients with various nonlinear models. If all infections are originally at one location, then the disease gradient represents the new infections *beyond* the mother infection(s), an indicator of spread. Thus, in addition to reproduction, dispersal is the other underlying process of epidemic expansion. The greater the reproduction and the greater the distance traveled by infectious units, the greater the epidemic expansion. The basic question is how the combination of reproduction (multiplication) and dispersal of the pathogen around infectious individuals adds up to the development of spatially expanding disease epidemics over time. There are several other related questions.

- How specifically does the rate of epidemic expansion (in, e.g., meters per day) depend on R_0?
- How does the rate of expansion depend of the particular form of the contact distribution, as discussed in section 7.3.4?
- Under what circumstances does the epidemic expand with a constant rate and when does the rate of expansion increase (or change) continuously?

All such questions are difficult, if possible at all, to investigate with just observations of disease at various times and locations. We will, therefore, in the next section develop a model to describe the spatial expansion of disease epidemics based on the underlying mechanisms of reproduction and dispersal. Our understanding of the underlying mechanisms is based on observations and analysis of disease increase in time and dispersal of the pathogen (as presented in Chapters 4, 5, and 7). Analyzing several variants of the model we will be able to theoretically answer the questions posed above. This understanding of spread can then guide us in how we monitor and quantify actual epidemics in time and space.

Our concern in this chapter is mostly on polycyclic diseases, where there is spread from plant to plant, or individual to individual. However, in one section we discuss some of the key issues involved in the spread of monocyclic diseases.

In the next section we utilize the terminology developed in Chapter 5, section 5.1. We suggest that readers review that material before proceeding further.

8.2 Models for Spatial Population Expansion

8.2.1 Introduction

In this section we will study the rate of epidemic expansion using a model. As in Chapter 5, we describe a model that captures the bare essentials behind spatial and temporal dynamics of the epidemic. We assume here that the epidemic does not run into so-called *density dependence* in terms of the healthy host. This amounts to saying that the density of healthy crop at any place and any time can be approximated by the density the crop had at the start of the epidemic (i.e., density of disease-free individuals, H_0). The model thus describes a *density independent* situation. Although this assumption may appear too unrealistic, it turns out, as will be discussed later, that this assumption is fully warranted for assessing expansion rates. This is because the density of infected individuals is generally low (or H_0 is high) at the forefront or leading edge of the epidemic wave. This can be seen for several observed disease gradients in Fig. 7.1, where Y (or $y = Y/H_0$) is small at large distances from the inoculum source.

The assumption about density independence implies that each infected individual (such as a plant, leaf, or unit area of leaf surface) causes, on average, a total of R_0 new infections over its entire infectious period (i.e., life-time). These infections are distributed through space in some fashion. In Chapter 7, we discussed disease gradients and contact distributions and we will use the concept of contact distribution to describe the pattern of the R_0 new infections around an infectious individual. Further simplifying assumptions we use here are: (i) that disease generations are separated in time, as we also assumed for the discrete-time (non-overlapping generation) model at the start of Chapter 5 (section 5.2.1); and (ii) that space is one-dimensional. For practical purposes, the second assumption is consistent with a line source for the initially infected individuals (such as a row of diseased plants). A schematic of a line source is given in Fig. 7.2. In Chapter 7 we assumed that dispersal or spread was from the "edge" of the area of interest, so that distance (s) was exclusively a non-negative number ($s \geq 0$). Here we assume that spread can be to the "left" and "right" of the line source; in other words, we assume that the original and subsequent infections can be in any location in the area. With this assumption, s can be a positive or negative value.

In section 8.3 we will discuss all these assumptions, describe how (or if) they affect the results we will obtain for the simple model, and refer to the relevant literature dealing with these issues. These assumptions can all be relaxed as needed.

8.2.2 Model derivation

To derive the model we show how the density of infectious individuals at position s in disease generation $n + 1$, $Y_{n+1}(s)$, is related to the density of infectious individuals in generation n at *all* possible positions. We use the symbol ξ for any possible (arbitrary) location and s for the specific location of interest. With this notation, $Y_n(\xi)$ is the density of infectious individuals at location ξ at generation n. With a one-dimensional process, $|s - \xi|$ then is the distance between location s and location ξ Note that here we assumed all diseased (infected) individuals are infectious. Thus, $Y = L + I + R = 0 + I + 0 = I$. For ease of comparison with later sections, we express the models in terms of Y (although this is equivalent to I here).

First we calculate the contribution of infectious individuals at location ξ in one generation to the density of infectious individuals in the next generation at distance s. We will use for this the notation $Y_{n+1}(s)\backslash Y_n(\xi)$. Because we are considering the situation in which density dependence does not play a role (e.g., that each spore has no chance to land on an infected individual), each infected individual causes R_0 new infections in the next disease generation; these new infections can be anywhere in the system. The total density of new infections in generation $n + 1$ caused by all infectious individuals present at position ξ in generation n thus is given by $R_0 Y_n(\xi)$. We now have to calculate how many of these infectious individuals will develop at position s. The *contact distribution*, as defined in section 7.3.4, is a scaled version of the dispersal-gradient model that describes how the new infectious individuals in generation $n + 1$ are distributed around the place where the infectious individual was in generation n. Equivalently, the contact distribution, $D(s - \xi)$, multiplied by a scaling constant, is interpreted as the probability of an infectious unit produced at location ξ coming in contact with a host individual at s. (Strictly speaking, $D(s - \xi)$ is the *fraction* per unit area of the *new* infections (i.e., in generation $n + 1$) at position s due to an infectious individual at position ξ). The continuous distribution can also be called a probability density function (*pdf*). Fig. 8.4 represents the process for a hypothetical case of infected individuals at two locations, ξ_1 and ξ_2. The simplest situation would be for the infectious individuals being at one location (e.g., $\xi = 0$) at generation n.

We can thus conclude that the density of infectious individuals in generation $n + 1$ at position s due to all

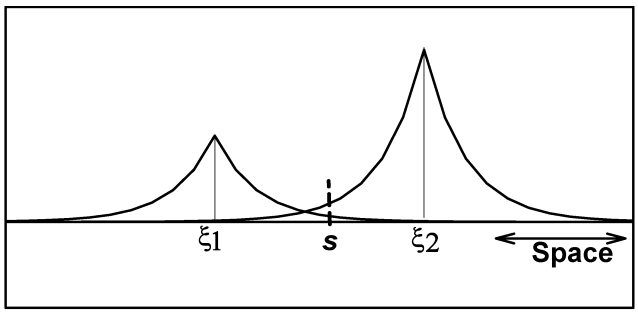

FIG. 8.4. Schematic of spatial spread of new infections around inoculum sources at positions ξ_1 and ξ_2. Number of new infections at position s is the sum of the contributions of new infections from the two sources. See text for further explanation.

infectious individuals in generation n at position ξ is given by:

$$Y_{n+1}(s) \backslash Y_n(\xi) = R_0 Y_n(\xi) D(s - \xi) \tag{8.1}$$

To calculate the total density of infectious individuals at position s in disease generation $n + 1$, we have to add up all contributions to infections at s from *all* individuals in generation n at *all* possible locations ξ. Adding up over a continuous variable, in this case continuous space, amounts to integrating over ξ. The model equation, therefore, takes the final form:

$$Y_{n+1}(s) = R_0 \int_{-\infty}^{\infty} Y_n(\xi) D(s - \xi)\, d\xi \tag{8.2}$$

This is a spatial extension of the purely temporal model for disease increase over discrete generations discussed in Chapter 5 (equation 5.4) for the situation when healthy host density is not limiting. Note that infected individuals at any location, including at location s, can contribute to infections at s. The magnitude of the contribution from each location depends on the distance separating s and ξ, as quantified by $D(s - \xi)$.

As with the discrete-time model for the temporal dynamics of an epidemic derived in Chapter 5 (equations 5.2a and 5.2b), model equation 8.2 is, in principle, easily solved, although the computations are tedious. Substituting the contact distribution and the value of R_0, it is possible to calculate the density of infectious individuals in generation $n = 1$, $Y_1(s)$ if we know the initial condition $Y_0(s)$. For instance, $Y_0(s)$ could be zero at all ξ values except one (say, at $\xi = 0$). The integral might be difficult to work with, but numerical procedures are available in several computer packages. To calculate the second generation, $Y_2(s)$, we substitute $Y_1(\xi)$ again in equation 8.2. This is continued for multiple generations.

The reader is reminded here that this is a density independent variant of a model first studied by Aronson and Weinberger (1978) and Weinberger (1978, 1982). Recently this model was studied by Kot et al. (1996) and versions including more aspects of demography by Neubert and Caswell (2000).

8.2.3 Rates of expansion in relation to contact distributions

In this section we will consider examples of rates of disease expansion based on four different contact distributions. The four distributions are given in Table 8.1. We will discuss each of these expansion-rate results in sequence. Each epidemic is started with Y_0 individuals at position $s = 0$; that is, $Y_0(s) = 0$, except for $Y_0(0) > 0$.

Before discussing spatial dynamics based on the model, we first discuss some properties of the selected contact distributions that are relevant for the interpretation in the next sections. It is convenient to discuss these distributions in terms of spore flight distances, although the contact distribution applies to more than just spore movement (Chapter 7). In the previous chapter, we exclusively considered situations with $s \geq 0$, and mostly dealt with two contact distributions, the exponential (equation 7.8) and Pareto (equation 7.9). Now, we allow s to be positive or negative ($-\infty \leq s \leq +\infty$), to allow for, for example, spore arrival from both "sides" of the location of interest (as shown schematically in Fig. 8.4).

Fig. 8.5 plots realizations of the four contact distributions, with parameter values chosen such that the predicted values had a fair degree of overlap (although total overlap is impossible). Other distributions could also be used, but these four encompass a sufficiently wide range of possibilities. The *double exponential distribution*, sometimes known as the Laplace distribution, is a generalization of the exponential when s can take on any value from positive to negative infinity. Likewise, the *modified power-law distribution* is a generalization of the Pareto distribution, which was used in Chapter 7, for the case of positive and negative s. The *Gaussian distribution* is the classical normal distribution for the random variable s. Finally, the *root distribution* is a nonstandard model for the distribution of s that serves a useful purpose in the comparison of contact distributions, as will become clearer below.

Statistical distributions are often characterized by their first two moments, that is, the mean (expected value) and the variance. However, all distributions do not have finite or definable means and variances, as discussed by Minogue (1989) in one particular context. In Chapter 7 it was pointed out that the Pareto distribution (i.e., for the case with the distance variable being nonnegative) only has a finite mean and variance with suitably large Ψ. The distributions in Table 8.1 are specifically defined here to be symmetrical around distance 0, so they all have means of 0. The variance of the Gaussian, double exponential, and root distributions are σ^2, σ^2, and $\Gamma(6)/\Psi^6$, respectively (where $\Gamma(\bullet)$ is the gamma function). The modified power law distribution

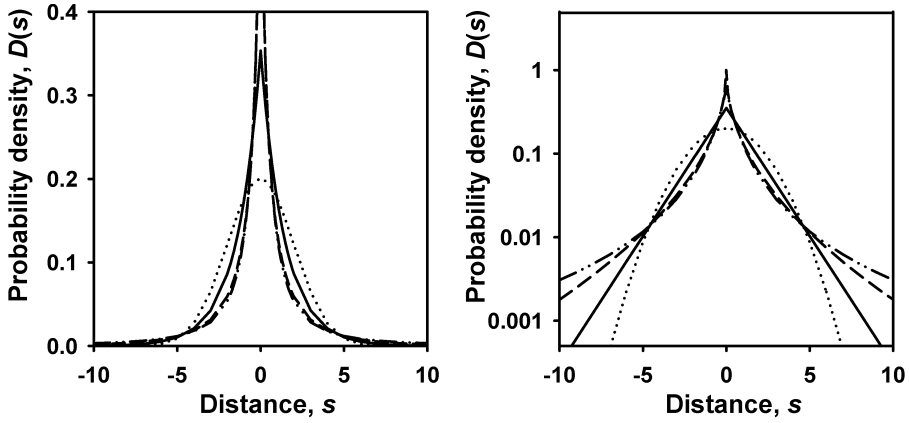

FIG. 8.5. Realizations of four contact distributions, $D(s)$ (or $D(s-\xi)$, with $\xi = 0$). Horizontal axis is distance from source, s, and vertical axis is either the probability density function for the contact distribution listed in Table 8.1 (left-hand graph), or the probability density on a log scale (right-hand graph). Dotted line is for the Gaussian, solid line is for the double exponential, dashed line is for the root, and dash-dotted line is for the modified power-law distribution.

TABLE 8.1. Four contact distributions $[D(s)]$[a] of relevance in the spread of plant diseases.

Common name	$D(s) =$[b]	$ln[D(s)] =$	*Disease expansion rate* $[C_n$ *or* $C(t)]$[c]
Gaussian	$\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{s^2}{2\sigma^2}\right)$	$\ln\left(\frac{1}{\sigma\sqrt{2\pi}}\right)-\frac{s^2}{2\sigma^2}$	Approaches constant (C)
Double exponential	$\frac{1}{\sigma\sqrt{2}}\exp\left(-\frac{\lvert s\rvert\sqrt{2}}{\sigma}\right)$	$\ln\left(\frac{1}{\sigma\sqrt{2}}\right)-\frac{\sqrt{2}\lvert s\rvert}{\sigma}$	Approaches constant (C)
Root	$\frac{\Psi^2}{4}\exp\left(-\Psi\sqrt{\lvert s\rvert}\right)$	$2\ln(\Psi)-\ln(4)-\Psi\sqrt{\lvert s\rvert}$	Increases linearly
Modified power law	$\frac{\Psi}{2}\lambda^{\Psi}(\lambda+\lvert s\rvert)^{-(\Psi+1)}$	$\ln(\Psi/2)+\Psi\ln(\lambda)-(\Psi+1)\ln(\lambda+\lvert s\rvert)$	Increases exponentially

[a]Distributions written with $\xi = 0$. For a more general version of the models, one can replace s by $s-\xi$ throughout. The range for s is negative to positive infinity.
[b]Parameter units: distance for σ in double exponential and Gaussian distributions; distance$^{-1/2}$ for Ψ in the root distribution; unitless for Ψ, and distance for λ in the modified power law distribution. All parameters are positive.
[c]Change in distance per generation or per time period where a low value of Y (Y') is found. See equation 8.6 for general formula.

only has a finite variance at $\Psi > 2$, and its form is so cumbersome as to be intuitively uninformative. For the four distributions here, however, σ, $1/\Psi$ (root), and λ/Ψ (modified power law) have the same general (but not exact) interpretation in terms of contact and dispersal: larger values of σ, $1/\Psi$ (root), and λ/Ψ (modified power law) mean greater distances of spore movement or larger spatial ranges of contact between host individuals.

The σ parameter of the Gaussian and double exponential distributions, the Ψ parameter of the root distribution, and the Ψ and λ parameters of the modified power-law distribution were chosen for Fig. 8.5 so that $D(s)$ values all intersect at the same s value. The vertical axis is on a linear scale (left-hand side) or a logarithmic scale (right-hand side). On a linear scale, the four distributions look quite different near the center (at distances near 0), but appear quite different in the tails (i.e., values relatively far from $s = 0$).

The differences between the tails of the four contact distributions become very clear when using a log scale for the ordinate (Fig. 8.5). The double exponential distribution consists of straight lines. The tail of the Gaussian distribution decreases faster with distance from the origin than the exponential distribution. Both the double exponential and the Gaussian distribution are said to be *exponentially bounded* (see section 7.3.4), meaning that the tails fall off (declines) faster than, or as fast as, an exponential function. In contrast, the root and the modified power-law distributions clearly decrease less than exponentially with distance. In particular, the rate of decline in $\ln[D(s)]$ decreases with increasing distance from the origin (Fig. 8.5). This is *not* dependent on the chosen values for Ψ and/or λ. Distributions with this property are said to be non-exponentially bounded, "fat-tailed", or "heavy tailed".

As discussed in Chapter 7, it was emphasized that there are many examples of dispersal gradients described by the exponential gradient model and many other examples of gradients described by the power-law gradient model. This means there are numerous examples of contact distributions that are either exponentially bounded or not exponentially bounded. In the next sections we show that there is an essential difference between the spatial expansion of the epidemic for an exponentially bounded and for a non-exponentially bounded contact distribution. These mechanistic results add to the more descriptive results given in the previous chapter.

8.2.3.1 Gaussian contact distribution. Model equation 8.2 with the Gaussian contact distribution (Table 8.1) can be solved explicitly (N. L. Johnson et al., 1994, 1995) for disease density at location s in generation n. The solution is:

$$Y_n(s) = Y_0 R_0^n \frac{1}{\sigma\sqrt{2\pi n}} \exp\left(-\frac{s^2}{2\sigma^2 n}\right) \qquad (8.3)$$

Fig. 8.6 shows an example realization of the solution of this model. How will the front of the epidemic wave move through space over time (i.e., between generation)? To calculate this we determine from equation 8.3 the position of a pre-specified density of infected individuals, Y', at time (generation) n, relative to the starting point of the epidemic, $s_{Y',n}$. To simplify the presentation of the equations and other calculations, we simplify the notation somewhat and use s'_n for $s_{Y',n}$. In Fig. 8.6A this concept is illustrated by the identification of Y' on the ordinate and the position of this value of disease for each generation on the abscissa. The line of fixed disease level is known as an *isopath* (see section 7.6.3), so s'_n is an isopath. Calculating the position of the small density of infectious individuals requires solving the equation $Y' = Y_n(s'_n)$. Using Y' and s'_n results in:

$$Y' = Y_n(s'_n) = Y_0 R_0^n \frac{1}{\sigma\sqrt{2\pi n}} \exp\left(-\frac{(s'_n)^2}{2\sigma^2 n}\right) \qquad (8.4)$$

and after some algebraic manipulation we find:

$$s'_n = \sqrt{2\sigma^2 n \left\{\ln(Y_0) - \ln(Y') - \ln\left(\sigma\sqrt{2\pi n}\right) + n\ln(R_0)\right\}} \qquad (8.5)$$

Fig. 8.7 shows s'_n as function of generation, n. The figure shows that after a very short initial phase, the distance of Y' to the starting point of the epidemic increases

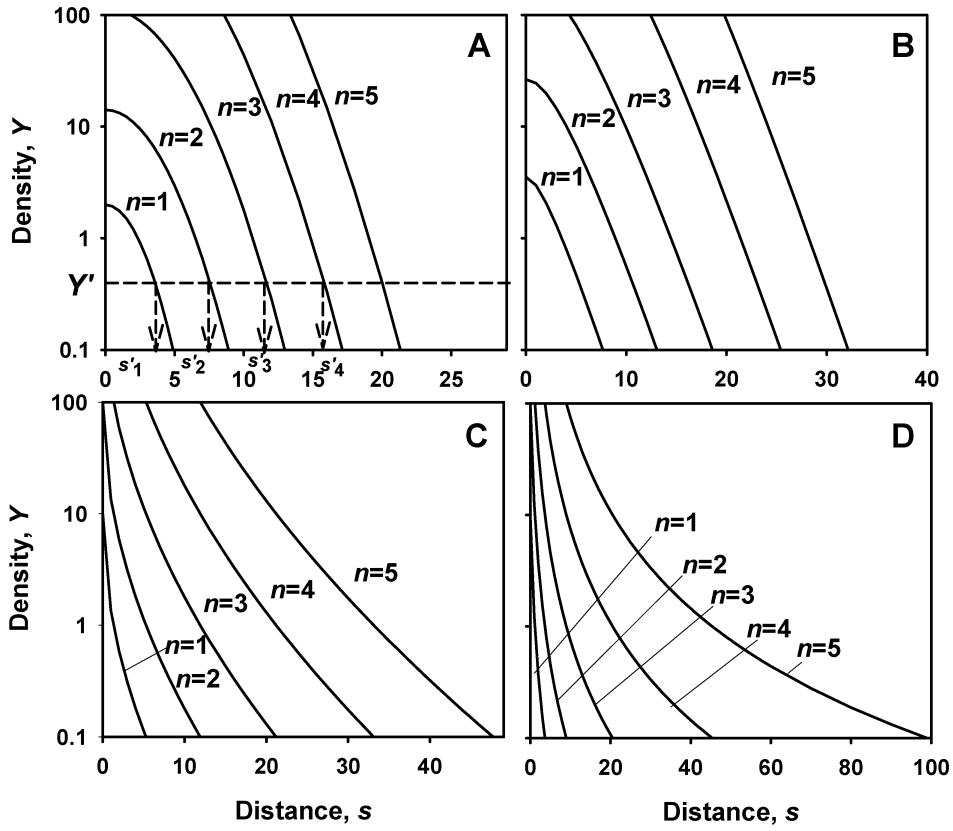

FIG. 8.6. Spatial expansion of disease density, $Y_n(s)$, as described by equation 8.2, corresponding to the four contact distributions depicted in Fig. 8.5 and listed in Table 8.1 It is assumed that the original inoculum source is at distance 0. Vertical axis is density of diseased individuals, and horizontal axis is distance from the starting point of the epidemic. Only positive distances are shown, although spread can be in both directions. (**A**) Gaussian contact distribution; (**B**) double exponential contact distribution; (**C**) root contact distribution; and (**D**) modified power law contact distribution.

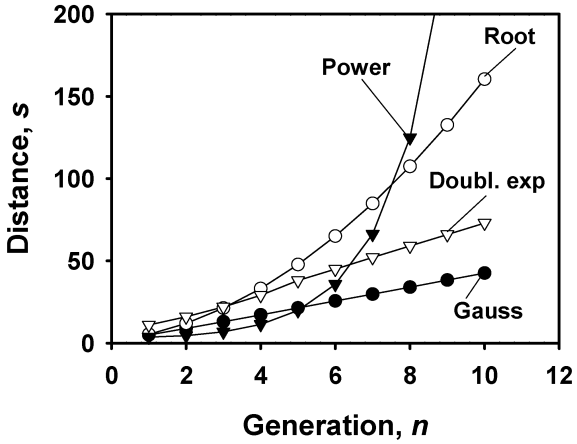

FIG. 8.7. Distance between the starting point of the epidemic (original inoculum source) and the position of a specified low density of disease, Y' (see Fig. 8.6A), for a sequence of non-overlapping generations. Results based on solving equation 8.2 for distance, s'_n ($=s_{Y',n}$), where Y' is found, for each of the contact distributions in Table 8.1.

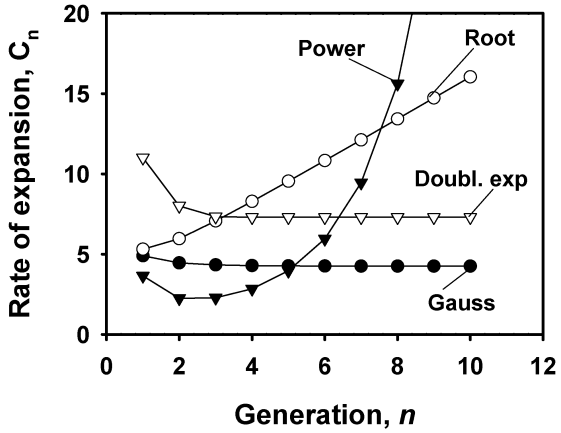

FIG. 8.8. Mean rate (or velocity) of epidemic expansion, C_n, between the beginning of the epidemic (generation 0) and the current generation (n), based on distances shown in Fig. 8.7. Results obtained using equation 8.6 for each of four contact distributions.

linearly with time. This implies that the rate of expansion is a constant. To calculate the rate we will consider the average rate of expansion between the start of the epidemic, $n = 0$, and the present generation n. This average rate of expansion, denoted by C_n, is calculated from:

$$C_n = \frac{s'_n}{n} \tag{8.6}$$

Substituting equation 8.5 in this expression we find as the rate of epidemic expansion:

$$C_n = \frac{s'_n}{n} = \sqrt{\frac{2\sigma^2\left\{\ln(Y_0) - \ln(Y') - \ln\left(\sigma\sqrt{2\pi n}\right)\right\}}{n} + 2\sigma^2\ln(R_0)} \tag{8.7}$$

In Fig. 8.8 the average rate of epidemic expansion is plotted as a function of time, for the same parameters and initial condition as those used in Fig. 8.6. C_n is the rate (or velocity or speed) of isopath movement (as in section 7.6.3.2). We see from this figure that the rate of expansion reaches a constant rate, C, very soon after the start of the epidemic. This result also holds for other parameter values than the ones used to calculate Fig. 8.8. This constant rate of expansion can be calculated mathematically from equation 8.7 by letting n go to infinity (i.e., increase indefinitely). We immediately see from this equation that the whole first term under the square-root sign vanishes. The constant rate of expansion thus is:

$$C = \sigma\sqrt{2\ln(R_0)} \tag{8.8}$$

Fig. 8.9A shows the dependency of C on the parameter values R_0 and σ. C increases linearly with the standard deviation of the Gaussian distribution; that is, increasing σ, which means a shallower gradient, results in a higher C. Moreover, C increases with R_0 meaning that increasing the number of new infections resulting from each infection also results in a faster velocity of isopath movement. For large values of R_0, say $R_0 > 10$, C approximately increases linearly with the logarithm of R_0. The per unit increase in C with R_0 is higher when R_0 is low than when it is high.

Before assessing the model (equation 8.2) with the other contact distributions, we discuss some numerical examples. Suppose we have done a series of experiments and found that the contact distribution of the disease under consideration is Gaussian and has a standard deviation of 10 cm. This value could be obtained, for instance, with a fungal disease dispersed by spores, by introducing infectious plants in one row in a field (for generation 0), and counting new infections at several distances from the origin the equivalent of one generation later (at $n = 1$). As long as $Y_1(s)$ is not high, one can regress $\ln[Y_1(s)]$ on s^2, and the slope is an estimate of $1/(2\sigma^2)$ (see Table 8.1). Furthermore, suppose that an experiment on focus expansion of this disease, as discussed in sections 8.1.2 and 7.6.5, shows an expansion rate at around 12 cm/day over a few generations. This can be done by identifying a particular density of disease (say, $Y' = 1$), determining where this density is found for each generation (s'_n), and then using equation 8.6. What does this tell us about the value of R_0? Solving R_0 from equation 8.8 we find:

$$R_0 = \exp\left(\frac{1}{2}\left(\frac{C}{\sigma}\right)^2\right) \tag{8.9}$$

Substituting $\sigma = 10$ and $C = 12$ we find that $R_0 = 2.05$. Suppose we can choose between two focus expansion

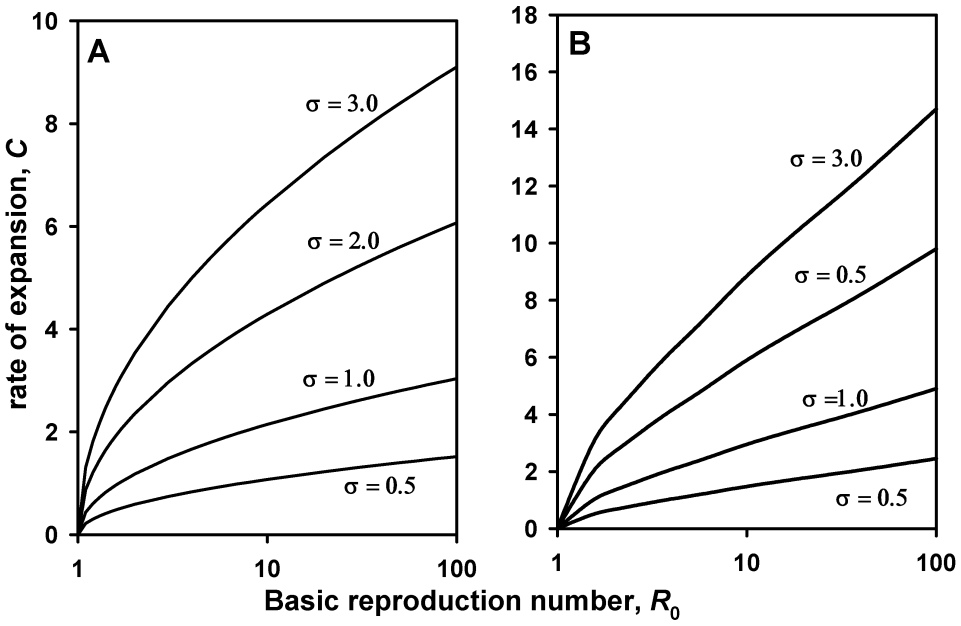

FIG. 8.9. The final (or asymptotic) rate of disease expansion, C, for epidemics with a (**A**) Gaussian contact distribution, and (**B**) double exponential contact distribution, in relation to the basic reproduction number (R_0; log scale) and standard deviation of the contact distribution (σ).

management methods. First, that it is possible to reduce σ by 50%. This can be done, for example, by planting a different crop variety whose plant architecture is such that the crop is more dense, implying less air movement in the crop and consequently infectious units being dispersed over shorter distances. Substituting this into equation 8.8 we find that the rate of disease expansion is reduced to 6 cm/day. We are assuming here that R_0 is not affected by this management tactic. Our second disease management method involves the use of fungicides that reduces R_0 by 50%. Substituting this new R_0 into equation 8.8, we find that the rate of focus expansion is reduced to 2.2 cm/day. For small values of R_0 it is more efficient (at least for this Gaussian contact-distribution example) to restrict disease expansion by reducing the number of new infections from each mother infection than it is to change the length scale of the contact distribution (σ).

Instead of using equation 8.6, one could obtain estimates of expansion velocity for each *successive* pair of generations. For a given Y', we can define the generational rate (or velocity) as:

$$C'_n = s'_{n+1} - s'_n \qquad (8.10)$$

where s'_n and s'_{n+1} are the distances where Y' is found in generations n and $n + 1$, respectively. One could plot these, and if no trend is apparent (as expected here), then an average of the C'_n values could be an estimate of C. Equation 8.10 produces "instantaneous" values of velocity (as also utilized in Chapter 7), whereas equation 8.6 produces an overall average from the first to nth generation. We emphasize the use of equation 8.6 here, since the biological interpretation is the same as that obtained with the mean of equation 8.10.

8.2.3.2 Double exponential contact distribution. Model equation 8.2 with the double exponential contact distribution (Table 8.1) can be solved explicitly (N. L. Johnson et al., 1994, 1995), although the solution is very messy:

$$Y_n(s) = Y_0 R_0^n \frac{1}{\phi} \frac{e^{-|s|/\phi}}{2^{2n-1}(n-1)!} \sum_{j=0}^{n-1} \frac{2^j (2n-j-2)!}{j!(n-j-1)!} \left|\frac{s}{\phi}\right|^j \qquad (8.11)$$

where $\phi = \sigma/\sqrt{2}$. Fig. 8.6B shows a solution of the model for the same parameter values as used for the solution of the model with the Gaussian contact distribution. The figure shows that the solutions for density of the infected individuals are very similar to those using the Gausian density as contact distribution, except that the rate of expansion of the model with the exponential contact distribution seems to be higher.

Unfortunately, equation 8.11 is so complicated that it turns out not to be possible to explicitly derive an equation for s'_n nor for the average rate of expansion C_n, as was done for the model with the Gaussian contact distribution. From equation 8.11, distance s'_n and the rate C_n *can* be calculated numerically (i.e., iteratively), and Figs. 8.7 and 8.8 show the results of this calculation. As for the model with the Gaussian contract distribution, we see from the figure that C_n approaches a constant value (C) just a few disease generations after the start of the epidemic. Without further proof we mention that this is not due to the particular

set of parameter values chosen. For each possible set of parameter values the rate of expansion approaches a constant.

Fig. 8.9B shows C as a function of the logarithm of R_0 for various values of σ of the double exponential distribution. This figure shows that the rate of expansion for the model with the double exponential distribution is larger than for the model with the Gaussian distribution. This is due to the fact that the exponential distribution has thicker (heavier, fatter) tails than the Gaussian distribution (but still exponentially bounded), resulting in more new infections further away from an infected individual. For values of R_0 larger than approximately 2, the rate of expansion is close to linear in the logarithm of R_0.

8.2.3.3 Root contact distribution. When using the root contact distribution in model equation 8.2, it is not possible to solve the model explicitly. Kot et al. (1996) have, however, introduced a very useful method to calculate approximations to the solution. Applying their method we find as approximate solution:

$$Y_n(s) \cong Y_0 R_0^n \frac{\Psi^2}{4} e^{-\Psi\sqrt{|s|}} \tag{8.12}$$

Fig. 8.6C shows a solution of the model. The figure suggests that the rate of epidemic expansion does not apparently converge to a constant. In fact, the rate of expansion keeps increasing over time (i.e., generation) as indicated by increasing distances between low Y values from generation to generation. This is a different qualitative behavior than found with the Gaussian and the double exponential contact distributions.

The average rate of disease expansion, C_n, can be calculated using the same procedure as for the model with the Gaussian contact distribution. This procedure yields:

$$\begin{aligned} C_n &= \frac{s_n'}{n} \\ &= \frac{1}{\Psi^2}\left[\frac{\ln(Y_0) - \ln(Y') + \ln(\Psi^2/4)}{\sqrt{n}} + \sqrt{n}\ln(R_0)\right]^2 \end{aligned} \tag{8.13}$$

For large values of n, this becomes:

$$C_n = n\frac{1}{\Psi^2}(\ln(R_0))^2 \tag{8.14}$$

This last expression shows that the rate of epidemic expansion indeed increases linearly with time (generation), with a slope of $[\ln(R_0)]^2/\Psi^2$. Specifically, the rate of expansion increases per generation with the square of the logarithm of the basic reproduction number and with the square of the distances traveled by spores (related to $1/\Psi$). The plot of s_n' versus n (Fig. 8.7) shows the expected upward curvature, indicating increasing rate of disease expansion. The plot of C_n versus n (Fig. 8.8) confirms the *linear* increase in the expansion rate after a few generations. The implication of equation 8.14 is that the farther the distance from the inoculum source, the greater the expansion rate of the epidemic.

It was pointed out before that, all other things being equal, C_n (or C) is higher with a double exponential than with a Gaussian distribution for contact in the epidemic model. However, there is no simple overall comparison of C_n obtained with a root and the previous contact distributions. This is because the expansion rate increases with time (generation) for the root distribution. Depending on the value of Ψ, C_n may be lower or higher than the corresponding rates for the previous two contact distributions when n is not large. With enough time, however, the expansion rate for a root distribution will (must!) surpass that obtained with the previous two distributions, since C_n is directly proportional to n. With sufficient time, C_n for a root distribution can be considerably larger than that obtainable with the Gaussian or double exponential distribution.

8.2.3.4 Modified power law contact distribution. Using the modified power law contact distribution in equation 8.2, it is not possible to solve the model explicitly. Applying the method of Kot et al. (1996) again we find as approximate solution:

$$Y_n(s) \cong Y_0 R_0^n \frac{\Psi}{2} \lambda^{\Psi} (\lambda + |s|)^{-(\Psi+1)} \tag{8.15}$$

Fig. 8.6D shows a solution of the model for one set of parameters. The figure suggests that, as for the model with the root contact distribution, that the rate of epidemic expansion does not converge to a constant, but that the rate of expansion keeps increasing over time (as demonstrated by the increasing distance between a selected low disease density for successive pairs of generations). In fact, with this example, it appears that the rate of expansion is increasing faster between generations than with the root contact distribution. How does the rate of epidemic expansion, C_n precisely increase with time? To answer this question, C_n was calculated using the same procedure as for the model with the Gaussian contact distribution. This procedure yields:

$$C_n = \frac{s_n'}{n} = \lambda\left(\frac{Y_0\Psi}{2\lambda Y'}\right)^{(1/(\Psi+1))}\left(\frac{R_0^{(n/(\Psi+1))}}{n}\right) - \frac{\lambda}{n} \tag{8.16}$$

For large times (generations), equation 8.16 approaches:

$$C_n = A\frac{R_0^{(n/(\Psi+1))}}{n} = \frac{A}{n}\exp\left(\frac{n}{\Psi+1}\ln(R_0)\right) \tag{8.17}$$

where the constant A is given by:

$$A = \lambda\left(\frac{Y_0\Psi}{2\lambda Y'}\right)^{1/(\Psi+1)} \tag{8.18}$$

From equation 8.17 we see that the rate of epidemic expansion increases as an exponential function of time (n). However, the relationship is a little more complicated, since n also appears in the denominator outside the exponential or power function in equations 8.16 and 8.17. This means that at small n (i.e., when equation 8.16 applies), the expansion rate *could* initially decrease with increasing time (depending on the other parameters), but the exponential function of n will dominate after a few disease generations, giving an increasing C_n.

Figs. 8.7 and 8.8 demonstrate how s'_n and C_n, respectively, are related to time (generation) with a particular combination of R_0 and contact-distribution parameters. The distance where Y' is found (s'_n) is clearly seen to increase over all times as the focus expands (Fig. 8.7), although there is a slight decrease in the rate of disease expansion for the first two generations, followed by an increasing rate (with no upper bound) (Fig. 8.8). Unlike the case for a root contact distribution, the increase in disease expansion rate is more than linear. That is, the change in C_n ultimately becomes larger with every successive pair of generations.

As discussed for the root distribution, there is no simple comparison of C_n between models with the modified power-law contact distribution and with the Gaussian and double exponential distributions. During the early stages of an epidemic, C_n for a modified power-law distribution may be smaller or larger than the C_n for a Gaussian or double exponential distribution (depending on the other parameters). However, with sufficient time, with the modified power-law distribution will be larger, possibly much larger, than the others.

8.2.4 Comparisons

Most plant diseases do not increase in discrete and non-overlapping generations (see Chapters 4 and 5). However, we showed in section 5.2.1 that use of discrete-generation models can readily lead to a general understanding of temporal dynamics of diseases. In fact, key results obtained with such a simple model—thresholds for epidemics, final disease levels, exponential increase early in the epidemic—all carry over to more realistic (but more complicated models) (sections 5.2.2–5.2.5). Here we utilize the simplicity of a discrete generation model to explore the consequences of different forms of the contact distribution (a scaled version of the disease gradient when healthy host density is not limiting) on spatio-temporal dynamics of disease. As required, the increase in disease density from generation to generation is dependent on R_0. Moreover, we found that the expansion rate of a disease focus (C_n or C) is also dependent on R_0 (the descriptor of reproduction, in the general sense) for all the contact distributions considered. The functional relationship *does* depend on the contact distribution, but typically C_n is proportional to $\ln(R_0)$. Since $\ln(R_0)$ is only positive when $R_0 > 1$, it is clear that for C_n to be positive in the long term (e.g., Gaussian distribution), or for the increase in C_n to be positive in the long term (e.g., root and modified power-law distributions), then R_0 must exceed 1.

The results of the previous sections show that there are two very distinct qualitative classes of disease expansion in epidemics. The first one is the class of *traveling epidemic waves*, where it can be said that spread is in a wave-like manner (see section 7.6.3). A traveling wave occurs when the contact distribution is Gaussian, double exponential, or any distribution that has an exponentially bounded tail. Mollison (1972) expressed this as a more formal rule based on the moment generating function of the contact distribution. The traveling waves approach a constant rate of expansion. As far as our understanding goes, this constant rate is approached very quickly, but this does depend on the parameter values. Traveling waves for populations have been described by mathematicians as early as 1937 (Fisher, 1937). The concept has been applied to: ecology for population expansion (Skellam, 1951; Hengveld, 1989; van den Bosch et al., 1992; Kot et al., 1996; Shigesada and Kawasaki, 1997; Neubert and Caswell, 2000); animal epidemiology for disease expansion (Diekmann, 1978, 1979; Thieme, 1977, 1979; Mollison, 1972); and to genetics for the rate of expansion of a gene (Fisher, 1937). Traveling epidemic waves also have been modelled for plant disease epidemic models (Kampmeijer and Zadoks, 1977; Zadoks and Kampeijer, 1977; Minogue and Fry, 1983a, b; van den Bosch et al., 1988a–c, 1990a, b, 1999).

The second class of disease expansion is of *dispersive epidemic waves*. A dispersive wave occurs with the root and modified power-law contact distributions, or with any distribution that does not have an exponentially bounded tail (see also section 7.6.3). Waves of this type continuously keep increasing in their rate of expansion (C_n never approaches a constant). Mechanistic, statistical, and meteorological arguments for this type of disease expansion were provided by Ferrandino (1993), Minogue (1989), and Aylor (1999). Recently, the idea of dispersive waves has also received attention both in plant pathology (Aylor, 1999; Cowger et al., 2005; Frantzen and van den Bosch, 2000) and in ecology (Kot et al., 1996). As our results for the root and modified power-law distributions demonstrate, the rate of expansion of dispersive waves can increase with time in various ways. We do not know of any thorough investigation about the formal relation between the shape of contact distributions and the way the rate of epidemic expansion increases for dispersive waves, and much still has to be discovered about this.

Traveling waves and dispersive ways were introduced in the previous chapter (section 7.6.3) in terms of observed epidemics and using a more descriptive model. We discuss some other features of the spatio-temporal model of Chapter 7 in more detail later, but here we point out an aspect of the spatial component of that model. Equations 7.33 and 7.35 were not based on *explicit* consideration of the contact distribution, but dealt directly with the disease gradient (Y or y in relation to s), with correction for a maximum disease level of 1. Moreover, although the equations of the previous chapter could describe gradients of disease in relation to the original inoculum source at multiple times, they could not adequately represent spread of disease between any two arbitrary locations away from the original inoculum source. Basically, of the possible models that could have been used for the spatial component, in order to provide general agreement with observed gradients at several times in an epidemic, equations were chosen specifically to obtain either a constant expansion rate or one that increased with time. Using a model for spread in this chapter, where contact between host individuals is directly incorporated at the level of each infected and healthy individual, we showed how traveling and dispersive wave can (must) naturally develop.

The more realistic spatial component here also shows that more complicated spread behavior can occur during the very early stages of polycyclic epidemics, before the spread "settles down" into its long term behavior of either a constant expansion rate or accelerating rate. Interestingly, in the simpler modeling of Chapter 7, we showed that the expansion rate was a function of the temporal rate parameter (e.g., r_L for the logistic model). Moreover, in Chapter 5 we showed that r_E (same as r_L when y is low for a logistic epidemic) is related to R_0. Thus, the results of the previous chapter are broadly consistent with the ones here, where C is seen to be a function of $\ln(R_0)$.

8.3 Some Extensions

8.3.1 One-dimensional versus two-dimensional epidemic expansion

Model equation 8.2 describes an epidemic in a one-dimensional space. Epidemics generally develop in two dimensions, and we need to know how the rate of epidemic expansion in one dimension relates to the rate of epidemic expansion in two dimensions. If the epidemic develops from a line source which is long compared to the distance infectious units disperse (see Chapter 7), the density of infectious units (and resulting infected host individuals) at a given distance perpendicular to this line source will be the same (on average) for every position along the line source. For instance, in Fig. 8.10, the density of infected individuals at point A will be the same as the density of infected individuals at point B, and the same will hold for any point that is a perpendicular distance d from the line source. This implies that we are dealing with a one-dimensional spatial epidemic, at least in terms of the model, where the space direction is perpendicular to the line source. Only at the edges of the line source we will see differences. If the line source is long compared to the standard deviation of the contact distribution (where there is a definable standard deviation, as with the Gaussian distribution), then the complications of an edge effect will not be of concern.

Now consider an epidemic developing from a point source, and assume that the two-dimensional space is homogeneous, meaning that all environmental circumstances are the same at every location, and the density of the crop is the same at every point. Under this assumption, the epidemic will be *rotationally symmetric* at every moment in time (at least in the model). Fig. 8.10 shows the epidemic developing from a point source (similar to Fig. 7.17). Contour lines for a fixed density of infected individuals are drawn for various points in time. These contour lines are concentric circles. When these concentric circles are large enough, locally they look like a straight line. That is, in the vicinity of any point on a circle, the circle appears a straight line in the same way as any uninformed observer would consider the world to be a flat disk based on local observations. This implies that such an epidemic can *eventually* be approximated (after enough time so that the circle circumference is not too small) by an epidemic expanding in one dimension, in which the direction is parallel to the circle. The one-dimensional model of this chapter thus applies to epidemics that develop from a point source as well, if we take care to use the proper form of the contact distribution, as will be discussed below.

There is, however, still the question left of what the general relation is between the two-dimensional contact distribution around an infectious individual and the one-dimensional contact distribution perpendicular to a line of infectious individuals. Consider a two-dimensional contact distribution $D(s_1, s_2)$. This distribution gives the scaled probability per unit area of an infectious unit originating at $(0, 0)$ being at (s_1, s_2). Equivalently, $D(s_1, s_2)$ is the fraction per unit area of the new infections at position (s_1, s_2) due to an infectious individual at $(0, 0)$. We could have just as easily taken any other position as the origin (ξ_1, ξ_2); then the contact distribution is written as $D(s_1 - \xi_1, s_2 - \xi_2)$. Let us assume a line source is parallel to the ordinate of our coordinate system. Fig. 8.10 shows an example in the lower right corner. Such a line source can be thought of as composed of a linear series of point sources. Drawn in Fig. 8.10 is a series of contour lines depicting $D(s_1, s_2)$ distributions around arbitrary points on the line. The contact distribution for the line source in *just* the s_1 direction, $D(s_1)$, is the scaled probability of infectious units on a line perpendicular to the line source. The figure

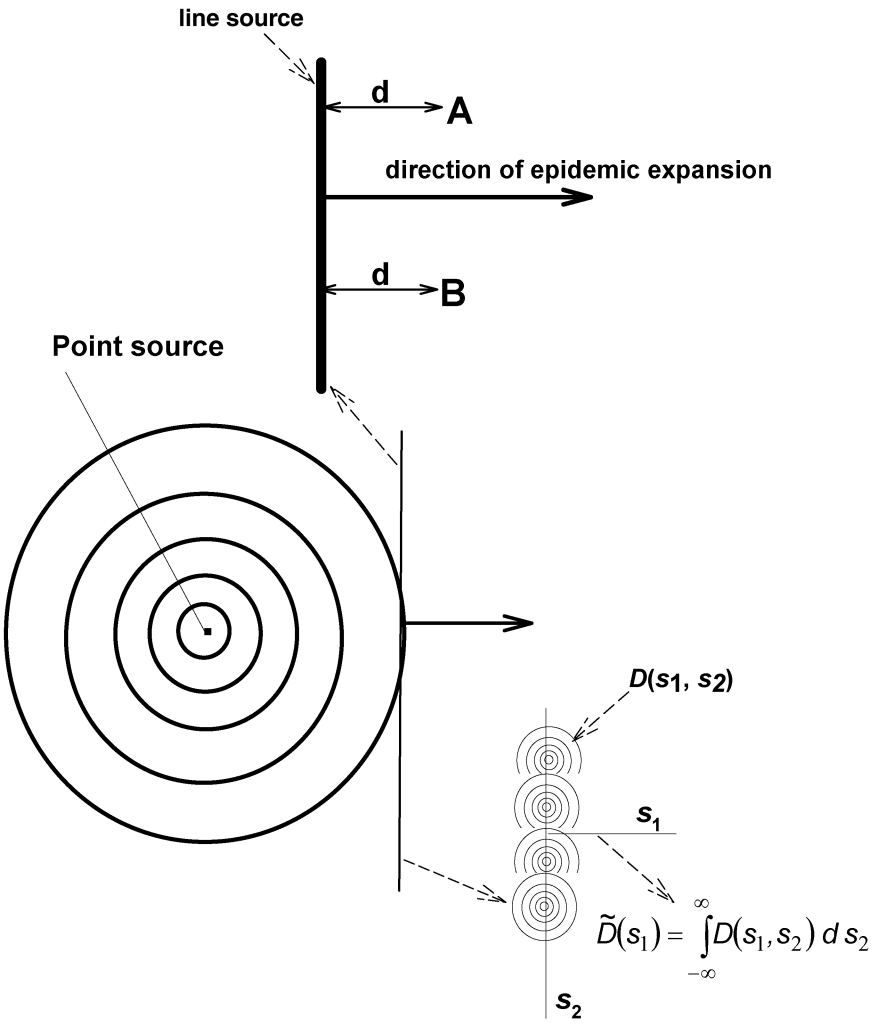

FIG. 8.10. Schematic for the relation between one-dimensional and two-dimensional disease spread in epidemics. See text for explanation.

attempts to make clear that the one-dimensional contact distribution can be calculated by adding up all contributions of all point sources, which is equivalent to the summation over all possible values of contact probability across the second spatial dimension (s_2). Assuming continuous space, we use integration instead of summation and write:

$$\tilde{D}(s_1) = \int_{-\infty}^{\infty} D(s_1, s_2) ds_2 \qquad (8.19)$$

$\tilde{D}(s_1)$ is known in statistics as the *marginal distribution* of the bivariate distribution $D(s_1, s_2)$.

Each contact distribution of Table 8.1 can be considered the marginal distribution of some bivariate distribution. Although one can, in principle, determine the marginal distribution for a given bivariate distribution, the opposite is not necessarily true—one cannot necessarily determine the bivariate distribution for a given marginal distribution. For the Gaussian distribution, the relationship is straightforward—the marginal and bivariate forms are both Gaussian (see Johnson and Kotz, 1970).

Some papers that deal with two-dimensional models for contact distributions include: Aronson and Weinberger (1978) and Weinberger (1978, 1982). We do not further work with explicit two-dimensional contact distributions in this chapter.

8.3.2 Continuous time and more realistic temporal dynamics

8.3.2.1 Model and simulations. Model equation 8.2 can be extended to represent epidemics with: overlapping disease generations (i.e., with continuous time); latent and infectious periods; dependence of the rates on healthy host density (i.e., for the density-dependent situation, when $H(t)$ is not considered fixed at H_0); and healthy host-density growth in the absence of disease. These extensions are similar to the way our simplest epidemic progress model (equations 5.2a and 5.2b) was extended in Chapter 5. The H-I-R, H-L-I-R, Vanderplank differential-delay, and the Kermack-McKendrick models can all be extended to allow for an explicit spatial component (Diekmann, 1978, 1979; Thieme, 1977, 1979).

For brevity, we consider only a spatial extension of the H-L-I-R model discussed in section 5.2.3. The densities of healthy, latent, infectious and removed individuals are now a function of both time, t, and position (location) in space, s, $H(t, s)$, $L(t, s)$, $I(t, s)$, and $R(t, s)$, respectively. We now assume that t and s are both continuous variables. The model is given in pseudo-equation form as:

$$\begin{aligned}
\frac{\partial H(t,s)}{\partial t} &= -\begin{bmatrix}\text{rate at which healthy}\\ \text{individuals become infected}\end{bmatrix}\\
\frac{\partial L(t,s)}{\partial t} &= \begin{bmatrix}\text{rate at which healthy}\\ \text{individuals become infected}\end{bmatrix} - \omega L(t,s)\\
\frac{\partial I(t,s)}{\partial t} &= \omega L(t,s) - \mu I(t,s)\\
\frac{\partial R(t,s)}{\partial t} &= \mu I(t,s)
\end{aligned} \tag{8.20}$$

where the $\partial/\partial t$ notation means the *partial* derivative with respect to time, with the same interpretation as in the differential equation models of Chapter 5. The rate of infection term at position s now depends, as in the discrete-generation equation 8.2, on infectious units (e.g., spores) that are produced by infectious individuals at all possible positions ξ. The density of infectious individuals at time t at position ξ is represented by $I(t, \xi)$. The probability that an infectious unit (e.g., spore) produced by an infectious individual at ξ will contact a host individual at position s is now given by the contact distribution, $D(s - \xi)$, multiplied by θ'. Spores may be deposited between plants, for instance. Thus, θ' reflects the fact that some infectious units never contact a host anywhere. If each infectious individual produces α spores per time unit, the total influx of spores at position s per time unit due to all infectious individuals at position ξ is given by $\alpha I(t, \xi)D(s - \xi)\theta'$. To calculate the total influx rate of spores at position s due to all infectious individual in the crop we need to add up all contributions of all infectious individuals at all possible positions ξ. This amounts to integrating over ξ and we have:

$$\begin{bmatrix}\text{rate of spore influx}\\ \text{at postition } s\end{bmatrix} = \alpha\theta' \int_{-\infty}^{\infty} I(t,\xi)D(s-\xi)\,d\xi \tag{8.21}$$

Further, noting that each infectious unit in contact with a host has a probability ψ to cause an infection, and following the same reasoning as for the derivation of the infection term in the H-L-I-R model in section 5.2.3.1, we find that:

$$\begin{bmatrix}\text{rate at which healthy}\\ \text{individuals become infected}\end{bmatrix} = \alpha\psi\theta' H(t,s) \int_{-\infty}^{\infty} I(t,\xi)D(s-\xi)\,d\xi \tag{8.22}$$

The model equations become:

$$\begin{aligned}
\frac{\partial H(t,s)}{\partial t} &= -\alpha\psi\theta' H(t,s)\int_{-\infty}^{\infty} I(t,\xi)D(s-\xi)\,d\xi\\
\frac{\partial L(t,s)}{\partial t} &= \alpha\psi\theta' H(t,s)\int_{-\infty}^{\infty} I(t,\xi)D(s-\xi)\,d\xi - \omega L(t,s)\\
\frac{\partial I(t,s)}{\partial t} &= \omega L(t,s) - \mu I(t,s)\\
\frac{\partial R(t,s)}{\partial t} &= \mu I(t,s)
\end{aligned} \tag{8.23}$$

We can simplify the expression by defining the transmission rate as $\beta = \alpha\theta'\psi$, as done in Chapter 5, and use β in equation 8.23. It is important to note that θ' may not always be explicitly used in the model formulation. That is, any scaling of the contact distribution simply becomes an implicit part of the transmission parameter (together with sporulation rate and probability of infection).

Note that this model *does* include density dependence in terms of disease-free host density. Minogue and Fry (1983a, b) analyzed a model including similar extensions of a purely temporal model to a spatio-temporal model, but they used the Vanderplank specification of latent and infectious periods (see equation 5.44). Models for the spatial expansion of epidemics also have been developed to include the more general sporulation ("infectivity") curve approach of Kermack and McKendrick (1927) for the latent and infectious periods as described in section 5.2.5. These models were first developed and analyzed by Diekmann (1978, 1979) and Thieme (1977, 1979) and were subsequently operationalized by van den Bosch et al. (1988a–c, 1990a, b, 1992, 1999) to analyze biological invasions and the expansion of plant disease foci.

By summing the partial differential equations for latent, infectious, and removed disease states, one can determine the equation for the change in total disease at a given time and location:

$$\frac{\partial Y(t,s)}{\partial t} = \beta H(t,s)\int_{-\infty}^{\infty} I(t,\xi)D(s-\xi)\,d\xi \tag{8.24}$$

When density of diseased individuals is low, $H(t, s)$ is near the initial total host density, H_0 (and we assume here that H_0 does not depend on location). We can thus assume that $H(t, s) \cong H_0$. Then, equation 8.24 can be approximated by:

$$\frac{\partial Y(t,s)}{\partial t} = \beta H_0\int_{-\infty}^{\infty} I(t,\xi)D(s-\xi)\,d\xi \tag{8.25}$$

It is helpful to compare equation 8.24 with the one for a purely temporal epidemic given in the previous chapter (section 5.2.3) for the H-L-I-R model. In equation 5.31, the density of new infections at a given time is based on the density of disease-free individuals and infectious

individuals for the entire population, $H(t)$ and $I(t)$, respectively, with $\beta H(t)I(t)$ as the transmission term. In equation 8.24, the density of new infections at a given time at a given location depends on healthy host density at that location, $H(t, s)$, and on infectious units over the entire area (all possible ξ values) in which the epidemic occurs. By the nature of contact distributions (see Fig. 8.5), infectious individuals at or near location s contribute more to new infections than infectious units farther away.

The basic reproduction number for this model is, as with the non-spatial situation, defined as $R_0 = \beta H_0/\mu$. This is because one infected individual in an otherwise disease-free population of H_0 individuals produces a total of α/μ $(=\alpha \cdot \iota)$ spores (for the example of a spore-producing fungus) during its infectious period, with $\alpha\theta' H_0/\mu$ spores contacting a host (somewhere), and with $\alpha\theta'\psi H_0/\mu$ resulting infections. In other words, the contact distribution does not affect the calculation of R_0.

There is no analytical solution to equation 8.23; rather, numerical solutions must be used, based on use of a computer program such as MATHCAD or MATHEMATICA. Because the epidemic model now involves partial derivatives, the required numerical integration methodology is more involved than that used for numerically solving the ordinary differential equations of Chapter 5. An example solution is shown for selected times as the heavy (thick) curves in Fig. 8.11, based on the double-exponential contact distribution (see Table 8.1). The pattern of disease gradients is very similar to those found for the more descriptive logistic-logistic spatio-temporal model of Chapter 7 (see Fig. 7.14) and, for values of Y not close to H_0, for the non-overlapping generation (and density-independent) epidemic model (Fig. 8.6B). Density of diseased individuals approaches the final disease level, which depends on R_0. When the ordinate is plotted on a log scale, a straight-line relation with distance is seen, except for densities above 100 or so, where the density dependence is most apparent. The simulation also shows an apparent equal distance between the gradient lines at low values of $Y(t, s)$, implying that the rate of disease expansion is constant (with this contact distribution). Because we are now dealing with continuous time, we refer to the velocity of disease expansion as $C(t)$ instead of as C_n, which is equivalent to $\partial s/\partial t$ discussed in the previous chapter (change in distance where a given level of disease is found with change in time).

When the modified power law distribution is used, a different pattern is found for the gradients (Fig. 8.12). In particular, even when $Y(t, s)$ is plotted on a log scale, there is strong curvature, especially at large distances from the source. The pattern is very similar to that found for the non-overlapping generation epidemic model with the same contact distribution (Fig. 8.6D), except for large Y. As seen in Fig. 8.6D, there is a very apparent increasing distance between the gradient curves at low values of $Y(t, s)$.

8.3.2.2 Disease expansion rates—traveling waves. In the temporal models discussed in Chapter 5, we took great care to describe the density of healthy individuals and how this density affects the number of new infections

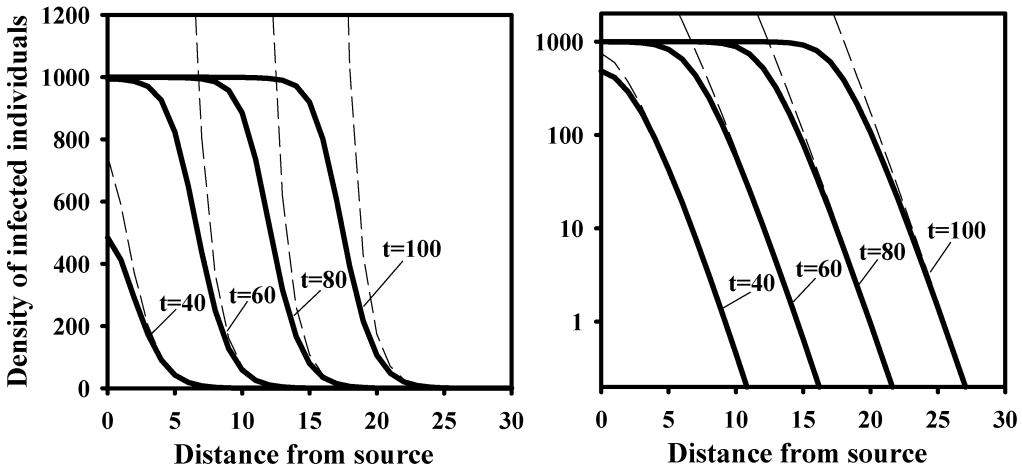

FIG. 8.11. Numerical solution of the spatial H-L-I-R epidemic model, based on use of a double-exponential contact distribution (with $\sigma = 1$) (see Table 8.1). Vertical axis is disease density, $Y(t, s)$, and horizontal axis is distance from the original inoculum source. Thick solid lines are gradients of disease for the density-dependent situation (equation 8.23), at selected times, t. Thin dashed lines are disease gradients for the density-independent situation (equation 8.23, with $H_0 = 1000$ used instead of $H(t, s)$). Note that in the front of the epidemic (large distances with low $Y(t, s)$), the two sets of gradients merge. Disease spread was in both directions from the original source (i.e., for positive and negative s), but only positive distances are shown. Other parameter values were: $\beta H_0 = 0.6$, $\omega = 1/7$, and $\mu = 1/10$.

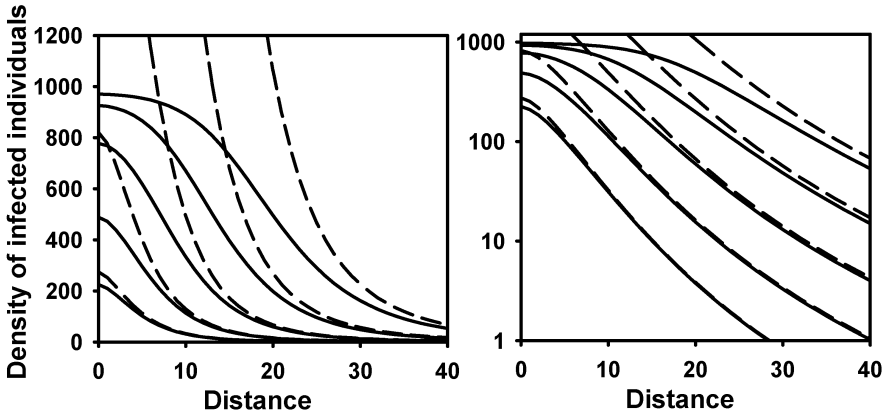

FIG. 8.12. Numerical solution of the spatial H-L-I-R epidemic model, based on use of a modified power law contact distribution (with $\Psi = \lambda = 3$) (see Table 8.1). Vertical axis is disease density, $Y(t, s)$, and horizontal axis is distance from the original inoculum source. Thick solid lines are gradients of disease for the density-dependent situation (equation 8.23), at selected times, t. Thin dashed lines are disease gradients for the density-independent situation (equation 8.23, with $H_0 = 1000$ used instead of $H(t, s)$). Note that in the front of the epidemic (large distances with low $Y(t, s)$), the two sets of gradients merge. Disease spread was in both directions from the original source (i.e., for positive and negative s), but only positive distances are shown. Other parameter values were: $\beta H_0 = 0.6$, $\omega = 1/7$, and $\mu = 1/10$.

produced per unit area per time unit. This is, of course, for the obvious reason that if the density of healthy individuals decreases during the epidemic, the number of new infections (for a given density of infectious individuals) will automatically also decrease. In other words, there is less opportunity for new infections as $H(t)$ declines (e.g., more and more spores land on diseased individuals). The effect of such density dependence on the rate of epidemic expansion has been studied by mathematicians in quite some detail. The work on this goes back to Kolmogorov et al. (1937) and Bramson (1983). Density dependent extensions of model equation 8.2 have been studied in detail by Weinberger (1978, 1982), and continuous-time versions of the model have been studied by Diekmann (1978, 1979), Thieme (1977, 1979) and Mollison (1972).

These studies considered the situation where the tail of the contact distribution falls off faster than exponentially with distance from the source. No particular contact distribution is used, so that general results are valid for *any* contact distribution with exponentially bounded tails. As we have discussed in section 8.2 for non-overlapping generations, this type of contact distribution implies that a traveling wave develops when there is no density dependence. A remarkable achievement of these cited theoretical studies is that *the rate of epidemic expansion in the density-dependent models is the* same *as the rate of expansion calculated from the density-independent models*. In other words, the results show that $C(t)$ approaches a constant velocity (C) and this velocity is the same in density dependent and density independent models. All the studies use a similar type of density dependence in the model formulation, which specifies that: the number of new infections produced per infectious individual at places and times where the density dependence matters is always less than when density dependence does not matter. In other words, $\partial Y(t, s)/\partial t$ decreases as $H(t, s)$ decreases for a given density of infectious individuals (since there are fewer opportunities for infection at location s). Moreover, the mathematical work assumed that the form of the contact distribution is not influenced by the density of healthy, infected and removed individuals. (Actually these authors use an even more general assumption, which we do not deal with here.)

The highly significant result can easily be understood intuitively by returning to Fig. 8.11. Superimposed on the disease density curves generated from density-dependent equation 8.23 are the disease gradients for the density independent case (i.e., with H_0 for $H(t, s)$ in equation 8.23), drawn as thin broken lines. Put another way, the thick (heavy) lines are the solutions of more complex equation 8.24, and the thin lines are the solutions of the simpler equation 8.25 for $\partial Y(t, s)/\partial t$. All parameter values are the same for both epidemic models. As expected, the use of the density-independent model ultimately results in meaningless results close to the original inoculum source when, in fact, there is an upper limit to disease density (H_0). In particular, prediction of $Y(t, s)$ can exceed H_0, given sufficient time. However, the fronts of the epidemics (i.e., Y values relatively far from the origin) are virtually identical for the two situations.

It may be helpful to the reader to refer back to the theory of polycyclic epidemics described in Chapter 5. Even for models of temporal disease progress for polycyclic diseases, an exponential growth model provided a good approximation when disease density was low. In a general sense, this approximation at low Y carries over to spatio-temporal models, for both the time and space dimensions. That is, because $H_0 \approx H(t, s)$ at small $Y(t, s)$, simplifications are possible. So, no matter how

complicated the model is for dynamics at moderate or high disease density, much simpler models are applicable far from the original inoculum source. The expanding epidemic is drawn forward through space by the infectious individuals at the forefront of the epidemic. It is exactly in the front of the epidemic where the density of infected individuals is so small that the dynamics of the epidemic is not affected by any (healthy-host) density-dependent process.

Several important mathematical results are obtainable based on the use of a density-independent model (Diekmann, 1978, 1979; Thieme, 1977, 1979) for the epidemic. Without mathematical proof, we emphasize the key ones related to the velocity or rate of disease expansion, $C(t)$. A major result is that after an initial period of time, $C(t)$ approaches an asymptotic constant velocity (C) when the contact distribution is exponentially bounded. As with the model with discrete generations, and as demonstrated in Fig. 8.11 for one set of parameters, an exponentially bounded contact distributions leads to a traveling epidemic wave. In particular, C can be written as an approximate function of R_0:

$$C \cong A + B\ln(R_0) \tag{8.26}$$

in which A and B are constants. These constants depend on other model parameters, such as σ and ω ($1/\wp$), but the exact functional dependencies will be influenced by the contact distribution used. We have not derived the relations for various contact distributions here but merely used the close-to-linear relation between C and $\ln(R_0)$ seen in Fig. 8.9. Based on equation 8.26, one can see that the farther spores (or other infectious units) travel, as indicated by σ, and the greater the new infections per each infectious unit, as indicated by R_0, the greater the velocity of focus expansion. From a control perspective, one can limit the invasion of a disease into new areas by reducing σ (perhaps with barriers) or by reducing R_0.

An important point is that it does take some time for the constant velocity, C, to be approached. The time period is not easy to define exactly (van den Bosch et al., 1990b), but depends on all model parameters. No estimates of convergence time exist and further work in this area is needed. Observed epidemics confirm that there is a time period before a constant expansion rate is found (Minogue and Fry, 1983b; van den Bosch et al., 1988c).

8.3.2.3 Disease expansion rates—dispersive traveling waves. It appears that the density-independent version of equation 8.23 (H_0 used for $H(t, s)$) accurately describes the front of disease expansion also when the contact distribution is not exponentially bounded. This is demonstrated in Fig. 8.12, where the thin broken lines correspond to numerical solutions of the non-density dependent version of equation 8.25. However, the mathematical theory has not been adequately developed for these types of contact distributions to give more detailed results. By referring back to the non-overlapping generation model, one should expect that the increase in $C(t)$ is proportional to the scale of contact distance ($1/\Psi$ in root, and λ/Ψ in modified power law distributions). Based on continuous-time results for exponentially-bounded distributions, one would also expect that $C(t)$ is inversely related to mean latent period. More research is clearly needed here, because many diseases are often characterized by power-law type gradients (e.g., Sackett and Mundt, 2005a), as discussed and demonstrated in the previous chapter. Consistent with a power law type of gradient, calculated velocities often do increase with time (Cowger et al., 2005; Madden et al., 1990a; Sackett and Mundt, 2005b).

8.3.2.4 Multi-seasonal epidemic expansion. van den Bosch et al. (1999) developed a model to describe both the within growing season and the multi-seasonal expansion of plant disease epidemics on a continental scale. We do not give the theory here, but the model describes the dynamics of a polycyclic epidemic during the crop growth season and combines this with a simple description of pathogen survival through the season where no crop host is present.

8.3.3 Disease expansion with monocyclic diseases

The basic concepts used here for continuous time and for discrete non-overlapping generations can easily be applied to monocyclic diseases, in which the only infections occur from the primary inoculum (the primary infectious units, such as overwintering spores in the soil or spores produced on weeds on the edge of a field). We consider only the continuous time scenario, and build on a portion of the model introduced in section 5.4 for the situation where primary infectious occur throughout the epidemic. Readers should re-study that section before reading the following material.

We can define $x(t, s)$ as the density of primary infectious units available for infection ("released spores") at time t and location s (with units of 1/time). We can consider $x(t, s)$ to be fixed, or to change over time. $\theta' D(s - \xi)$ is the probability that a primary infectious unit originating at ξ comes in contact with a host individual at s, and ψ_p is the probability that an infectious unit in contact with a host causes an infection. If there are no secondary infections ($\beta = 0$), one can write:

$$\frac{\partial Y(t, s)}{\partial t} = \theta' \psi_p H(t, s) \int_{-\infty}^{\infty} x(t, \xi) D(s - \xi)\, d\xi \tag{8.27}$$

A comparison of this equation with equation 8.24 shows that $x(t, s)$ is used here instead of spores produced by

infectious diseased individuals [$\alpha I(t, \xi)$]. For the simpler situation where there is a single inoculum source, at $s = 0$, the equations simplifies to:

$$\frac{\partial Y(t, s)}{\partial t} = \theta' \psi_p H(t, s) x(t, 0) D(s) \qquad (8.28)$$

No integration is necessary over space, since inoculum arriving at s can only be coming from *one* location (s units away). An even simpler model is possible if one can assume that inoculum density does not change, so that $x(t, 0)$ is a constant ($\equiv x_0$). Obviously, a fixed density of available inoculum is not to be expected exactly, but it may be quite a reasonable assumption over the time period of concern. Then one can write:

$$\frac{\partial Y(t, s)}{\partial t} = \theta' \psi_p x_0 D(s) H(t, s) \qquad (8.29)$$

Equation 8.29 is a spatial expansion of the monocyclic model of Chapter 4 (equation 4.8), in terms of Y (rather than y), where $\theta' \psi_p x_0$ is analogous to the simple-interest rate parameter r_M. Minogue (1986) showed how to use and interpret this model in terms of disease expansion. Interestingly, there is an analytical solution for this model with a single inoculum source.

$$Y(t, s) = H_0(1 - \exp(-\theta' \psi_p x_0 D(s) t)) \qquad (8.30)$$

Equation 8.30 assumes that $Y(0, s)$ ($= Y_0$) is 0 everywhere and the epidemic starts with all inoculum at $s = 0$. A numerical example is shown in Fig. 8.13 for the double-exponential contact distribution (see Table 8.1) and one set of parameters. An interesting result is that the rate of disease expansion ($C(t)$) appears to decline over time; that is, the distance between the curves at a single (low) value of Y declines over time.

Because there is an analytical solution for this model, one can readily determine the rate of disease expansion. With the double-exponential contact distribution, a little algebraic manipulation of equation 8.30 results in:

$$\ln(-\ln(1 - y(t, s))) = \ln\left(\frac{\theta' \psi_p x_0 t}{\sigma\sqrt{2}}\right) - \frac{\sqrt{2}}{\sigma} | s | \qquad (8.31)$$

in which $y(t, s) = Y(t, s)/H_0$. The left-hand side of equation 8.31 is the so-called complementary log-log transformation (CLL), which is used more extensively in Chapter 9 for other purposes. The CLL declines with distance, with a slope of $\sqrt{2}/\sigma$, and increases with the log of time. To determine $C(t)$, one can first define a fixed value of $y(t, s)$, say y', and find the value of s that gives y' ($s_{y'(t)}$, or s' for short). This results in an equation for the isopath in relation to time:

$$s' = \frac{\sigma}{\sqrt{2}}\left(\ln\left(\frac{\theta' \psi_p x_0 t}{\sigma\sqrt{2}}\right) - \ln(-\ln(1 - y'))\right) \qquad (8.32)$$

The rate of disease expansion ($C(t)$) can be obtained by determining the partial derivative of s' with respect to t ($\partial s'/\partial t$). For equation 8.32, the partial derivative is given by:

$$\frac{\partial s'}{\partial t} = C(t) = \frac{\sigma}{t\sqrt{2}} \qquad (8.33)$$

Based on equation 8.33, one can see that $C(t)$ is directly proportional to σ, as found for polycyclic diseases (see section 8.2)—the longer the distance traveled by spores, for example, the greater the rate of disease expansion.

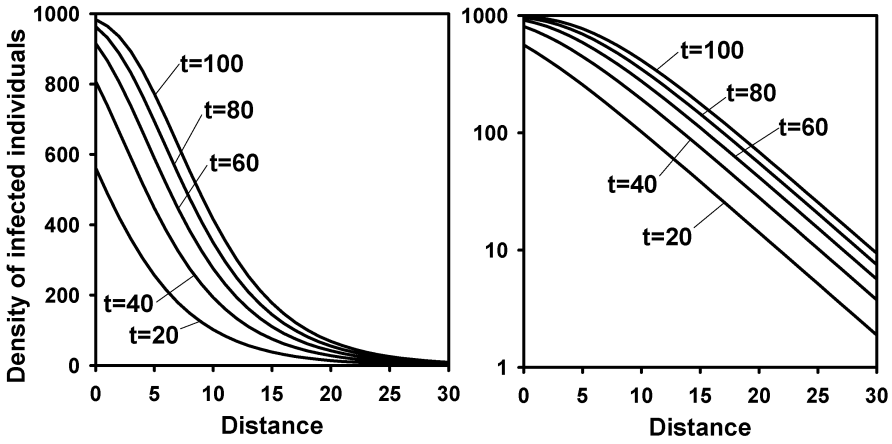

FIG. 8.13. Numerical solution of the spatial monocyclic epidemic model for a single inoculum source at distance 0 (equation 8.30), based on use of a double-exponential contact distribution (with $\sigma = 7$) (see Table 8.1). Vertical axis is disease density, $Y(t, s)$, and horizontal axis is distance from the inoculum source. Solid lines are gradients of disease at selected times, t. Other parameter values: $H_0 = 1000$, and $\theta' \psi_p x_0 = 0.4$.

Moreover, $C(t)$ is inversely proportional to time since the start of the epidemic. This declining expansion rate makes intuitive sense, because infections do not lead to further infections with monocyclic diseases, and diseased individuals at the front of the epidemic do not move the disease forward in space. In other words, when all infections beyond the origin originate exclusively from inoculum at the origin, there is a decline in expansion over time (even with a fixed inoculum density at the source).

Equation 8.30 is applicable for any of the contact distributions used in this chapter. However, equations 8.31–8.33 are specific for the double-exponential distribution. Other contact distributions will lead to different formulations for s' and $C(t)$, sometimes with more complicated equations for $C(t)$.

The results here should be contrasted with those obtained in Chapter 7. It was found in section 7.6.4 to be challenging to link the monocyclic temporal model with a logical disease-gradient model and obtain an easy-to-use, realistic, and interpretable combined model. By being more realistic in the model formulation, equation 8.30, and using contact distributions consistent with observed dispersal gradients, an interpretable spatio-temporal model is readily obtained.

8.3.4 Multiple foci and temporal dynamics

The previous sections of this chapter showed that a considerable understanding of the spread of disease can be obtained using a range of models that capture the essential biological features of epidemics in time and space. Depending on the model formulation, exact or approximate equations can be developed to predict the rate of disease expansion from a single focus. Of course, many epidemics are initiated from several foci instead of one, but, fortunately, the principles of spread expressed in this chapter hold for any number of foci. With multiple original foci, the concept of a simple single expansion rate for the epidemic no longer applies. There is, in fact, expansion of each disease focus, possibly with mergers of originally distinct foci, as well as a resulting "global" expansion rate at the edge of the full spatial extent of the epidemic.

Models such as the spatial H-L-I-R one in equation 8.23 can be numerically solved using mathematics computer programs for any set of initial conditions $[Y(0, s)]$. For example, infectious individuals can be located at any number of locations over the spatial domain. With multiple foci, graphs of observed disease gradients or predicted disease gradients from equation 8.23 (such as in Figs. 8.11 and 8.12) are of less direct interpretation. This is because there is no longer an unambiguous $s = 0$ point in terms of the starting conditions with multiple original foci. For this situation, maps of (observed or predicted) disease density at selected times are often produced instead of simple curves of density versus distance. Maps of disease intensity will be considered in detail starting in the next chapter.

One can also use spatio-temporal models (such as equation 5.23) to determine disease progress curves for the *entire* field or region in which the epidemic is occurring, not just for specific locations, based on any starting conditions. As done in Chapters 4 and 5, we can define $Y(t)$ as the total number of diseased individuals over all locations at time t. With explicit consideration of space, $Y(t)$ is given by:

$$Y(t) = \int_{-\infty}^{\infty} Y(t, s)\, ds.$$

One can then explore the effects of temporal parameters [e.g., β $(=\alpha\theta'\psi)$, ω $(=1/\wp)$, and μ $(=1/\iota)$], spatial parameters (σ for certain contact distributions), and initial conditions on the disease progress curves. This is demonstrated as line 3 (l_3) in Fig. 8.14 for an example

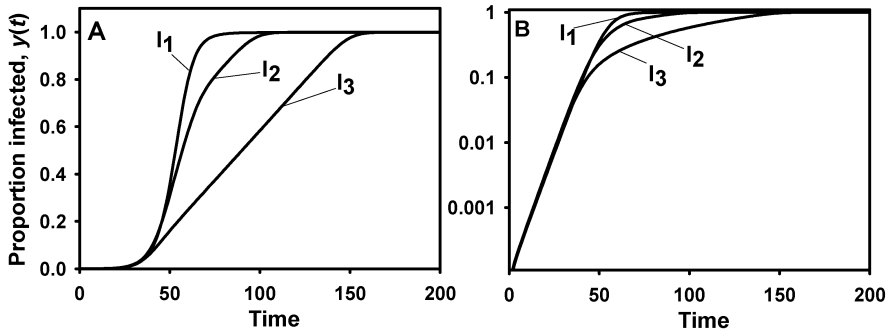

FIG. 8.14. Total proportion of diseased individuals across all locations, $y(t)$ $[=Y(t)/H_0]$, versus time, t, for a spatial H-L-I-R model (equation 8.23) with a double-exponential contact distribution ($\sigma = 1$; see Table 8.1), or aspatial model (equation 5.30). (**A**) arithmetic scale and (**B**) log scale. Line l_1: realization for equation 5.30, which corresponds to homogeneous mixing. Line l_2: realization for equation 8.23, with three evenly inoculum line sources. Line l_3: realization of equation 8.23, with a single inoculum line source in the center of the spatial range. Total density of diseased individuals at $t = 0$ was the same for all three epidemics. Other parameter values were: $\beta H_0 = 0.6$, $\omega = 1/7$, and $\mu = 1/10$.

epidemic where s ranges from 0 to 60, with one initial focus in the field center. Results are expressed as proportions, $y(t) = Y(t)/H_0$. The double exponential contact distribution was used, and the parameters are the same as used in generating the gradients in Fig. 8.11. Line 1 (l_1) is corresponds to the $y(t)$ values for predicted based on the aspatial H-L-I-R model of Chapter 5 (equation 5.30) with the same initial density of infected individuals and the same model parameters. With the aspatial model, homogeneous mixing of the infectious units throughout the host population is implied (i.e., contact does not depend on position).

As seen in the graph, aggregation of disease into a single initial focus results in a slower increase in total disease density (not necessarily at any given location) than when there is homogeneous mixing of inoculum, at least for the chosen value of σ. This is because, for the typical situation of a contact distribution that declines rapidly with increasing distance, infectious units produced by an infectious individual are likely to come in contact with a previously infected individual when the diseased individuals are mostly near each other to start with. Careful evaluation of Fig. 8.14 shows that the difference between the two cases is virtually nonexistent at low $y(t)$, since there are many disease-free host individuals at all locations to come in contact with inoculum. Then, the differences in the two cases become apparent as the effect of the declining disease-free hosts is manifested. Ultimately, the same final disease density is reached for both cases. It should be noted that a larger value of σ would diminish the effect shown in the figure. This is because infectious units travel greater distances, on average, resulting is a lower likelihood that an infectious unit will come in contact with a diseased host.

The line 2 (l_2) in Fig. 8.14 corresponds to the same situation as the l_3 line, but with three evenly separated foci at time 0, instead of one focus. Initial total density of diseased individuals at time 0 was the same. The rate of disease increase was, again, lower than for the aspatial (homogeneous mixing) case (l_1), but the difference was not as great as between the aspatial case and the single-focus case. With well separated initially diseased individuals, there is less likelihood that spores will contact diseased host individuals than when all diseased individuals are initially together.

Various efforts have been made to account for some aspects of the spatial component of epidemics using strictly temporal models of disease progress (Gubbins and Gilligan, 1997; Hughes et al., 1997; Liu et al., 1987; Madden and Campbell, 1990; Waggoner and Rich, 1981; Yang and TeBeest, 1992; Zhang et al., 2000b). In particular, the proposed models account for, in principle, the non-homogeneous mixing of the infectious units without explicit consideration of the spatial location of individuals or the contact distribution. Sometimes the equations are called *spatially implicit models* of disease progress. The basic approach is to replace $I(t)$ and $H(t)$ in the transmission term ($\beta I(t)H(t)$) in temporal models (e.g., in equation 5.30) with nonlinear functions of infectious (or more simply, infected [$Y(t)$]) and/or healthy host density, $f_I(I(t))$ and $f_H(H(t))$, respectively. The equation component for $dH(t)/dt$ in the H-L-I-R model would then be written as:

$$\frac{dH(t)}{dt} = -\beta^* f_I(I(t)) \cdot f_H(H(t)) \qquad (8.34)$$

This type of modification can also be used in models with a dynamic host crop in the absence of disease (see section 6.2 of Chapter 6), as well as in models that include vector dynamics (see section 6.4). The "transmission parameter" is written as β^* because it is no longer simply interpreted just as $\alpha\theta'\psi$. Power functions are common choices for $f_I(I(t))$ and $f_H(H(t))$, such as I^a and H^b (Liu et al., 1987). Then, a special case is when $a = b = 1$, where the standard (homogeneous mixing) temporal model is obtained. Values of a or b less than 1 are indicative of non-homogeneous mixing. For instance, with $b = 0.8$, the reduction in healthy plants per unit time (or the increase in infected plants) will be less than when $b = 1$ (homogeneous mixing), since $H^{0.8} < H$. This would be, in principle, accounting for the fact that there are fewer contacts between infectious units and healthy plants with non-homogeneous mixing than with homogeneous mixing.

More complicated functions can also be used in the transmission expression (see equations 6 and 7 in Hughes et al., 1997). Although equation 8.34 and similar models can provide good fits to observed or simulated disease progress curves, there is no formal quantitative linkage between a and b (or more generally, the explicit functions of I and H) and the underlying contact-distribution parameters and the initial spatial conditions (i.e., locations of initially diseased individuals). Thus, this approach should only be considered an approximation to represent the increase in density of diseased individuals for explicit spatio-temporal epidemic processes.

As will be shown in Chapter 9, there are useful alternatives to the methods discussed in this and the previous chapter for characterizing and quantifying the spatial component of epidemics. These approaches are based on *statistical* models of patterns of disease intensity rather than mathematical models of contact distributions and disease gradients. These statistical models also lead to modifications of temporal epidemic models that are in the same format as general equation 8.34 (Hughes et al., 1997; McRoberts et al., 1996; Waggoner and Rich, 1981) although we do not give more details in this book. The general conclusion from this work is that increasing aggregation of disease leads to slower rates of increase in $Y(t)$ over time, as found here based on explicit consideration of the spatial component of the epidemic (Fig. 8.14).

Modification of an H-L-I-R (or similar) model with power functions is analogous to modification of the

$1 - y$ term in the logistic equation (equation 4.14) to obtain the Richards growth model (see equation 4.22). By adding one or two parameters to the model, one would expect better fits to the data than obtained without the new parameters. This is, indeed, generally found with experimental data (e.g., Gubbins and Gilligan, 1997; Zhang et al., 2000b). The next research step will be to develop spatially implicit models (which could include modifications of the ones currently in use) that can be more closely related to the parameters of the contact distribution and the initial spatial conditions.

8.4 An Application

The effect of mixing two or more cultivars to control plant diseases is well documented (Browning and Frey, 1969; Mundt and Browning, 1985; Wolfe, 1985). In this section, we consider the rate of epidemic expansion in cultivar mixtures to demonstrate an application of the above model. We consider a mixture of a susceptible and a completely resistant cultivar and study how the fraction of the crop consisting of the susceptible cultivar, φ, influences the rate of expansion (C). We use results for two well-studied pathosystems, stripe rust (caused by *Puccinia striiformis*) of wheat, and rust (caused by *Uromyces appendiculatus*) on bean (van den Bosch et al., 1988c, 1990a; Assefa et al., 1995). For these pathosystems we assume here that the spatial expansion of the disease may be described by a traveling wave (which is consistent with some of the published studies).

In this section, we will derive an equation that relates the fraction of susceptible crop, φ, in a mixture of susceptible and completely resistant cultivars, to the rate of focus expansion. It turns out that the other model parameters, such as R_0, σ, and $\wp$ can be condensed into two constants, and that their numerical values are not needed to either derive our hypothesis about the relation between φ and C, nor to test the hypothesis we derive.

As indicated above, C for both the Gaussian and the double exponential contact distributions can be predicted by equation 8.26. Suppose that our mixture is a random mixture of susceptible and resistant plants. The infectious units produced by an infectious individual are then distributed over the two cultivars in proportion to their fractions in the mixture. This implies that a fraction φ of the infectious units end up on the susceptible cultivar and a fraction $1 - \varphi$ on the resistant cultivar. Only the infectious units on the susceptible cultivar cause infections and, therefore, the basic reproduction number realized in the mixture is:

$$R_0 = \varphi \tilde{R}_0 \tag{8.35}$$

where we use $\tilde{R}_0$ to represent the basic reproduction number of the disease when the crop consists only of susceptible cultivar. Substituting equation 8.35 into 8.26 gives:

$$C = A + B \ln(\varphi \tilde{R}_0) \tag{8.36}$$

which can also be written as:

$$C = \tilde{C} + B \ln(\varphi) \tag{8.37}$$

where $\tilde{C} = A + B \ln(\tilde{R}_0)$, the (asymptotic) constant rate of expansion when the crop consists of susceptible cultivar only.

Equation 8.37 suggests that the rate of disease expansion is linearly related to the logarithm of the fraction of susceptible cultivar in the crop. This is a hypothesis that can be tested experimentally. Plots have to be established with a range of seed mixtures to cover the range of proportions, from $\varphi = 1$ down to $\varphi \approx 0$; a plot with φ equal to 0 provides no information since then there is no epidemic. The center plant of each plot (the center line could also be used) is inoculated by the pathogen to serve as the initial source of infectious units, and the developing epidemic is observed and disease gradients are determined over the time period the disease is spreading through the plot. Using the empirical methods in section 7.6.3 of the previous chapter, one can determine the distance where a small value of disease is found at each time, $Y'(t, s')$. In brief, this can be done by drawing a horizontal line through a gradient graph (as done in Fig. 8.6) at Y', and finding the distance where this value occurs at each time. To avoid interpolation, one could also fit an empirical gradient model to the data at each time [making sure to use a gradient model that adjusts for a maximum y of 1 (see section 7.5.2)], and rearrange the prediction equation for distance (s'). From these maps the rate of epidemic expansion is calculated. If the hypothesis is correct we will find a straight line relation between the rate of expansion and the logarithm of the fraction of susceptible crop (assuming that the contact distribution is exponentially bounded, and that the contact distribution does not depend on φ.

Fig. 8.15 shows the results of the experiments of van den Bosch et al. (1999) and Assefa et al. (1995). From the figure we see that the experimental results agree with the hypothesized relation, although there is considerable variation in C values for a given fraction of susceptibles in the mixture. For applications, it is helpful to have rules of thumb telling how certain epidemiological quantities, such as the rate of epidemic expansion, varies with the fraction of susceptible plants in the crop. Generalizing from the theoretical considerations in this section and the experiments done on two different plant pathogen systems we can state the rule of thumb: In a mixture of susceptible and resistant plants, the rate of epidemic expansion increases linearly with the logarithm of the fraction of susceptible plants in the crop.

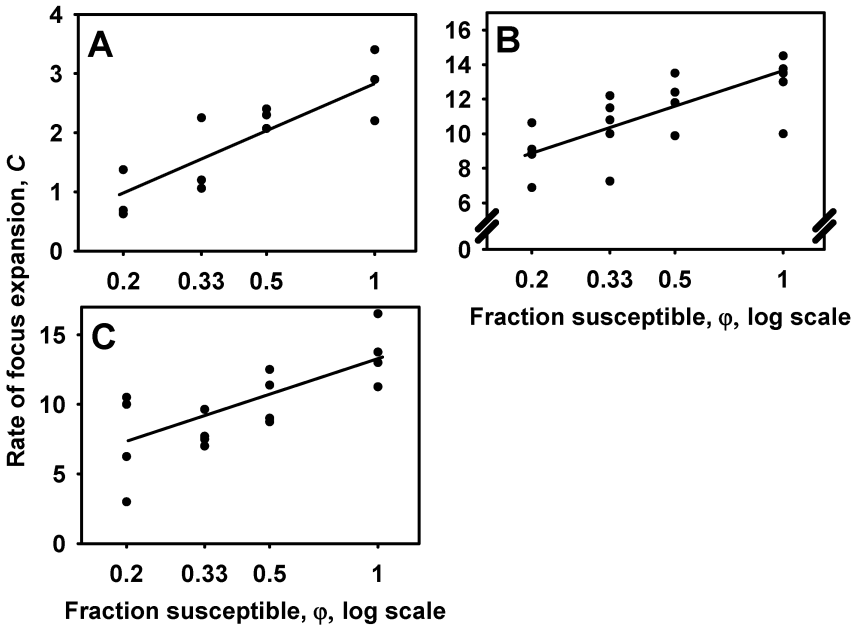

FIG. 8.15. Observations on the rate of disease expansion in relation to the fraction of susceptible crop in a mixture. Drawn line is the fit of equation 8.36 based on least squares regression. (**A**) stripe rust of wheat, caused by *Puccinia striiformis* (data from van den Bosch et al., 1990b). (**B**) and (**C**) bean rust, caused by *Uromyces appendiculatus* in (**B**) 1990 and (**C**) 1991 (data from Assefa et al., 1995).

Complicating the issue is the fact that there is evidence from other studies (done in other locations) that power law type contact distributions may best describe stripe rust and bean rust disease gradients (Maffia and Berger, 1998; Sackett and Mundt, 2005a, b; see also Figs. 7.11 and 7.13). In particular, Cowger et al. (2005) recently showed that the rate of focus expansion increased exponentially with time for stripe rust, which is consistent with a power-law contact distribution (equation 8.17). Thus, the experimental data supports the epidemiological theory for their pathosystem.

8.5 Concluding Remarks

In this chapter we derived a simple model formulation for the spatial expansion of polycyclic (compound interest) plant disease. The model is an integration of: (i) reproduction (in the general sense)—i.e., infections from each infected individual, including the production of infectious units and their probability of causing an infection when in contact with a disease-free host individual—and (ii) the contact distribution, the measure of dispersal of infectious units (equation 8.2 or 8.23). Analyzing the model we found that there are two essentially different types of expansion rates for the front of the epidemic, or for a specified isopath of low disease intensity (Y'). The categories are determined entirely by the form of the contact distribution. A traveling wave, in which the front moves at constant rate, is found for contact distributions with exponentially bounded tails. A dispersive wave, in which the front moves out with increasing velocity, is found for contact distributions with non-exponentially bounded tails. For dispersive waves, the rate of expansion of dispersive waves (acceleration) can be linear or nonlinear (e.g., exponential) functions of time. In all situations, the expansion rate is a positive function of the basic reproduction number, R_0.

The basic results for polycyclic diseases were fairly easy to show using a simple non-overlapping-generation model with no (host) density dependence (equations 8.2 and 8.3). The mathematical complexity increases greatly with consideration of latent and infectious periods, overlapping generations, and density dependence, but various authors have shown that the basic results hold for expansion of the epidemic (Diekmann, 1978, 1979; Thieme, 1977, 1979). In contrast to polycyclic diseases, the mathematical formulation is much simpler for monocyclic (simple interest) diseases, even when there is density dependence. Depending on the contact distribution, it can be shown that the expansion rate may decline with time.

Direct fitting of spatio-temporal models such as equation 8.23 to data is more difficult than fitting of the equivalent purely temporal models. This is because of the partial derivatives instead of ordinary derivatives that must be numerically integrated. For practical purposes, it is probably most useful to estimate model parameters from ancillary studies and then use these parameters to predict $Y(t, s)$ for a specific situation using numerical methods. For instance, gradients of disease intensity, especially early in the epidemic, provide the required information for calculating contact-distribution parameters (e.g., σ). Furthermore, C can be estimated by measurements of disease intensity at the front of the epidemic

at several times (see Chapter 7). Depending on the contact distribution, R_0 can then be estimated from σ and C.

There are several other approaches to model spatial dynamics of populations. Ecologists, theoretical biologists and sometimes plant pathologists use diffusion-type models (e.g., Gilligan, 1995; White and Gilligan, 1998). The basic assumption underlying diffusion equations is that the individuals, or the entities being dispersed (e.g., spores), move randomly through space (Minogue, 1989). This assumption may be a reasonable approximation, at least over certain spatial and temporal scales of interest, especially for animal species. This approach could, for example, be useful to model virus diseases with arthropod vectors, where the vector can reasonably be assumed to move randomly through the crop. A good introduction to the use of diffusion models is found in Okubo (1988). As a general rule, however, a diffusion-type model is unrealistically simplistic for many plant disease epidemics (see discussion in Minogue, 1986; and also Mollison, 1972), so we did not discuss development and applications of this model in the chapter.

Spatial structure can also be incorporated into epidemic models using spatially explicit simulation models (Kampmeijer and Zadoks, 1977; Sackett and Mundt, 2005b). As discussed in regards to temporal models (see section 5.6.3), simulation is especially useful for determining specific outcomes under specific conditions, such as in studying particular profiles of environmental conditions on spatio-temporal dynamics. Simulation is also valuable to explore disease outcomes at all locations and times when there are multiple initial inoculum sources, and also when researchers wish to study stochastic aspects of epidemics. The spatio-temporal stochastic model of Xu and Ridout (1998) is especially useful in this regard. For a more mathematical treatment of spatio-temporal dynamics for a lattice of plants, using stochastic modeling, readers should see Filipe and Gibson (1998) and Filipe and Maule (2003, 2004). Among several other results, Filipe and Maule (2004) developed a spatially implicit epidemic model (see equation 8.34) that is different from the power-function expression explained in section 8.3.4. Keeling (1999) presents an alternate way of addressing the problem of spatial structure of disease in temporal dynamics of epidemics without the need for simulation. Additional valuable references include Durrett and Levin (1993), Park et al. (2001), Stacey et al. (2004), and Gilligan (2002).

Although the rate of focus expansion is determined by the contact distribution, it appears that different contact distributions may be appropriate for the same disease under different circumstances (Assefa et al., 1995; Maffia and Berger, 1998; Sackett and Mundt, 2005a, b; van den Bosch et al, 1988c, 1990a). Results on realized disease gradients, and corresponding contact distributions, may reflect the scale of the experiments used to quantify disease spread. As discussed in Chapter 7, disease intensity must be measured at relatively large distances from the inoculum source in order to clearly distinguish among contact-distribution types. Moreover, based on Aylor (1999), based on physical principles of spore transport, it could be argued that different contact-distribution types may be of most relevance at different distance ranges from the source for a given pathogen. Clearly more research is needed in this area since results have a major consequence on predictions of disease spread into new areas.

Finally, a subject that is widely acknowledged for its importance in the spatial dynamics of populations, but is very little studied, is heterogeneity of the environment. Shigesada and Kawasaki (1997) and Shigesada et al. (1986) are excellent references on this topic.

8.6 Selected Readings

Aylor, D. E. 1986. A framework for examining inter-regional aerial transport of fungal spores. *Agric. For. Meteorol.* 38: 263–288.

Kot, M., Lewis, M. A., and van den Driessche, P. 1996. Dispersal data and the spread of invading organisms. *Ecology* 77: 2027–2042.

Shigesada, N., and Kawasaki, K. 1997. *Biological Invasions: Theory and Practice*. Oxford Series in Ecology and Evolution. Oxford University Press, New York.

van den Bosch, F., Metz, J. A. J., and Zadoks, J. C. 1999. Pandemics of focal plant disease, a model. *Phytopathology* 89: 495–505.

Zadoks, J. C., and van den Bosch, F. 1994. On the spread of plant disease: a theory on foci. *Annu. Rev. Phytopathol.* 32: 503–521.