O.A. van Herwaarden

Analysis of unexpected exits using the Fokker-Planck equation

Promotor: dr. ir. J. Grasman
hoogleraar in de Wiskunde, inclusief de Numerieke Wiskunde

NN08201, 2058

O.A. van Herwaarden

Analysis of unexpected exits using the Fokker-Planck equation

Proefschrift
ter verkrijging van de graad van doctor
in de landbouw- en milieuwetenschappen
op gezag van de rector magnificus,
dr. C.M. Karssen,
in het openbaar te verdedigen
op maandag 11 maart 1996
des namiddags te vier uur in de Aula
van de Landbouwuniversiteit te Wageningen.

ISN 922443.

BIBLIOTHEEK
LANDBOUWUNIVERSITEIT
WAGENINGEN

CIP-DATA KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Herwaarden, O.A. van

Analysis of unexpected exits using the Fokker-Planck
equation / O.A. van Herwaarden. - [S.l. : s.n.]. - Ill.
Thesis Landbouwuniversiteit Wageningen. - With ref. - With
summary in Dutch.
ISBN 90-5485-493-6
Subject headings: Fokker-Planck equation / exit problems
/ stochastic epidemics

NN08201, 2058

Stellingen

1. In de uitdrukking voor de conditionele verwachte aankomsttijd van een toestand bij een gegeven gedeelte van de rand van een domein is in Mangel (1979) ten onrechte een deling door de aankomstkans achterwege gelaten.
   Mangel, M. (1979), *Small fluctuations in systems with multiple steady states,* SIAM J. Appl. Math., 36, pp. 544-572.
   Dit proefschrift.

2. Voor een een-dimensionale stroming met een kleine diffusie is de conditionele verwachte reistijd naar een rand tegen de stroom in in eerste orde gelijk aan de verwachte reistijd in omgekeerde richting.
   Dit proefschrift.

3. $$\int_0^{\sqrt{(\beta/\alpha)}} \exp[-\alpha t - \beta/t]\,dt = \sqrt{\beta/\alpha}\,K_1(2\sqrt{\alpha\beta}\,) - \frac{1}{2\alpha}\exp[-2\sqrt{\alpha\beta}\,] \qquad (\alpha,\ \beta > 0),$$
   waarin K_1 een gemodificeerde Besselfunctie is.

4. In het wiskunde-onderwijs is het voor de overdracht van een nieuw begrip noodzakelijk de definitie vooraf te laten gaan door een voldoende aantal toepasselijke voorbeelden.

5. Eén van de belangrijkste problemen bij het wiskunde-onderwijs aan studenten uit niet-wiskunde studierichtingen is het motiveren van deze studenten. Het gebruik van toepassingen uit de desbetreffende richtingen is een belangrijk hulpmiddel bij de aanpak van dit probleem.
   Arnoldussen - van der Lugt, A., en O.A. van Herwaarden (1990), *Wiskunde in de landbouwwetenschappen*, Euclides, 66, pp. 105-110.

6. Analyse van schilderijen van Hollandse meesters uit de 17e eeuw laat zien dat correcte toepassing van de lineaire perspectief zich bij de meeste schilders beperkt tot het gebruik van één verdwijnpunt.

7. In de Bijbel gaat het niet om de vraag: wat is de waarheid?, maar: Wie is de waarheid?

8. Het christelijke geloof is niet een religie, maar een relatie.

O.A. van Herwaarden
Analysis of unexpected exits using the Fokker-Planck equation
Wageningen, 11 maart 1996

Voorwoord

Dit proefschrift is het resultaat van het onderzoek dat ik vanaf september 1989 heb verricht als medewerker van de Vakgroep Wiskunde van de Landbouwuniversiteit Wageningen. Graag wil ik op deze plaats een aantal personen bedanken die op verschillende wijzen hebben bijgedragen aan de totstandkoming van dit proefschrift.

Allereerst wil ik mijn promotor Johan Grasman van harte bedanken. Zonder zijn deskundige en stimulerende begeleiding zou dit proefschrift er niet zijn gekomen. Ik dank hem dat hij mij de mogelijkheid heeft geboden om naast mijn onderwijstaak ook onderzoek te doen. Niet alleen zijn inhoudelijke begeleiding wil ik hier noemen. Hij was ook in staat om mij op de juiste momenten en op de juiste wijze te inspireren en te motiveren. Bovendien heeft hij, toen het onderzoek naar de verwachte uitsterftijd van een infectieziekte vast zat, zich uitermate ingespannen om het weer vlot te trekken.

Mijn vrouw Elsa wil ik van harte bedanken voor haar grote steun en zorg. Deze zijn onmisbaar geweest bij het tot stand komen van het proefschrift. Een groot gedeelte van het onderzoek heeft plaatsgevonden buiten de reguliere werktijden. Dit is slechts mogelijk geweest doordat Elsa de zorgen thuis, vooral die voor onze kinderen Ada en Arjen, voor haar rekening heeft genomen. Ada en Arjen bedank ik dat ze mijn afwezigheid zo vaak 'voor lief' hebben genomen.

Mijn ouders wil ik bedanken voor hun zorg door de jaren heen. Mijn vader zou het schrijven en voltooien van een proefschrift graag hebben meegemaakt.

Mijn collega's en oud-collega's van de Vakgroep Wiskunde dank ik voor de goede werksfeer. Hun belangstelling voor het vorderen van het proefschrift heb ik altijd gewaardeerd. Een aantal van mijn collega's heb ik met name genoemd in de acknowledgements van diverse hoofdstukken. Van mijn oud-collega's wil ik hier met name noemen Annie Arnoldussen-van der Lugt en Mat Hendriks, bij wie ik zowel voor als na hun pensionering (resp. VUT) altijd met vragen over het werk terecht kon.

Ook het plezier bij mijn overige werkzaamheden, vooral het verzorgen van het onderwijs, is van grote invloed geweest op het werken aan mijn onderzoek. In het bijzonder wil ik de studenten uit de studierichting Moleculaire Wetenschappen bedanken voor de prettige verstandhouding. Ik beschouw het als een grote eer dat zij mij het afgelopen jaar hebben verkozen tot hun 'docent van het jaar'.

Graag wil ik ook Bea en Erwin van Wonderen-Tettelaar van harte danken. Onze vriendschap en verbondenheid in geloof zijn mij ook bij het werk steeds tot grote steun geweest. Ik ben dankbaar dat zij mijn paranimfen willen zijn. Verder wil ik Pieter van Kampen en Henk Geertsema bedanken voor het doorspreken van een aantal van de stellingen.

Bij dit alles besef ik hoe betrekkelijk de kennis uit dit proefschrift is. Graag verwijs ik naar mijn beide laatste stellingen. Het is belangrijker, met de woorden van de apostel Paulus, de liefde van Christus te kennen, die alle kennis te boven gaat (Efeziërs 3:19a).

Wageningen, 27 december 1995

Abstract

In this thesis exit problems are considered for stochastic dynamical systems with small random fluctuations. We study exit from a domain in the state space through a boundary, or a specified part of the boundary, that is unattainable in the underlying deterministic system. We analyze diffusion approximations of the dynamical systems. The processes are described with a Fokker-Planck equation in a continuous state space. Taking the diffusion parameter as the small parameter, we determine asymptotic expressions for the probability of exit and the (conditional) expected exit time.

We consider applications in groundwater flow and epidemiology. For a contaminant in an advective-dispersive groundwater flow asymptotic expressions are derived for the probability of arrival at a well and the expected arrival time. For a stochastic *SIR*-model describing the spread of an infectious disease in a population we determine asymptotic expressions for the following quantities: the probability that a major outbreak occurs upon the introduction of the disease into the population, the probability of extinction of the disease at the end of a major outbreak, and the expected extinction time of the disease for an initial state in the stable equilibrium. Finally, for an interval in a one-dimensional stochastic system we study the expected exit time at precisely that end of the interval where exit is not likely, including in our analysis initial states outside a boundary layer.

Contents

Chapter 1

Introduction

A large number of phenomena in nature can be modelled as a dynamical system. Frequently, an essential feature of these phenomena is formed by stochastic fluctuations. In that case we incorporate these fluctuations in the model and study a stochastic system. The stochastic system may explain features of phenomena that the deterministic system can not account for. In this thesis we study stochastic dynamical systems in which the random fluctuations are small.

In each chapter attention is focused on exit problems. A domain in the state space of the dynamical system is considered and exit from this domain through the boundary is studied. In particular, we are interested in exit through a boundary, or a specified part of the boundary, that is unattainable in the underlying deterministic system. For these 'unexpected exits' the probability of exit and the (conditional) expected exit time form the main points of interest.

In this study we analyze diffusion approximations of the dynamical systems. The processes are described with a Fokker-Planck (or Kolmogorov) equation in a continuous state space. The first order derivatives in this differential equation are related to the deterministic motion. The second order derivatives correspond to the (small) random perturbations. The probability of exit through a specified part of the boundary of a domain in state space, and the (conditional) expected exit time, are determined by Dirichlet problems for the backward Fokker-Planck equation. Using the theory of singular perturbations, see, e.g., Eckhaus (1973), Kevorkian and Cole (1981) and O'Malley (1991), we asymptotically solve these Dirichlet problems, taking the diffusion parameter as the small parameter. The asymptotic solutions are compared with results obtained by random walk simulations.

The research of this thesis can be seen as an extension of and a sequel to the research of exit problems performed in the last twenty years, see, e.g., Ludwig (1975), Schuss (1980), Gardiner (1983) and Roozen (1990). Chapters 2, 3 and 4 broaden the field of applications. Chapter 5 brings a new element in the more theoretical aspects of exit problems. It opens a field of future research that is expected to be interesting both for its mathematical aspects and its applications in other sciences.

We now give an outline of the following chapters. In chapter 2 a study is made of the transport of pollution in groundwater. It is not sufficient to model the transport of particles by advection only. In addition macroscopic dispersion has to be taken into consideration. This accounts for the random motion of individual particles in the flow. Because of dispersion pollution may enter a region that is unattainable for the advective flow only, e.g., a region with flow towards a well. We formulate and asymptotically solve Dirichlet problems for the probability of arrival and the expected arrival time of contaminated particles at a well in an arbitrary background flow. Since dispersion contributes considerably less to the displacement of a particle than advection, the analysis starts with the advective flow pattern. In addition inside a boundary layer dispersion is taken into account. Using the boundary layer solution and the advective travel time we also construct a composite expansion for the expected arrival time that is valid in the region of advective flow towards the well. We note that this study can be extended to include loss of contamination by, for example, adsorption or radioactive decay, see Van Kooten (1994). It is also noted that chapter 2 contains a derivation of the Dirichlet problem for the (conditional) expected exit time through a specified part of the boundary of a domain in a more dimensional state space.

This mathematical analysis of exit problems can be used in a broad field of applications. In chapters 3 and 4 we study a two-dimensional stochastic system modelling the spread of an infectious disease. In this model we consider a population that is divided in three classes: susceptibles, infectives and removed. The population is renewed at a constant rate. We study the case where in the underlying deterministic system the disease becomes endemic. In the stochastic system the disease can disappear from the population because of stochastic fluctuations. In chapter 3 we first study the probability that a major outbreak of the disease does not occur upon the entry of one or a few infectives into the population. This probability is determined by formulating and asymptotically solving an exit problem. In the asymptotic analysis it is assumed that $1/N$ is a small parameter, where N is the size of the population when the disease is absent. We also study the expected extinction time of the disease given that it has become endemic. The asymptotic solution of the Dirichlet problem for this expected extinction time contains one unknown constant. To determine this constant the WKB-method is applied to the forward Fokker-Planck equation, the resulting ray equations are numerically solved, and use is made of the divergence theorem.

In chapter 4 we study for this epidemiological model the following question: given that a major outbreak occurs upon the entry of one or a few infectives into the population, what is the probability that the disease will disappear from the population (directly) at the end of the major outbreak? Of particular importance in answering this question of epidemic fade-out is the rate at which the population is renewed. For a large renewal rate it is likely that the disease will become endemic after the major outbreak. If the renewal rate is very small, the disease will die out at the end of the major outbreak with probability close to one. There is a range of renewal rate values such that the extinction probability varies from close to zero to close to one. In this chapter we deal with this transitional case. We first derive local asymptotic expansions for the trajectories of the underlying deterministic system, in particular for the deterministic trajectory starting in the saddle point. For small renewal rate values the susceptible population is restored only slowly after the major outbreak. At this stage of the process diffusion plays an important role and extinction of the disease can occur in the stochastic system. Using an approximation of the backward Fokker-Planck equation we formulate a boundary value problem for the probability that the disease dies out during this part of the process. From the solution we obtain an expression for the probability of extinction of the disease at the end of the major outbreak.

After this analysis of exit problems in several, quite different applications, we address a more theoretical aspect of exit problems in chapter 5. This aspect came up in studying the expected arrival time of pollution at a well in a dispersive groundwater flow, see chapter 2. For starting points outside the region of advective flow towards the well the asymptotic solution for the expected arrival time derived in that chapter is restricted to starting points inside a boundary layer along the separating streamline. Outside the boundary layer more refined asymptotic methods are required to approximate the small probability that pollution enters the well and the expected arrival time. The study of chapter 5 can be seen as a first step to approach this kind of problems. For an interval in a one-dimensional state space exit is studied at precisely that boundary where exit is not likely. Three fundamental cases are considered, the interval containing, respectively, no equilibrium, an unstable equilibrium and a stable equilibrium. It is noted that the asymptotic expressions reveal essential features of the expected exit time that remain hidden in the exact solutions.

Finally, we note that chapters 2, 3, 4 and 5 have originally been written as independent papers. They are, therefore, self-contained and can be read independently.

References

Eckhaus, W. (1973), *Matched Asymptotic Expansions and Singular Perturbations*, North Holland, Amsterdam.

Gardiner, C.W. (1983), *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, Springer-Verlag, Berlin.

Kevorkian, J. and J.D. Cole (1981), *Perturbation Methods in Applied Mathematics*, Springer-Verlag, New York.

Ludwig, D. (1975), *Persistence of dynamical systems under random perturbations*, SIAM Rev., 17, pp. 605-640.

O'Malley, R.E. (1991), *Singular Perturbation Methods for Ordinary Differential Equations*, Springer-Verlag, New York.

Roozen, H. (1990), *Analysis of the exit problem for randomly perturbed dynamical systems in applications*, PhD Thesis, Wageningen.

Schuss, Z. (1980), *Theory and Applications of Stochastic Differential Equations*, Wiley, New York.

Van Kooten, J.J.A. (1994), *Groundwater contaminant transport including adsorption and first order decay*, Stoch. Hydrol. Hydraulics, 8, pp. 185-205.

Chapter 2

Spread of pollution by dispersive groundwater flow[1]

Abstract

A study is made of the transport of pollution in groundwater. The probability that in groundwater a contaminated particle crosses the boundary of a protected zone is computed by solving asymptotically the Dirichlet problem for the backward Kolmogorov equation describing the random motion of the particle. The randomness in the displacement of the particle is due to the dispersive properties of the flow. The expected arrival time at the boundary is computed from a corresponding nonhomogeneous Dirichlet problem.

1. Introduction

In this study we consider the flow of groundwater which is confined in a layer, called aquifer. The thickness of this layer is assumed to be small, so the transport will be modelled as a 2D-flow problem. For the study of groundwater pollution it is not sufficient to consider the transport of particles by advection only. We also have to take into consideration the mechanism of macroscopic dispersion. This accounts for the random motion of individual particles in the flow. We assume the dispersion to be proportional to the velocity with coefficients a_L in the longitudinal direction and a_T in the transversal direction, see Bear and Verruijt (1987). In general the value of a_T is smaller than the value of a_L.

A situation typical for the discharge of groundwater from the aquifer is shown in Figure 1. The region of advective flow towards the well is bounded by two separating streamlines which end in a stagnation point. When the groundwater is polluted at a site outside the domain of advective flow to the well and far away

[1]Published in: SIAM Journal of Applied Mathematics, 54 (1994), pp. 26-41

from the separating streamlines, the pollution will not reach the well, because the advective component of the flow is much larger than the dispersive component. When the site of pollution is just outside this stream domain, contaminated particles may enter the well because of the dispersiveness of the flow. The probability for a particle to reach the well changes rapidly from close to 0 to close to 1 for starting points within a boundary layer along a separating streamline. The asymptotic method we present gives detailed results on this rapid change.

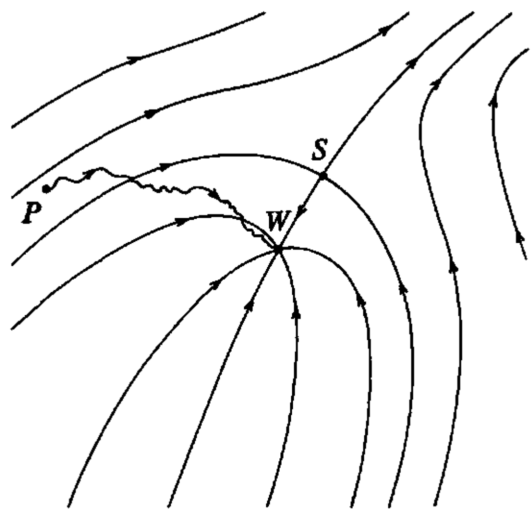


Figure 1. *By dispersion a contaminated particle released in P may enter the well W. The region of advective flow towards the well is bounded by two separating streamlines ending in the stagnation point S.*

Our analysis of the dispersion problem starts from the advection-dispersion equation which can be seen as the Fokker-Planck equation corresponding with the random motion of a particle, see Uffink (1989). The advection-dispersion equation can be obtained, see Bear and Verruijt (1987), by constructing a macroscopic mass balance equation and considering an average flux in which are included the following macroscopic fluxes: an advective flux and a dispersive flux. It is assumed that the dispersive flux can be expressed as a Fickian type law. The advective flux is based on the average velocity of the flow; Bear and Bachmat (1990) determine an equation of motion for this velocity by averaging the microscopic momentum balance equation for a Newtonian incompressible fluid (i.e., Navier-Stokes equation) filling the pores. When the advection-dispersion equation is used to model solute transport at larger scales, the dispersion arises from heterogeneities of the hydraulic conductivity at a scale which may be large with respect to the macroscopic scale, but which is small compared to the geometry of the region that is considered, see Figure 1. This assumption is not realistic in all situations, see e.g. Dagan (1989). The expression (2.2) for the elements of the dispersion matrix is obtained from Bear and Verruijt (1987) for an isotropic porous medium.

Because dispersion contributes considerably less to the displacement of a particle than advection, the dispersion problem is approached by perturbation techniques. The advective flow yields a first order approximation. Inside the boundary layer a second order approximation is computed by taking into account the dispersion. The objective of this study is to compute the probability that contaminated particles cross a boundary, e.g., are pumped up at a well. Moreover, the expected time of arrival at the boundary is approximated asymptotically. Compared with Van Herwaarden and Grasman (1991) the present paper yields an improved asymptotic approximation of the problem. In particular, in the asymptotic formula for the arrival time with leading order term of $O(-\ln a_T)$ it is important to include the order unity contribution.

2. The Fokker-Planck equation

Let $p(x, t)$ be the probability density function to find a particle at a point x at time t. Then the function $p(x, t)$ satisfies the forward Kolmogorov of Fokker-Planck equation

$$\frac{\partial p}{\partial t} = Mp, \tag{2.1a}$$

$$M = -\frac{\partial}{\partial x_i}(v_i\,.) + \frac{\partial}{\partial x_i}(D_{ij}\frac{\partial\,.}{\partial x_j}), \tag{2.1b}$$

where v is the velocity vector and D is the dispersion matrix. With respect to the indices the common summation convention is used. The entries of the symmetric matrix D are given by

$$D_{ij} = a_T|v|\delta_{ij} + (a_L - a_T)v_i v_j/|v|. \tag{2.2}$$

Related to the forward operator M is the backward operator

$$L = v_i\frac{\partial\,.}{\partial x_i} + \frac{\partial}{\partial x_i}(D_{ij}\frac{\partial\,.}{\partial x_j}). \tag{2.3}$$

L is the formal adjoint of the forward operator: $L = M^*$. It plays an important role in exit problems. Let us consider a domain Ω with the boundary $\partial\Omega$ composed of two parts $\partial\Omega_0$ and $\partial\Omega_1$. We are interested in the probability $u(x)$ that a particle released in $x \in \Omega$ reaches $\partial\Omega$ the first time at the part $\partial\Omega_1$. This function $u(x)$ satisfies the Dirichlet problem

$$Lu = 0 \quad \text{in} \quad \Omega, \tag{2.4a}$$

$$u = 0 \quad \text{at} \quad \partial\Omega_0, \qquad u = 1 \quad \text{at} \quad \partial\Omega_1, \tag{2.4b}$$

see, e.g., Gardiner (1983). For a particle starting in $x \in \Omega$ we are also interested in the expected arrival time $T_1(x)$ at $\partial\Omega$ with the condition that exit from Ω takes place at $\partial\Omega_1$. In order to solve this problem we make use of the probability density function $r(\bar{x}, x)$ of leaving Ω at $\bar{x} \in \partial\Omega$ and the expected arrival time $\tau(\bar{x}, x)$ at $\partial\Omega$ conditional that exit occurs at $\bar{x} \in \partial\Omega$ for a starting point $x \in \Omega$. For these functions it is shown in Gardiner (1983) that

$$L(\tau r) = -r \quad \text{in} \quad \Omega, \tag{2.5a}$$

$$\tau r = 0 \quad \text{at} \quad \partial\Omega. \tag{2.5b}$$

Defining

$$T(x) = \int_{\partial\Omega_1} \tau(\bar{x}, x) r(\bar{x}, x)\, dS \tag{2.6}$$

we derive from (2.5) the nonhomogeneous Dirichlet problem

$$LT = -u \quad \text{in} \quad \Omega, \tag{2.7a}$$

$$T = 0 \quad \text{at} \quad \partial\Omega. \tag{2.7b}$$

For the expected arrival time $T_1(x)$ at $\partial\Omega$ with the condition that exit from Ω takes place at $\partial\Omega_1$ we have

$$T_1(x) = \int_{\partial\Omega_1} \tau(\bar{x}, x) r(\bar{x}, x)/u(x)\, dS, \tag{2.8}$$

so we find

$$T_1(x) = T(x)/u(x) \tag{2.9}$$

with $T(x)$ satisfying (2.7) and $u(x)$ satisfying (2.4). It is noted that in other publications (Mangel (1979)) the term $u(x)$ is not present in (2.9).

3. Symmetric flow near a stagnation point

As a starting point we consider the symmetric flow field with velocity vector

$$v(x, y) = (-x, y), \tag{3.1}$$

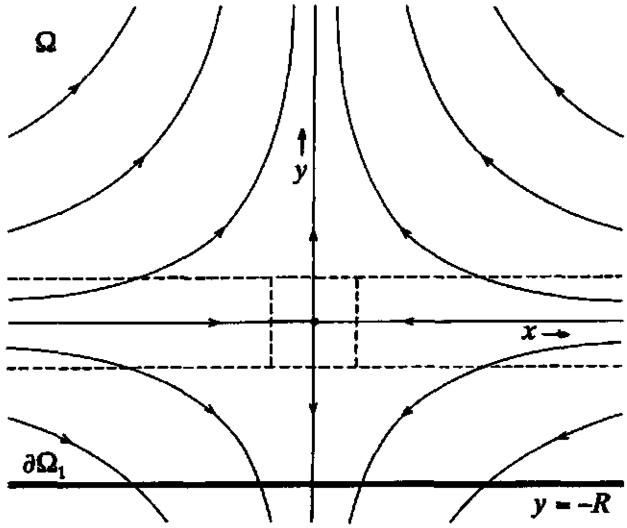


Figure 2. *Symmetric flow near a stagnation point. The probability of arrival and the expected arrival time at the boundary $\partial\Omega_1$ are computed.*

which has (0, 0) as stagnation point, and the domain

$$\Omega = \{(x, y) \mid y > -R\} \tag{3.2}$$

with $R > 0$ and $y = -R$ away from $y = 0$, see Figure 2. As a result of dispersion a contaminated particle released in a point (x, y) with $y > 0$ may cross the separating streamline $y = 0$ and reach the boundary $\partial\Omega_1$: $y = -R$. The probability $u(x, y)$ that a particle starting in $(x, y) \in \Omega$ reaches $y = -R$ satisfies the differential equation

$$-x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} + \frac{\partial}{\partial x}(D_{xx}\frac{\partial u}{\partial x}) + \frac{\partial}{\partial x}(D_{xy}\frac{\partial u}{\partial y}) + \frac{\partial}{\partial y}(D_{xy}\frac{\partial u}{\partial x}) + \frac{\partial}{\partial y}(D_{yy}\frac{\partial u}{\partial y}) = 0 \tag{3.3a}$$

with boundary conditions

$$u(x, -R) = 1, \qquad \lim_{y\to\infty} u(x, y) = 0. \tag{3.3b}$$

The dispersion terms are given by

$$D_{xx} = a_T|v| + (a_L - a_T)x^2/|v|, \tag{3.4a}$$

$$D_{xy} = -(a_L - a_T)xy/|v|, \tag{3.4b}$$

$$D_{yy} = a_T|v| + (a_L - a_T)y^2/|v|. \tag{3.4c}$$

Since the dispersive component of the flow is of a smaller magnitude than the advective component, we expect u to change from about 0 to 1 in a small region along the x-axis. Therefore, we assume a boundary layer to be present at this place. Outside this boundary layer we may neglect the dispersion and approximate (3.3a) by its advective part

$$-x\frac{\partial \bar{u}}{\partial x} + y\frac{\partial \bar{u}}{\partial y} = 0. \tag{3.5}$$

For a particle released in a point (x, y) in the upper half plane outside the boundary layer this yields the approximate solution $\bar{u} = 0$ and for a particle starting in the lower half plane outside the boundary layer $\bar{u} = 1$.

To solve the dispersion problem for a particle released in a point (x, y) inside the boundary layer we apply a coordinate stretching procedure (Kevorkian and Cole (1981))

$$\eta = y/\sqrt{a_T}. \tag{3.6}$$

We will analyze this boundary layer problem for $x > 0$. The case $x < 0$ is treated in the same way. In our analysis we exclude a small neighbourhood of the origin where another approximation has to be made. Its outcome has only very local importance and is therefore not analyzed. Substitution of (3.6) in (3.3a) yields for $a_T, a_L \to 0$

$$-x\frac{\partial \bar{u}}{\partial x} + \eta\frac{\partial \bar{u}}{\partial \eta} + x\frac{\partial^2 \bar{u}}{\partial \eta^2} = 0. \tag{3.7}$$

Matching of the function $\bar{u}(x, \eta)$ to the outer solutions $\bar{u} = 0$ and $\bar{u} = 1$ along the characteristics xy = constant of (3.5) leads to the solution

$$\bar{u}(x, \eta) = \frac{1}{\sqrt{2\pi}} \int_{\eta/\sqrt{(2x/3)}}^{\infty} \exp[-t^2/2]\, dt. \tag{3.8}$$

We will now solve asymptotically problem (2.7) for the expected arrival time at the boundary $y = -R$. For starting points (x, y) in the upper half plane outside the boundary layer we may approximate (2.7) by

$$-x\frac{\partial T}{\partial x} + y\frac{\partial T}{\partial y} = 0, \tag{3.9a}$$

$$\lim_{y\to\infty} T(x, y) = 0, \tag{3.9b}$$

where the boundary condition is obtained by applying (2.7b) to a boundary $y = \bar{R}$ and letting $\bar{R} \to \infty$. Eqs. (3.9) are satisfied by

$$T = 0. \tag{3.10}$$

For particles starting in the lower half plane outside the boundary layer we have the approximation

$$-x\frac{\partial T}{\partial x} + y\frac{\partial T}{\partial y} = -1, \tag{3.11a}$$

$$T(x, -R) = 0. \tag{3.11b}$$

This is satisfied by

$$T(x, y) = -\ln(-y) + \ln R, \tag{3.12}$$

which equals the advective travel time to the boundary $y = -R$. Inside the boundary layer T is computed from

$$-x\frac{\partial T}{\partial x} + \eta\frac{\partial T}{\partial \eta} + x\frac{\partial^2 T}{\partial \eta^2} = -\bar{u}(x, \eta), \tag{3.13a}$$

where T must match with the outer solutions (3.10) and (3.12) along the characteristics xy = constant. This leads to the matching conditions

$$T(x, \eta) = 0 \quad \text{for} \quad \eta \to \infty \quad \text{and} \quad x\eta = c > 0, \tag{3.13b}$$

$$T(x, \eta) \sim -\ln(-\eta) - \tfrac{1}{2}\ln a_T + \ln R \quad \text{for} \quad \eta << -1 \quad \text{and} \quad x\eta = c < 0. \tag{3.13c}$$

A particular solution for (3.13a) is easily found

$$T_p(x, \eta) = \bar{u}(x, \eta) \ln x. \tag{3.14}$$

Setting

$$T(x, \eta) = T_p(x, \eta) + T_h(x, \eta) \tag{3.15}$$

and introducing new coordinates

$$\tau = \tfrac{1}{3}x^3 \quad \text{and} \quad \zeta = x\eta, \tag{3.16}$$

we obtain the following homogeneous initial value problem for the function $T_h(\tau, \zeta)$

$$\frac{\partial T_h}{\partial \tau} = \frac{\partial^2 T_h}{\partial \zeta^2}, \tag{3.17a}$$

$$T_h(0, \zeta) = 0 \quad \text{for} \quad \zeta > 0, \tag{3.17b}$$

$$T_h(0, \zeta) = -\tfrac{1}{2}\ln a_T + \ln R - \ln(-\zeta) \quad \text{for} \quad \zeta < 0. \tag{3.17c}$$

This is satisfied by

$$T_h(\tau, \zeta) = \frac{1}{\sqrt{2\pi}} \int_{\zeta/\sqrt{(2\tau)}}^{\infty} (-\tfrac{1}{2}\ln a_T + \ln R - \ln(-\zeta + t\sqrt{2\tau}\,))\, \exp[-t^2/2]\, dt. \tag{3.18}$$

Finally, returning to x, η-coordinates and using (2.9), we find the expected arrival time for starting points inside the boundary layer

$$T_{\text{bound}}(x, \eta) = -\tfrac{1}{2}\ln a_T + \ln R + \tag{3.19}$$

$$- \int_{\eta/\sqrt{(2x/3)}}^{\infty} \ln(-\eta + t\sqrt{2x/3})\exp[-t^2/2]\, dt \Big/ \int_{\eta/\sqrt{(2x/3)}}^{\infty} \exp[-t^2/2]\, dt.$$

We remark that this solution substantially improves the result derived in Van Herwaarden and Grasman (1991), where in the matching procedure not all terms of order unity have been taken into account.

4. Arbitrary flow with a stagnation point

In this section we generalize the method used in the previous example for an arbitrary flow near a separating streamline ending in a stagnation point. The stagnation point is an interior point of the domain Ω. We assume the flow to be free of sources and sinks and irrotational in Ω, for which the streamlines leading towards and away from the stagnation point are perpendicular. For example, in the case of discharge of water from a well we may take Ω as depicted in Figure 3a. The well is excluded from Ω by a circular domain with boundary $\partial\Omega_1$, away from the separating streamlines. (For $\partial\Omega_1$ the point denoting the position of the well may be taken.) The solution $u(x, y)$ of the Dirichlet problem (2.4) equals the probability that a particle released at $(x, y) \in \Omega$ reaches $\partial\Omega$ the first time at the part $\partial\Omega_1$. If the distance of $\partial\Omega_0$ to the well is sufficiently large and (x, y) is not near $\partial\Omega_0$, then $u(x, y)$ also equals the probability that a particle released at (x, y) ends in the well.

4.1 Probability of arrival at $\partial\Omega_1$

To analyze the probability u that a particle released at a point near a separating streamline ending in a stagnation point reaches $\partial\Omega_1$, we introduce new coordinates ρ and ν. The coordinate $\rho > 0$ is the coordinate along the separating streamline and ν the one perpendicular to it, see Figure 3b. The stagnation point is in $(\rho, \nu) = (0, 0)$. Let the velocity vector near the separating streamline be given in new coordinates by

$$(v(\rho, \nu), w(\rho, \nu)). \tag{4.1}$$

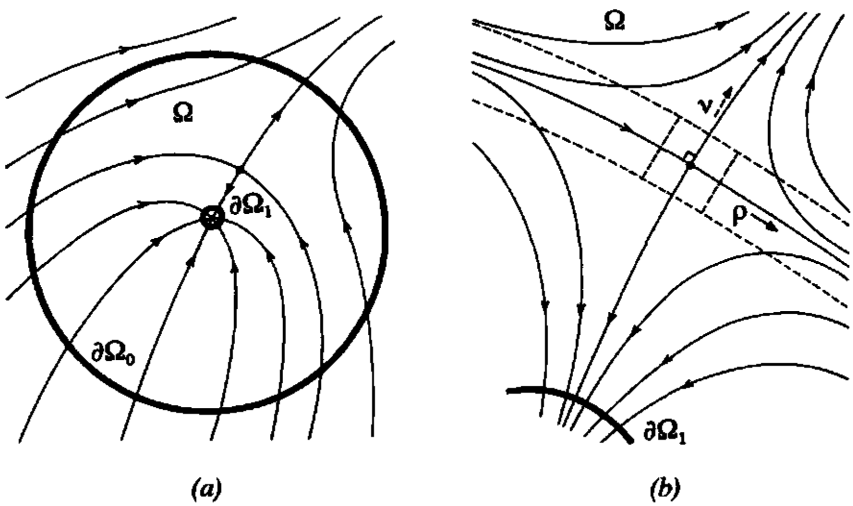


Figure 3. *Flow patterns near a well and a stagnation point. (a) Example of a flow field with a stagnation point in the case of discharge of water from a well. (b) Coordinates* ρ *and* ν *along and perpendicular to the separating streamline, used for starting points near this streamline.*

Along the separating streamline we assume a boundary layer of width $O(\sqrt{a_T})$. For starting points inside this boundary layer we introduce the local coordinate

$$\eta = \nu/\sqrt{a_T}. \tag{4.2}$$

Switching from the coordinates ρ and ν to ρ and η we obtain from (2.4) after letting a_T, $a_L \to 0$ the asymptotic approximation

$$v(\rho,\ 0)\frac{\partial \bar{u}}{\partial \rho} + w_\nu(\rho,\ 0)\eta\frac{\partial \bar{u}}{\partial \eta} - v(\rho,\ 0)\frac{\partial^2 \bar{u}}{\partial \eta^2} = 0 \tag{4.3}$$

with $\bar{u}(\rho, \eta)$ to be matched with the outer solutions $\bar{u} = 0$ and $\bar{u} = 1$ along the characteristics of the advective equation

$$v(\rho,\ \nu)\frac{\partial \bar{u}}{\partial \rho} + w(\rho,\ \nu)\frac{\partial \bar{u}}{\partial \nu} = 0. \tag{4.4}$$

Making use of the relation $w_\nu = -v_\rho$ at $\nu = 0$ for a flow that is free of sources and sinks we find the solution

$$\bar{u}(\rho,\ \eta) = \frac{1}{\sqrt{2\pi}}\int_{\eta q(\rho)}^{\infty} \exp[-t^2/2]\, dt \tag{4.5a}$$

with

$$q(\rho) = \left(2\int_0^\rho v(\tilde{\rho},\, 0)^2 / v(\rho,\, 0)^2 d\tilde{\rho}\right)^{-1/2}. \tag{4.5b}$$

It is noted here that the problem of crossing a separatrix for starting points in its neighbourhood is also considered by Mangel and Ludwig (1977) and by Mangel (1979), who obtain comparable results for the probability of arrival by constructing a formal asymptotic series solution suggested by the analysis of a one-dimensional version.

4.2 The expected arrival time at $\partial\Omega_1$

Starting points outside the boundary layer. To find the expected arrival time at $\partial\Omega_1$ we solve problem (2.7) asymptotically. For starting points in Ω outside the boundary layer and not at the same side of the separating streamline as $\partial\Omega_1$ we may neglect the dispersion terms. Approximating u by $\bar{u} = 0$ we obtain

$$T = 0. \tag{4.6}$$

For particles starting in Ω outside the boundary layer and at the same side of the separating streamline as $\partial\Omega_1$ we may neglect the dispersion terms and approximate u by $\bar{u} = 1$, which yields the advective travel time T_{adv} to the boundary $\partial\Omega_1$ as solution of (2.7). In particular, we are interested in the behaviour of T_{adv} near the boundary layer. An expression for this behaviour will be needed for matching purposes. To find this expression we will analyze $T_{adv}(\rho, \nu)$ for ν close to 0. In the analysis we also make use of new coordinates σ and μ near the streamline leading away from the stagnation point in the direction of $\partial\Omega_1$, see Figure 4. The coordinate $\sigma < 0$ is the coordinate along this streamline and μ the one perpendicular to it. The stagnation point is in $(\sigma, \mu) = (0, 0)$ and the boundary $\partial\Omega_1$ and this streamline intersect in the point $(\sigma, \mu) = (\sigma_b, 0)$. In these coordinates the velocity vector near this streamline is given by

$$(r(\sigma,\, \mu),\, s(\sigma,\, \mu)). \tag{4.7}$$

We now divide $T_{adv}(\rho, \nu)$ into three parts

$$T_{adv} = T_I + T_{II} + T_{III}, \tag{4.8}$$

where T_I is the travel time close to the separating streamline, T_{II} the travel time near the stagnation point and T_{III} the travel time near the streamline leading away from the stagnation point, see Figure 4, where δ and ε are close to 0. For $-1 << \nu < 0$ and ρ fixed and letting $\delta \downarrow 0$, $\varepsilon \uparrow 0$ we then obtain

$$T_{adv}(\rho,\, \nu) \sim \frac{1}{v_\rho(0,\, 0)} \ln(-\nu) + T_{reg}(\rho,\, \sigma_b) \tag{4.9a}$$

with

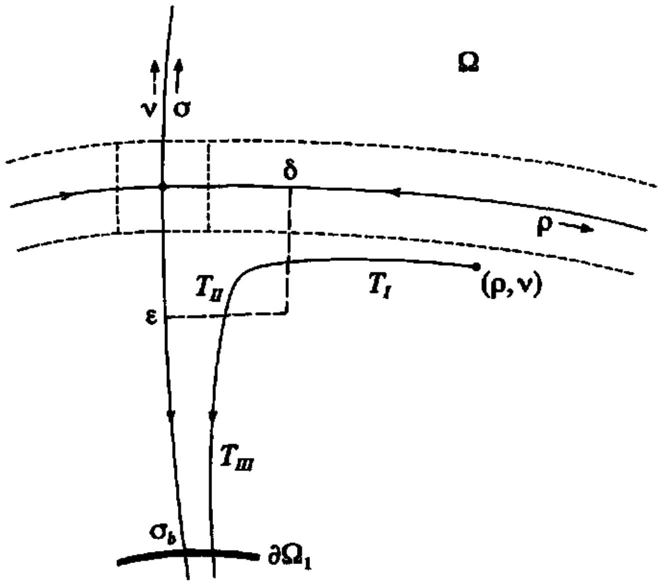


Figure 4. *The contributions to the advective travel time $T_{adv}(\rho, \nu) = T_I + T_{II} + T_{III}$ for small values of* ν, δ *and* ε.

$$T_{reg}(\rho, \sigma_b) = \frac{1}{v_\rho(0,\ 0)} \ln \frac{v(\rho,\ 0)}{v_\rho(0,\ 0)\rho.-\sigma_b} + \int_\rho^0 \frac{1}{v(\tilde{\rho},\ 0)} - \frac{1}{v_\rho(0,\ 0)\tilde{\rho}}\, d\tilde{\rho} +$$

$$+ \int_0^{\sigma_b} \frac{1}{r(\tilde{\sigma},\ 0)} - \frac{1}{r_\sigma(0,\ 0)\tilde{\sigma}}\, d\tilde{\sigma}. \tag{4.9b}$$

The boundary layer solution. Inside the boundary layer T is computed from

$$v(\rho,\ 0)\frac{\partial T}{\partial \rho} + w_v(\rho,\ 0)\eta\frac{\partial T}{\partial \eta} - v(\rho,\ 0)\frac{\partial^2 T}{\partial \eta^2} = -\bar{u}(\rho,\ \eta), \tag{4.10a}$$

where T is matched with the outer solutions (4.6) and (4.9a) along the characteristics $-\eta v(\rho, 0)$ = constant of the advective equation. This leads to the conditions

$$T(\rho,\ \eta) = 0 \quad \text{for } \eta \to \infty \quad \text{and} \quad -\eta v(\rho,\ 0) = c > 0, \tag{4.10b}$$

$$T(\rho,\ \eta) \sim \frac{1}{v_\rho(0,\ 0)} \ln(-\eta\sqrt{a_T}) + T_{reg}(\rho,\ \sigma_b) \quad \text{for } \eta << -1 \text{ and } -\eta v(\rho,\ 0) = c < 0. \tag{4.10c}$$

To solve this problem we can make use of the particular solution of (4.10a)

$$T_p(\rho, \eta) = -\int_\gamma^\rho \frac{1}{\nu(\tilde{\rho}, 0)} d\tilde{\rho} . \bar{u}(\rho, \eta), \tag{4.11}$$

where γ is an integration constant. Setting

$$T(\rho, \eta) = T_p(\rho, \eta) + T_h(\rho, \eta) \tag{4.12}$$

we obtain a homogeneous problem for $T_h(\rho, \eta)$. Introduction of the new coordinates

$$\tau = \int_0^\rho \nu(\tilde{\rho}, 0)^2 d\tilde{\rho}, \tag{4.13a}$$

$$\zeta = -\eta\nu(\rho, 0) \tag{4.13b}$$

yields for $T_h(\tau, \zeta)$ the following initial value problem

$$\frac{\partial T_h}{\partial \tau} = \frac{\partial^2 T_h}{\partial \zeta^2}, \tag{4.14a}$$

$$T_h(0, \zeta) = 0 \quad \text{for} \quad \zeta > 0, \tag{4.14b}$$

$$T_h(0, \zeta) = g(\zeta) \quad \text{for} \quad \zeta < 0 \tag{4.14c}$$

with

$$g(\zeta) = \frac{1}{\nu_\rho(0, 0)} \ln \frac{\zeta\sqrt{a_T}}{\nu_\rho(0, 0)\gamma.-\sigma_b} + \int_\gamma^0 \frac{1}{\nu(\tilde{\rho}, 0)} - \frac{1}{\nu_\rho(0, 0)\tilde{\rho}} d\tilde{\rho} +$$

$$+ \int_0^{\sigma_b} \frac{1}{r(\tilde{\sigma}, 0)} - \frac{1}{r_\sigma(0, 0)\tilde{\sigma}} d\tilde{\sigma}. \tag{4.14d}$$

This is satisfied by

$$T_h(\tau, \zeta) = \frac{1}{\sqrt{2\pi}} \int_{\zeta/\sqrt{(2\tau)}}^\infty g(\zeta - t\sqrt{2\tau}\,) \exp[-t^2/2]\, dt. \tag{4.15}$$

Bringing this solution in ρ, η-coordinates and using (4.12) we find the solution for $T(\rho, \eta)$. It is noted that by adding $T_p(\rho, \eta)$ the integration constant γ is removed. Applying (2.9) we finally obtain the boundary layer solution for the expected arrival time

$$T_{\text{bound}}(\rho, \eta) = \frac{1}{\nu_\rho(0, 0)} \ln(\sqrt{a_T}) + T_{\text{reg}}(\rho, \sigma_b) +$$

$$+ \frac{1}{v_\rho(0,\ 0)} \int_{\eta q(\rho)}^{\infty} \ln(-\eta + t/q(\rho)) \exp[-t^2/2]\, dt \Big/ \int_{\eta q(\rho)}^{\infty} \exp[-t^2/2]\, dt \qquad (4.16)$$

with $T_{reg}(\rho, \sigma_b)$ given by (4.9b) and $q(\rho)$ by (4.5b). We remark that this solution only depends on the point of intersection of the boundary $\partial\Omega_1$ with the streamline leading away from the stagnation point, and not on the particular shape of $\partial\Omega_1$, see σ_b in (4.16). This may be explained intuitively by noting that starting points inside the boundary layer are situated on streamlines that approach very closely the streamline leading away from the stagnation point. It is expected that in higher order approximations more properties of the boundary $\partial\Omega_1$ are contained.

A composite expansion of the expected arrival time. For particles released in Ω in the region of advective flow towards $\partial\Omega_1$ we now have found two approximations for the expected arrival time at $\partial\Omega_1$. For starting points inside the boundary layer we have the approximation T_{bound} given by (4.16). For particles starting outside the boundary layer we may use the approximation T_{adv}, that is the advective travel time. To avoid the difficulty of deciding whether a point is within or outside the boundary layer, we construct a composite expansion which is valid both inside and outside the boundary layer, see Van Dyke (1975),

$$T_{comp} = T_{adv} + T_{bound} - T_{match}, \qquad (4.17)$$

where T_{match} is given by the right hand side of Eq. (4.9a). Inside the boundary layer T_{match} cancels T_{adv}, so T_{comp} reduces to T_{bound}; outside the boundary layer T_{match} compensates T_{bound}, which results in $T_{comp} \sim T_{adv}$.

5. A well in a uniform background flow

As an example the method developed in the foregoing section will be applied to analyze the pollution of a well in a uniform 2D-background flow. The complex potential for the flow field is

$$\omega(z) = z - \ln(z + 1), \qquad z = x + iy. \qquad (5.1)$$

The velocity vector of the background flow is (1, 0). The well is located in (−1, 0). For the velocity components one can derive

$$v_x = \frac{dx}{dt} = \frac{x^2 + x + y^2}{(1 + x)^2 + y^2}, \qquad (5.2a)$$

$$v_y = \frac{dy}{dt} = \frac{-y}{(1 + x)^2 + y^2}. \qquad (5.2b)$$

The stagnation point is (0, 0) and the separating streamline is given by

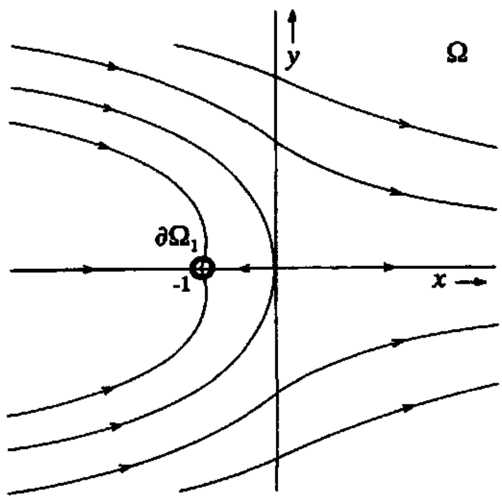


Figure 5. *Streamlines for a well in a uniform background flow. The well is excluded from Ω by a small circular domain with boundary $\partial\Omega_1$.*

$$x = -1 + \frac{y}{\tan y}. \tag{5.3}$$

For particles released inside the stream domain of the well the advective travel time to the well is given by

$$T_{adv}(x, y) = \ln\frac{y}{\sin y} - \ln\left(\frac{y}{\tan y} - x - 1\right) - x - 1, \tag{5.4}$$

see Van der Hoek (1992). We consider the region Ω, which is the x, y-plane from which we have excluded an arbitrarily small circular domain with boundary $\partial\Omega_1$ containing the well, see Figure 5. Because of the symmetry we can restrict our analysis to the upper half plane $y \geq 0$. A point (x, y) of the separating streamline corresponds in ρ, ν-coordinates with $(\rho, 0)$, where

$$\rho = \int_0^y \left(1 + \left(\frac{1}{\tan \bar{y}} - \frac{\bar{y}}{\sin^2 \bar{y}}\right)^2\right)^{1/2} d\bar{y}. \tag{5.5}$$

We first consider the probability $\bar{u}(\rho, \eta)$ that a particle starting in a point (ρ, η) of Ω ends in the well. It is given by Eqs. (4.5). Using (5.5) it is possible to express $q(\rho)$ in (4.5b) as a function of y. We obtain

$$q(\rho) = \left(2 \int_0^y Q(\tilde{y}) R(\tilde{y})\, d\tilde{y} / Q(y)\right)^{-1/2}, \tag{5.6a}$$

where

$$Q(y) = \sin^2 y \left(1 + \left(\frac{1}{y} - \frac{1}{\tan y}\right)^2\right) \tag{5.6b}$$

and

$$R(y) = \left(1 + \left(\frac{1}{\tan y} - \frac{y}{\sin^2 y}\right)^2\right)^{1/2}. \tag{5.6c}$$

For a particle starting in a point (ρ, η) of the boundary layer the expected arrival time $T_{\text{bound}}(\rho, \eta)$ at the well is given by Eq. (4.16) with $\sigma_b = -1$. It is possible to express $T_{\text{bound}}(\rho, \eta)$ as a function of y and η. We obtain

$$T_{\text{bound}}(\rho, \eta) = -\tfrac{1}{2} \ln a_T - \frac{y}{\tan y} - \tfrac{1}{2} \ln\left(1 + \left(\frac{1}{y} - \frac{1}{\tan y}\right)^2\right) +$$

$$- \int_{\eta q(\rho)}^{\infty} \ln(-\eta + t/q(\rho)) \exp[-t^2/2]\, dt \Big/ \int_{\eta q(\rho)}^{\infty} \exp[-t^2/2]\, dt. \tag{5.7}$$

This boundary layer solution may be used to construct a composite expansion T_{comp} according to (4.17), which is valid in the stream domain of the well both inside and outside the boundary layer. We obtain

$$T_{\text{comp}} = T_{\text{adv}} - \tfrac{1}{2} \ln a_T + \ln(-\nu) +$$

$$- \int_{\eta q(\rho)}^{\infty} \ln(-\eta + t/q(\rho)) \exp[-t^2/2]\, dt \Big/ \int_{\eta q(\rho)}^{\infty} \exp[-t^2/2]\, dt. \tag{5.8}$$

For the region outside the stream domain of the well we may use the boundary layer solution (5.7) as an approximation of the expected arrival time both inside and outside the boundary layer.

6. The random walk model

In this section we will compare the results of random walk simulations for the probability of arrival and the expected arrival time at a particular boundary $\partial\Omega_1$ with the analytical approximations derived in the foregoing sections. The position $(X(t), Y(t))$ of a particle for which the probability density function $p(x, y, t)$ of section 2 applies, satisfies the system of stochastic differential equations of Ito type

$$dX = (v_x + \frac{\partial}{\partial x}D_{xx} + \frac{\partial}{\partial y}D_{xy})dt + \sqrt{2a_L v}\,\frac{v_x}{v}dW_L + \sqrt{2a_T v}\,\frac{v_y}{v}dW_T, \quad (6.1a)$$

$$dY = (v_y + \frac{\partial}{\partial x}D_{xy} + \frac{\partial}{\partial y}D_{yy})dt + \sqrt{2a_L v}\,\frac{v_y}{v}dW_L - \sqrt{2a_T v}\,\frac{v_x}{v}dW_T, \quad (6.1b)$$

where $W_T(t)$ and $W_L(t)$ are independent Wiener processes. From these equations a system of stochastic difference equations can be derived. In Table 1 we give the results from $N = 1000$ simulations at different starting points $(X(0), Y(0))$ for the symmetric flow problem of section 3. These results are compared with the probabilities of arrival at the boundary $y = -R$ obtained from the asymptotic approximation (3.8) and with the expected arrival times obtained from (3.19). We remark that in this special case the composite expansion T_{comp} of (4.17) equals the boundary layer solution T_{bound} of (3.19).

In Figures 6 and 7 we present simulation results for a well in a uniform background flow. On the separating streamline the point C with coordinates $(\rho, \nu) = (2.958, 0)$ is considered, corresponding with $(x, y) = (-1.915, 2)$. For a number of starting points on the normal on the separating streamline through C a Monte Carlo simulation run ($N = 2000$) has been made. In this way an approximation has been found for the fraction of particles that reaches the well. In Figure 6 this fraction is compared with the asymptotic approximation for the probability of arrival at the well given by (4.5) and (5.6).

The simulation results for the expected arrival time at the well are presented in Figures 7. In Figure 7a they are compared with values of the boundary layer solution T_{bound} given by Eq. (5.7) and the advective travel time T_{adv} given by Eq. (5.4), which approximates the expected arrival time in the stream domain of the well outside the boundary layer. In Figure 7b these simulation results are compared with values of the composite expansion T_{comp} given by (5.8) for starting points inside the stream domain of the well, and again with values of T_{bound} outside this domain. For increasing $\nu > 0$ the simulation results are based on a decreasing number of particles that arrive at the well. Therefore the error in these results increases. It is noted that outside the boundary layer T_{adv} and T_{bound} do not differ much (see Figure 7a). This is due to the regular structure of the uniform background flow. It is expected that for irregular flow patterns this will not be the case. Then the composite formula is indispensable for making an accurate approximation.

Table 1. *The probability of arrival u and the expected arrival time T at the boundary $y = -R$ with $R = 2$ for the symmetric flow problem of section 3. The values of u_{simul} and T_{simul} have been obtained from $N = 1000$ simulations at each starting point (x, y) with $x = 4$; $a_L = .125$ and $a_T = .05$. The asymptotic approximations u_{asympt} and T_{asympt} have been computed from* (3.8) *and* (3.19).

y	u_{simul}	u_{asympt}	T_{simul}	T_{asympt}
0.4	.13	.14	2.84	2.84
0.3	.24	.21	2.63	2.73
0.2	.30	.29	2.59	2.60
0.1	.42	.39	2.44	2.47
0	.51	.50	2.35	2.32
−0.1	.62	.61	2.20	2.17
−0.2	.70	.71	2.01	2.01
−0.3	.80	.79	1.84	1.84
−0.4	.87	.86	1.66	1.67

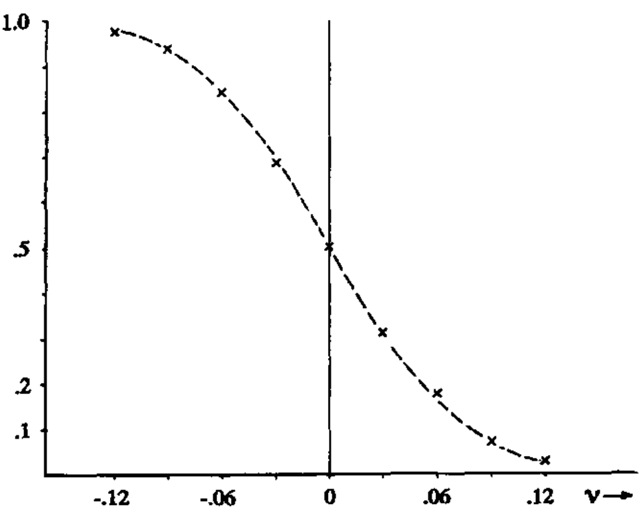
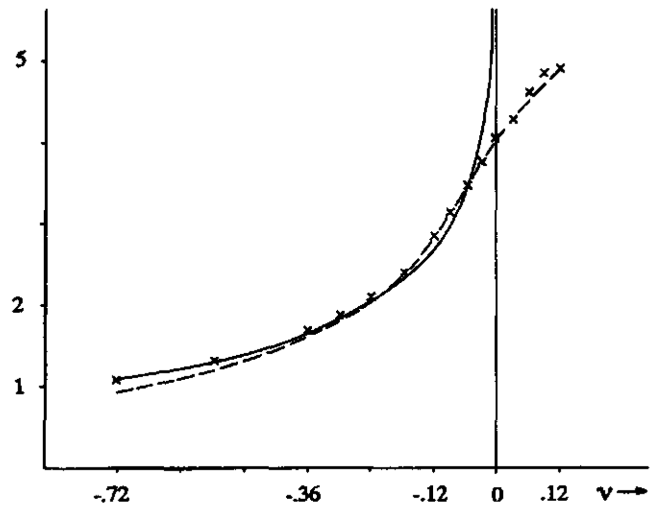
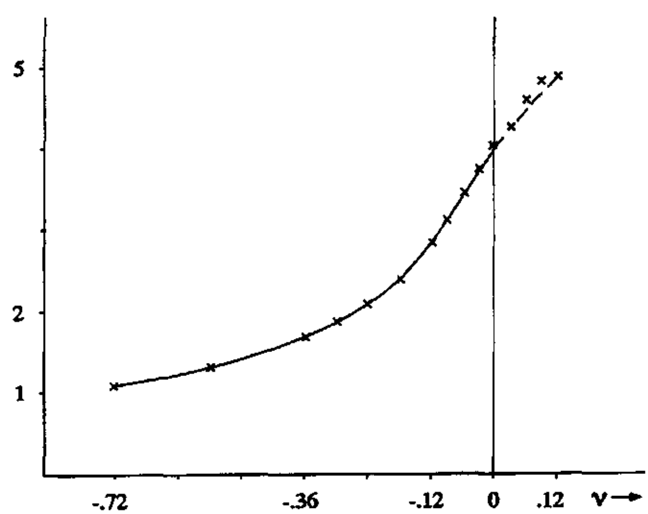


Figure 6. *The probability of arrival at a well in a uniform background flow from $N = 2000$ simulations at each point (ρ, ν) with $\rho = 2.958$ (×); $a_T = .001$ and $a_L = .01$. The asymptotic approximation (−−) has been calculated from* (4.5) *and* (5.6).



(a)



(b)

Figure 7. *The expected arrival time at a well in a uniform background flow from $N = 2000$ simulations at each point (ρ, ν) with $\rho = 2.958$; $a_T = .001$ and $a_L = .01$. Simulation results are compared with the asymptotic approximation T_{bound}, the advective travel time T_{adv} and the composite expansion T_{comp}, calculated from (5.7), (5.4) and (5.8), respectively. (a) Simulation results (×) compared with T_{adv} (—) and T_{bound} (– –). (b) Simulation results (×) compared with T_{comp} (—) and T_{bound} (– –).*

7. Conclusions

In this paper we have analyzed the dispersion problem by solving asymptotically a Fokker-Planck equation. The advective flow yields a first order approximation. In addition inside a boundary layer dispersion has been taken into account. For a flow field with a stagnation point a formula has been derived for the probability that a particle reaches a particular boundary, see Eqs. (4.5). Also for starting points inside the boundary layer an expression has been derived for the expected arrival time T_{bound} at the boundary, see Eq. (4.16). It is emphasized that this solution only depends on the point of intersection of the boundary with the streamline leading away from the stagnation point, and not on the particular shape of the boundary. The formula for T_{bound} may be used to construct a composite expansion T_{comp} which is valid in the entire region of advective flow towards the boundary, see Eq. (4.17).

For the example of a well in a uniform background flow Figures 6 and 7 show that there is a good correspondence between the simulation results obtained by the random walk method and the values of the asymptotic approximations for the probability of arrival at the well and the expected arrival time. For the symmetric flow problem of section 3 the simulation results are also in accordance with the asymptotic approximations, see Table 1. We note that in this example T_{comp} equals T_{bound}. The advantage of the method we have presented above, e.g., the random walk method is that the boundary layer structure is taken as starting point and that analytical expressions are obtained. From Eqs. (4.5) one can construct confidence domains where, e.g., 5% or less of the released pollution reaches the boundary of the protected zone.

The model described in this study may be extended to include loss of contamination by, for example, adsorption or radioactive decay, see Van Kooten (1994), who also considers the total flux of the polluting particles at the well. For the calculation of such a flux, see also Naeh et al. (1990).

Finally, it is noted that also for numerically obtained advective flow patterns the asymptotic solutions for the probability of arrival and the expected arrival time are given by the formulas (4.5) and (4.16), which can be computed by numerical integration. For that purpose only the velocity along the separating streamline has to be taken into account.

Acknowledgements

I want to thank Prof. Johan Grasman and Prof. Pieter A.C. Raats for the discussions concerning the subject treated in this paper and for their remarks on the text.

References

Bear, J. and Y. Bachmat (1990), *Introduction to Modeling of Transport Phenomena in Porous Media*, Kluwer, Dordrecht.

Bear, J. and A. Verruijt (1987), *Modeling Groundwater Flow and Pollution*, Reidel, Dordrecht.

Dagan, G. (1989), *Flow and Transport in Porous Formations*, Springer-Verlag, Berlin, Heidelberg.

Gardiner, C.W. (1983), *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, Springer-Verlag, Berlin.

Kevorkian, J. and J.D. Cole (1981), *Perturbation Methods in Applied Mathematics*, Springer-Verlag, New York.

Mangel, M. and D. Ludwig (1977), *Probability of extinction in a stochastic competition*, SIAM J. Appl. Math., 33, pp. 256-266.

Mangel, M. (1979), *Small fluctuations in systems with multiple steady states*, SIAM J. Appl. Math., 36, pp. 544-572.

Naeh, T., M.M. Klosek, B.J. Matkowsky and Z. Schuss (1990), *A direct approach to the exit problem*, SIAM J. Appl. Math., 50, pp. 595-627.

Saffman, P.G. (1960), *Dispersion due to molecular diffusion and macroscoping mixing in flow through a network of capillaries*, J. Fluid Mech., 7, pp. 194-208.

Uffink, G.J. (1989), *Application of Kolmogorov's backward equation in random walk simulations of groundwater contaminant transport*, in *Contaminant Transport in Groundwater*, H.E. Kobus and W. Kinzelbach, Eds., Balkema, Rotterdam, pp. 283-289.

Van Dyke, M. (1975), *Perturbation Methods in Fluid Mechanics*, The Parabolic Press, Stanford.

Van der Hoek, C.J. (1992), *Contamination of a well in a uniform background flow*, Stoch. Hydrol. Hydraulics, 6, pp. 191-207.

Van Herwaarden, O.A. and J. Grasman (1991), *Dispersive groundwater flow and pollution*, Math. Mod. Methods Appl. Sci., 1, pp. 61-81.

Van Kooten, J.J.A. (1994), *Groundwater contaminant transport including adsorption and first order decay*, Stoch. Hydrol. Hydraulics, 8, pp. 185-205.

Chapter 3

Stochastic epidemics: major outbreaks and the duration of the endemic period[1]

Abstract

A study is made of a two-dimensional stochastic system that models the spread of an infectious disease in a population. An asymptotic expression is derived for the probability that a major outbreak of the disease will occur in case the number of infectives is small. For the case that a major outbreak has occurred, an asymptotic approximation is derived for the expected time that the disease is in the population. The analytical expressions are obtained by asymptotically solving Dirichlet problems based on the Fokker-Planck equation for the stochastic system. Results of numerical calculations for the analytical expressions are compared with simulation results.

1. Introduction

In this study we consider a two-dimensional stochastic system that arises in epidemiology. It is a model for the spread of an infectious disease in which the population is divided in three classes: susceptibles, infectives and removed. We consider the case where in the corresponding deterministic system the disease becomes endemic. In the stochastic system the disease can disappear from the population because of stochastic fluctuations. With probability one this will happen within a finite time. In this paper we are interested in answering the following questions. If there are one or a few infectives in the population, what is the

---

[1]By O.A. van Herwaarden and J. Grasman. Published (in slightly abridged form) in: Journal of Mathematical Biology, 33 (1995), pp. 581-601

probability that a major outbreak of the disease will occur? And given that a major outbreak of the disease has occurred, what is the expected extinction time of the disease?

In the literature various stochastically perturbed dynamical systems with a stable deterministic equilibrium have been described for the case that the stochastic fluctuations are small. References are given in the introduction of Roozen (1989). In that paper a two-dimensional stochastic system from population dynamics is treated. The diffusion matrix of that system becomes singular at the boundary of the region under consideration. There, the normal components of the diffusion and the drift vanish linearly with the distance to the boundary. In our study the methods used in Roozen (1989) are applied to derive asymptotic expressions for the probability that a major outbreak of the disease will occur, and for the expected extinction time in case it has occurred. In our system the normal components of the diffusion also vanish at the boundary. It differs from the system described in Roozen (1989) at the following points: the diffusion matrix is nondiagonal and, moreover, at part of the boundary the deterministic vector field enters the region under consideration. It also turns out that the method cannot be applied to every stochastic population problem with a stable deterministic equilibrium.

In section 2 we describe the stochastic system, formulate the corresponding Fokker-Planck (or Kolmogorov) equations and analyze their use in exit problems. In section 3 a boundary value problem is formulated and solved asymptotically for the probability that a major outbreak of the disease will take place in case the initial number of infectives is small. Moreover, for the probability of a major outbreak a discrete approximation is given. In section 4 we formulate a boundary value problem. Its asymptotic solution forms an approximation of the expected extinction time of the disease in case a full epidemic has developed. In section 5 the analytical expressions are compared with results obtained by random walk simulations.

2. The stochastic model and the Fokker-Planck equations

2.1 The deterministic model

In the epidemic model of this study, the population is divided in three classes: susceptibles, infectives and removed. The sizes of these classes are S, I and R, respectively. In the deterministic version of the model (see, e.g., Edelstein-Keshet (1988)), it is assumed that the rate of transition from susceptibles to infectives is proportional with S and I with transmission rate constant β. The rate of transition from infectives to removed is proportional with I with rate constant γ, and renewal takes place with rate constant μ. The size of the total population is constant: $N = S + I + R$. Our asymptotic analysis is based on the assumption that N is a large parameter. See Figure 1 for a diagram of this model. We write

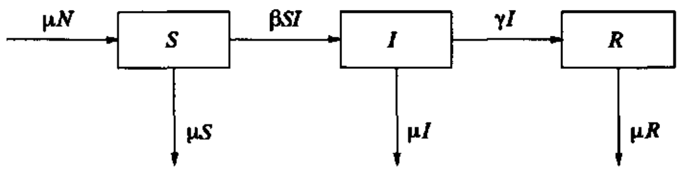


Figure 1. *Subdivision of the total population (N) into three classes: susceptibles (S), infectives (I) and removed (R). Transitions and renewal take place with rate constants* $\beta = \tilde{\beta}/N$, γ *and* μ.

$$\beta = \frac{\tilde{\beta}}{N} \tag{2.1}$$

and assume $\tilde{\beta}$, γ and μ to be of order $O(1)$, where $O(.)$ denotes the order symbol introduced by Landau. Thus, we compare the size of these three parameters with the large parameter N and fix them while increasing N indefinitely.

The dynamics of the deterministic system is, therefore, given by

$$\frac{dS}{dt} = \mu N - \frac{\tilde{\beta}}{N}SI - \mu S, \tag{2.2a}$$

$$\frac{dI}{dt} = \frac{\tilde{\beta}}{N}SI - \gamma I - \mu I, \tag{2.2b}$$

$$\frac{dR}{dt} = \gamma I - \mu R. \tag{2.2c}$$

From these equations, it is easily seen that the total population N is constant, indeed. We can, therefore, eliminate one variable, say R, and restrict our attention to the Eqs. (2.2a) and (2.2b) in the two unknowns S and I. Introducing the scaled population sizes

$$x_1 = \frac{S}{N}, \qquad x_2 = \frac{I}{N} \tag{2.3}$$

and a new time scale by substituting t for μt, we obtain

$$\frac{dx_1}{dt} = 1 - \kappa x_1 x_2 - x_1, \tag{2.4a}$$

$$\frac{dx_2}{dt} = (\kappa x_1 - \nu - 1)x_2, \tag{2.4b}$$

where κ and ν are dimensionless parameters

$$\kappa = \frac{\bar{\beta}}{\mu}, \qquad \nu = \frac{\gamma}{\mu}. \tag{2.5}$$

We now consider the state space $\bar{D}$ with

$$D = \{(x_1, x_2) \mid x_1 > 0, x_2 > 0\}. \tag{2.6}$$

It is noted that at the boundary $x_1 = 0$ the deterministic vector field enters the region D, and that at the boundary $x_2 = 0$ the normal component of the deterministic vector field vanishes. The system (2.4), in which the number of parameters has been reduced to two, has two equilibria:

$$P_1 = (1, 0), \qquad P_2 = (\bar{x}_1, \bar{x}_2) = (\frac{\nu + 1}{\kappa}, \frac{1}{\nu + 1} - \frac{1}{\kappa}). \tag{2.7}$$

In this study, we are interested in values of the (positive) parameters κ and ν for which the critical point P_2 is situated in the first quadrant, that is, we assume

$$\kappa > \nu + 1. \tag{2.8}$$

By this assumption, P_1 is a saddle point and P_2 is a stable equilibrium. Depending on the values of κ and ν, the stable steady point P_2 is a node, an improper node or a focus. This means for the deterministic model, that, once the number of infectives is not equal to 0, the state of the population will develop towards this steady state and the disease will not die out. We remark that for the unscaled populations S and I the critical points P_1 and P_2 correspond with the equilibria

$$(N, 0), \qquad (\frac{\gamma + \mu}{\beta}, \frac{\mu N}{\gamma + \mu} - \frac{\mu}{\beta}). \tag{2.9a,b}$$

2.2 The stochastic model

In this subsection a stochastic model is introduced corresponding with the deterministic model described above. In this model we consider deterministic inflow, as above, but the outflow and other transitions are considered to be stochastic. To be precise, we assume that in the small time interval $(t, t + \Delta t)$ S decreases by one and I increases by one because of a transition from the susceptibles to the infectives with probability $\beta SI\Delta t$. Furthermore, in the small time interval of length Δt, I decreases by one because of a transition from the infectives to the removed with probability $\gamma I\Delta t$, and because of outflow of an individual the sizes of the classes S and I decrease independently by one with probability $\mu S\Delta t$ and $\mu I\Delta t$, respectively. The probability of more than one transition is of order $O((\Delta t)^2)$, and is neglected for small Δt. The deterministic inflow into the class of susceptibles in the time interval Δt is $\mu N\Delta t$ with N fixed. We remark that in this stochastic model the parameter N takes a slightly different role: it is the size of the population if the disease is absent. Moreover, it is remarked that the method we use cannot be applied to a system with deterministic outflow $\mu S\Delta t$. In that case the type of singularity of the probability density function in $(N, 0)$ is such that the divergence theorem does not hold for that function.

Introducing again the scaled population sizes (2.3), and using (2.1), we obtain the following transition probabilities in the small time interval Δt:

$\tilde{\beta}Nx_1x_2\Delta t$	for the transition	$x_1 \to x_1 - 1/N$,	$x_2 \to x_2 + 1/N$,
$\gamma Nx_2\Delta t$	for the transition	$x_2 \to x_2 - 1/N$,	because of removal,
$\mu Nx_1\Delta t$	for the transition	$x_1 \to x_1 - 1/N$,	because of outflow,
$\mu Nx_2\Delta t$	for the transition	$x_2 \to x_2 - 1/N$,	because of outflow.

This yields for the first moments of the changes of x_1 and x_2 over the time interval Δt

$$E(\Delta x_1) = \mu\Delta t - \frac{1}{N}\tilde{\beta}Nx_1x_2\Delta t - \frac{1}{N}\mu Nx_1\Delta t, \tag{2.10a}$$

$$E(\Delta x_2) = \frac{1}{N}\tilde{\beta}Nx_1x_2\Delta t - \frac{1}{N}\gamma Nx_2\Delta t - \frac{1}{N}\mu Nx_2\Delta t, \tag{2.10b}$$

and for the second moments

$$E((\Delta x_1)^2) = (\mu\Delta t)^2 + \frac{1}{N^2}\tilde{\beta}Nx_1x_2\Delta t + \frac{1}{N^2}\mu Nx_1\Delta t, \tag{2.11a}$$

$$E((\Delta x_2)^2) = \frac{1}{N^2}\tilde{\beta}Nx_1x_2\Delta t + \frac{1}{N^2}\gamma Nx_2\Delta t + \frac{1}{N^2}\mu Nx_2\Delta t. \tag{2.11b}$$

We observe that the variances of Δx_1 and Δx_2 equal the second moments up to $O((\Delta t)^2)$. So the stochastic process is approximated by the system of stochastic differential equations of Ito type

$$dx_1 = (\mu - \tilde{\beta}x_1x_2 - \mu x_1)dt - \sqrt{\tilde{\beta}x_1x_2/N}\,dW_1 - \sqrt{\mu x_1/N}\,dW_2, \tag{2.12a}$$

$$dx_2 = (\tilde{\beta}x_1x_2 - \gamma x_2 - \mu x_2)dt + \sqrt{\tilde{\beta}x_1x_2/N}\,dW_1 - \sqrt{\gamma x_2/N}\,dW_3 - \sqrt{\mu x_2/N}\,dW_4 \tag{2.12b}$$

with dW_i the increments of the independent Wiener processes $W_i(t)$, $i = 1, \ldots, 4$.

From these equations we can obtain the forward Fokker-Planck (or forward Kolmogorov) equation, which is a differential equation for the probability density function $p(x_1, x_2, t)$ of finding the system in state (x_1, x_2) at time t. The Fokker-Planck equation that corresponds with the system of stochastic differential equations of Ito type

$$dx = \tilde{b}(x)dt + \tilde{\sigma}(x)dW(t), \tag{2.13}$$

is given by

$$\frac{\partial p}{\partial t} = -\sum_i \frac{\partial}{\partial x_i}(\tilde{b}_i p) + \frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial x_i \partial x_j}(\tilde{a}_{ij} p), \tag{2.14}$$

where

$$\tilde{a}_{ij} = (\tilde{\sigma}\tilde{\sigma}^{\mathrm{T}})_{ij}, \tag{2.15}$$

see, e.g., Gardiner (1983). Here $\tilde{\sigma}^{\mathrm{T}}$ denotes the transpose of $\tilde{\sigma}$. Writing (2.12) in the form (2.13) with x and $\tilde{b}(x)$ vectors in $\mathbb{R}^2$, $W(t)$ a 4-dimensional Wiener process and $\tilde{\sigma}(x)$ a 2×4 matrix, we obtain the forward Fokker-Planck equation

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x_1}\{(\mu - \tilde{\beta}x_1x_2 - \mu x_1)p\} - \frac{\partial}{\partial x_2}\{(\tilde{\beta}x_1 - \gamma - \mu)x_2 p\} + \tag{2.16}$$

$$+ \frac{1}{2N}\left[\frac{\partial^2}{\partial x_1^2}\{(\tilde{\beta}x_2 + \mu)x_1 p\} - 2\frac{\partial^2}{\partial x_1 \partial x_2}(\tilde{\beta}x_1x_2 p) + \frac{\partial^2}{\partial x_2^2}\{(\tilde{\beta}x_1 + \gamma + \mu)x_2 p\}\right].$$

Introducing again the new time scale t instead of μt and the dimensionless parameters (2.5), we obtain the forward Fokker-Planck equation

$$\frac{\partial p}{\partial t} = Mp \equiv -\sum_{i=1}^{2}\frac{\partial}{\partial x_i}\{b_i(x)p\} + \frac{1}{2N}\sum_{i,j=1}^{2}\frac{\partial^2}{\partial x_i \partial x_j}\{a_{ij}(x)p\}, \tag{2.17}$$

where the elements a_{ij} of the diffusion matrix are given by

$$a_{11} = (\kappa x_2 + 1)x_1, \quad a_{12} = a_{21} = -\kappa x_1 x_2, \quad a_{22} = (\kappa x_1 + \nu + 1)x_2, \tag{2.18}$$

and the elements b_i of the drift vector by

$$b_1 = 1 - \kappa x_1 x_2 - x_1, \quad b_2 = (\kappa x_1 - \nu - 1)x_2 \tag{2.19}$$

corresponding with the associated deterministic system (2.4).

2.3 The boundary value problems

Related to the forward operator M in (2.17) is the backward operator L, which is the formal adjoint of M. It plays an important role in exit problems. Let us consider a domain Ω with boundary $\partial\Omega$ and let $r(\tilde{x}, x)$ be the probability density function for the state of leaving Ω the first time at $\tilde{x} \in \partial\Omega$, if starting in $x \in \Omega$. Then between the stationary backward Fokker-Planck (or backward Kolmogorov) equation

$$0 = Lu \equiv \sum_{i=1}^{2} b_i(x)\frac{\partial u}{\partial x_i} + \frac{1}{2N}\sum_{i,j=1}^{2} a_{ij}(x)\frac{\partial^2 u}{\partial x_i \partial x_j} \quad \text{in} \quad \Omega \tag{2.20a}$$

with boundary condition

$$u(x) = h(x) \quad \text{at} \quad \partial\Omega, \tag{2.20b}$$

and $r(\tilde{x}, x)$ the following relation exists

$$u(x) = \int_{\partial\Omega} r(\tilde{x}, x)h(\tilde{x})\,dS, \tag{2.21}$$

see Schuss (1980). If the boundary $\partial\Omega$ is composed of two disjoint parts $\partial\Omega_0$ and $\partial\Omega_1$, this information can be employed to calculate the probability that the state reaches $\partial\Omega$ the first time at the part $\partial\Omega_1$, if starting in $x \in \Omega$. This probability $u(x)$ is given by Eq. (2.21) if $h = 1$ at $\partial\Omega_1$ and $h = 0$ at $\partial\Omega_0$. Therefore, $u(x)$ satisfies the Dirichlet problem

$$Lu = 0 \quad \text{in} \quad \Omega, \tag{2.22a}$$

$$u = 0 \quad \text{at} \quad \partial\Omega_0, \qquad u = 1 \quad \text{at} \quad \partial\Omega_1. \tag{2.22b}$$

Eqs. (2.22) will be used in section 3 to formulate a boundary value problem for the probability that a major outbreak of the disease will occur in case the number of infectives is small. The backward operator L can also be used to determine the expected time $T(x)$ for the state to leave the domain Ω with boundary $\partial\Omega$, if starting in $x \in \Omega$. This expected exit time satisfies

$$LT = -1 \quad \text{in} \quad \Omega, \tag{2.23a}$$

$$T = 0 \quad \text{at} \quad \partial\Omega, \tag{2.23b}$$

see Schuss (1980). Eq. (2.23a) is the Dynkin equation. The Dirichlet problem (2.23) will be used in section 4 to determine the expected extinction time in case the number of infectives in the population is not small anymore.

3. The probability of a major outbreak

In this section we formulate a Dirichlet problem that answers the question about the probability of a major outbreak of the disease, if initially there is a small number of infectives. In the deterministic model described above a major outbreak of the disease definitely occurs and the disease becomes endemic as soon as there is at least one infective. In the stochastic model the disease may disappear from the population before a full epidemic develops. In this section we derive an asymptotic expression, for large N, for the probability that a major outbreak of the disease does not occur in case the number of infectives is small. For that purpose we consider the domain

$$\Omega = \{(x_1, x_2) \mid x_1 > \frac{\nu + 1}{\kappa},\ x_2 > 0\} \tag{3.1}$$

and the Dirichlet problem

$$Lu = 0 \quad \text{in} \quad \Omega, \tag{3.2a}$$

$$u = 1 \quad \text{at} \quad x_2 = 0, \qquad u = 0 \quad \text{at} \quad x_1 = (\nu + 1)/\kappa \tag{3.2b}$$

with the elliptic operator L defined in (2.20a). The solution $u(x)$ of (3.2) equals the probability that the state of the system reaches the boundary $\partial\Omega$ the first time at the part $\partial\Omega_1$: $x_2 = 0$, if starting in $x \in \Omega$. If $x \in \Omega$ is close to $x_2 = 0$ (and not near $x_1 = (\nu + 1)/\kappa$), then $u(x)$ also equals the probability that an epidemic does not develop for a small initial number of infectives. In particular, we are interested in $u(x)$ for x close to the saddle point $P_1(1, 0)$ of the deterministic system, corresponding with a population in which the infection has just been introduced. Because of the direction of the characteristics of the deterministic vector field and because of N being large, we expect u to change rapidly from 1 to about 0 in a small region along the x_1-axis. Therefore, a boundary layer is expected to be present at this place. From boundary layer theory (see Kevorkian and Cole (1981)) it is concluded that there is a boundary layer of width $O(1/N)$ along the x_1-axis. Inside this boundary layer another boundary layer region of width $O(1/N^{1/2})\times O(1/N)$ is contained near the critical point P_1 of the deterministic system; see Figure 2. In our analysis we exclude a small region near the point $((\nu + 1)/\kappa, 0)$. It requires a separate asymptotic analysis, which is not important for the problem we presently address.

To solve problem (3.2) in the outer region we may neglect the diffusion terms and approximate (3.2a) by the equation with drift terms only:

$$\sum_{i=1}^{2} b_i(x)\frac{\partial u}{\partial x_i} = 0, \tag{3.3a}$$

$$u = 0 \quad \text{at} \quad x_1 = (\nu + 1)/\kappa, \tag{3.3b}$$

which has the solution

$$u(x) = 0. \tag{3.4}$$

Inside the boundary layer near P_1 we introduce the stretched coordinates

$$\xi = (x_1 - 1)\sqrt{N}, \qquad \eta = Nx_2. \tag{3.5}$$

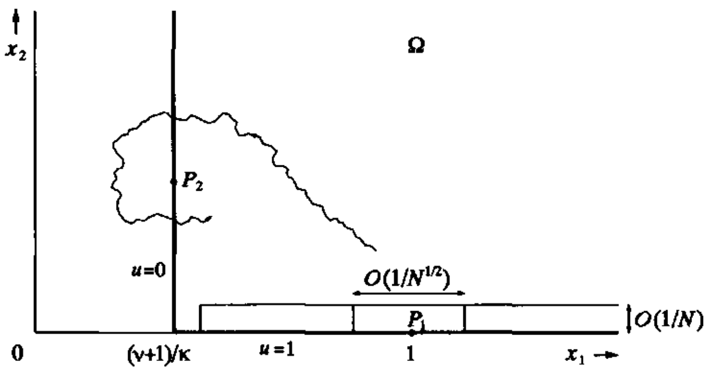


Figure 2. *Boundary layer regions for the Dirichlet problem* (3.2), *used to determine the probability that a major outbreak of the disease will not occur.*

Substitution of (3.5) in (3.2) yields for $N \to \infty$

$$-\xi\frac{\partial u}{\partial \xi} + (\kappa - \nu - 1)\eta\frac{\partial u}{\partial \eta} + \frac{1}{2}\frac{\partial^2 u}{\partial \xi^2} + \frac{1}{2}(\kappa + \nu + 1)\eta\frac{\partial^2 u}{\partial \eta^2} = 0, \quad (3.6a)$$

with the boundary condition

$$u(\xi, \eta) = 1 \quad \text{at} \quad \eta = 0 \quad (3.6b)$$

and the matching condition with the outer solution (3.4)

$$\lim_{\eta \to \infty} u(\xi, \eta) = 0. \quad (3.6c)$$

Separation of variables leads to the boundary layer solution

$$u(\xi, \eta) = \exp[-\alpha\eta] \quad (3.7a)$$

with

$$\alpha = \frac{2(\kappa - \nu - 1)}{(\kappa + \nu + 1)}, \quad (3.7b)$$

see Roozen (1989), where an analogous problem is solved. Inside the boundary layer along the x_1-axis, not near P_1, we introduce the stretched coordinate

$$\eta = Nx_2. \quad (3.8)$$

Substitution in (3.2) yields for $N \to \infty$

$$(1 - x_1)\frac{\partial u}{\partial x_1} + (\kappa x_1 - \nu - 1)\eta\frac{\partial u}{\partial \eta} + \tfrac{1}{2}(\kappa x_1 + \nu + 1)\eta\frac{\partial^2 u}{\partial \eta^2} = 0, \qquad (3.9a)$$

with the boundary condition

$$u(x_1, \eta) = 1 \quad \text{at} \quad \eta = 0. \qquad (3.9b)$$

The matching condition with the outer solution (3.4) is

$$\lim_{\eta\to\infty} u(x_1, \eta) = 0. \qquad (3.9c)$$

Substitution of the similarity solution

$$u(x_1, \eta) = \exp[-\phi(x_1)\eta] \qquad (3.10)$$

in (3.9a) yields

$$(1 - x_1)\phi' + (\kappa x_1 - \nu - 1)\phi - \tfrac{1}{2}(\kappa x_1 + \nu + 1)\phi^2 = 0. \qquad (3.11)$$

Dividing this equation by ϕ^2 we obtain a linear differential equation in $1/\phi$, which can be solved by the method of variation of constants. The solution of this first order differential equation contains an integration constant, which we determine by matching (3.10) with the solution (3.7) near P_1, that is by the condition

$$\lim_{x_1\to 1} \phi(x_1) = \frac{2(\kappa - \nu - 1)}{(\kappa + \nu + 1)}. \qquad (3.12)$$

In this way we obtain

$$\phi(x_1) = \frac{2(1 - x_1)e^{\kappa x_1}}{\int_{x_1}^{1} (\kappa s + \nu + 1)\{(1 - s)/(1 - x_1)\}^{\kappa - \nu - 2} e^{\kappa s}\, ds}. \qquad (3.13)$$

We note that the solution (3.7) in the boundary layer region near P_1 is contained in the solution (3.10), because of the relation (3.12). Therefore, the asymptotic solution (3.10) is valid in the boundary layer along the x_1-axis, the neighbourhood of P_1 included. For the probability $U(S_0, I_0)$ that a major outbreak of the disease will not occur, we now obtain from (3.10) the asymptotic approximation in the original population variables

$$U(S_0, I_0) \sim \exp[-\phi(S_0/N)I_0] \qquad (3.14)$$

where S_0 and I_0 are the initial numbers of susceptibles and infectives, respectively, with I_0 small and S_0 larger than and not close to $(\gamma + \mu)/\beta$.

We remark that, for large values of N and $S_0 >> (\gamma + \mu)/\beta$, we can find another approximation for this probability, using the discrete instead of the continuous model. Assuming that, for sufficiently large S_0, we can take S initially constant, we consider the probabilities p and q that I decreases and increases, respectively, by one individual. Using the transition probabilities given in subsection 2.2, we obtain

$$p = \frac{(\gamma + \mu)I}{(\beta S_0 + \gamma + \mu)I}, \qquad q = \frac{\beta S_0 I}{(\beta S_0 + \gamma + \mu)I}, \tag{3.15}$$

so p and q are independent of I. This process can be compared with a linear birth and death process with an absorbing state for $I = 0$; see, e.g., Karlin (1966) or Goel and Richter-Dyn (1974), who also refer to Kendall (1956) for the result (3.17), (3.19) below. For the probability $R(I_0)$ of extinction of the disease we find the relation

$$R(I_0) = pR(I_0 - 1) + qR(I_0 + 1), \tag{3.16}$$

which is a linear second order difference equation in I_0 with constant coefficients. Its solution is

$$R(I_0) = A\left(\frac{p}{q}\right)^{I_0} + B, \tag{3.17}$$

where the relation $q = 1 - p$ has been used. The coefficients A and B are determined by the conditions

$$R(0) = 1, \qquad \lim_{I_0 \to \infty} R(I_0) = 0. \tag{3.18}$$

Using $p < q$ for $S_0 > (\gamma + \mu)/\beta$, we obtain

$$A = 1, \qquad B = 0. \tag{3.19}$$

Assuming that, for small values of I_0, the extinction probability $R(I_0)$ of the disease approximates the probability $U(S_0, I_0)$ that a major outbreak of the disease will not occur, we find the approximation

$$U(S_0, I_0) \sim \left(\frac{\gamma + \mu}{\beta S_0}\right)^{I_0} \quad \text{for} \quad S_0 >> (\gamma + \mu)/\beta. \tag{3.20}$$

4. The expected extinction time

In this section we derive an asymptotic expression for the expected extinction time of the disease in the stochastic model for large values of N. In particular, we are interested in the expected extinction time for initial states in the neighbourhood of the stable equilibrium of the corresponding deterministic system, i.e., after a major outbreak of the disease has occurred and the disease has become endemic. In subsection 4.1 we asymptotically solve the Dirichlet problem (2.23) for the expected exit time. The solution contains an unknown constant. This constant is determined by solving the forward Fokker-Planck equation for the probability density function of the quasi-stationary state of the system in subsection 4.2, and applying its solution in the divergence theorem in subsection 4.3.

4.1 The backward equation

We consider the state space $\bar{D}$ with $D = \{(x_1, x_2) \mid x_1, x_2 > 0\}$. For starting points $x \in D$, we study the expected exit time $T(x)$ at the exit boundary $x_2 = 0$. We assume $T(x)$ to be of the form

$$T(x) = C(N)\tilde{T}(x) \tag{4.1}$$

with $C(N)$ exponentially large. Substitution of (4.1) in the backward Fokker-Planck equation (2.23a) with L defined in (2.20a) yields for $N \to \infty$ the reduced equation

$$\sum_{i=1}^{2} b_i(x)\frac{\partial \tilde{T}}{\partial x_i} = 0. \tag{4.2}$$

This equation is satisfied by a constant, that we can take 1 (any other constant can be taken up in $C(N)$):

$$\tilde{T}(x) = 1, \tag{4.3a}$$

corresponding with

$$T(x) = C(N). \tag{4.3b}$$

This is the outer solution, valid away from $x_2 = 0$. We can explain this solution as follows. For a starting point x, away from $x_2 = 0$, we expect the state of the system to approach first the stable equilibrium and next to circle in the neighbourhood of it for a long time. Large excursions from this equilibrium take place with small probabilities. Exit of the state at $x_2 = 0$ occurs during such an excursion and is independent of the starting point x.

Boundary layer analysis reveals the existence of a boundary layer of width $O(1/N)$ along the x_1-axis, in which is contained a boundary layer region of width $O(1/N^{1/2}) \times O(1/N)$ near the saddle point $P_1(1, 0)$ of the deterministic system; see Figure 3. Inside the boundary layer region near P_1 we introduce the stretched coordinates ξ and η, given by (3.5). Substitution of these coordinates and (4.1) into (2.23a) yields for $N \to \infty$

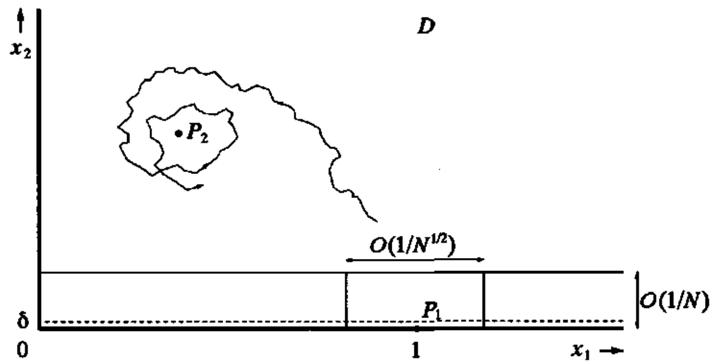


Figure 3. *Boundary layer regions used for the determination of the expected exit time $T(x)$ at the boundary $x_2 = 0$. The (dashed) line $x_2 = \delta$ with $0 < \delta << 1/N$ is used in the divergence theorem in subsection* 4.3.

$$-\xi\frac{\partial\tilde{T}}{\partial\xi} + (\kappa - \nu - 1)\eta\frac{\partial\tilde{T}}{\partial\eta} + \frac{1}{2}\frac{\partial^2\tilde{T}}{\partial\xi^2} + \frac{1}{2}(\kappa + \nu + 1)\eta\frac{\partial^2\tilde{T}}{\partial\eta^2} = 0. \tag{4.4a}$$

The boundary condition (2.23b), $T(x) = 0$, at the exit boundary $x_2 = 0$ leads to

$$\tilde{T}(\xi, \eta) = 0 \quad \text{at} \quad \eta = 0, \tag{4.4b}$$

and matching with the outer solution (4.3a) requires that the following matching condition is satisfied

$$\lim_{\eta\to\infty} \tilde{T}(\xi, \eta) = 1. \tag{4.4c}$$

With $\tilde{T}$ replaced by $1 - u$, these equations are similar to the Eqs. (3.6). We thus obtain the solution

$$\tilde{T}(\xi, \eta) = 1 - \exp[-\alpha\eta], \tag{4.5a}$$

corresponding with

$$T(x_1, x_2) = C(N)(1 - \exp[-\alpha N x_2]) \tag{4.5b}$$

with α given by (3.7b). Inside the boundary layer along the x_1-axis, not near P_1, we introduce the stretched coordinate η, given by (3.8). Substituting this coordinate and (4.1) into (2.23a), we obtain for $N \to \infty$

$$(1 - x_1)\frac{\partial\tilde{T}}{\partial x_1} + (\kappa x_1 - \nu - 1)\eta\frac{\partial\tilde{T}}{\partial\eta} + \tfrac{1}{2}(\kappa x_1 + \nu + 1)\eta\frac{\partial^2\tilde{T}}{\partial\eta^2} = 0, \tag{4.6a}$$

with the condition at the exit boundary

$$\tilde{T}(x_1, \eta) = 0 \quad \text{at} \quad \eta = 0 \tag{4.6b}$$

and the matching conditions with the outer solution (4.3a) and the boundary layer solution (4.5a) near P_1, respectively,

$$\lim_{\eta\to\infty} \tilde{T}(x_1, \eta) = 1, \tag{4.6c}$$

$$\lim_{x_1\to 1} \tilde{T}(x_1, \eta) = 1 - \exp[-\alpha\eta]. \tag{4.6d}$$

In section 3 this problem is solved with $\tilde{T}$ replaced by $1 - u$. We obtain the boundary layer solution

$$\tilde{T}(x_1, \eta) = 1 - \exp[-\phi(x_1)\eta] \tag{4.7}$$

with $\phi(x_1)$ given by (3.13). We note that the solution (4.5a) in the boundary layer region near P_1 is contained in (4.7). Therefore, summarizing (4.1), (4.5a) and (4.7), we have the expression for the expected exit time

$$T(x_1, x_2) = C(N)(1 - \exp[-\phi(x_1)Nx_2]) \tag{4.8}$$

valid in the boundary layer along the x_1-axis, the neighbourhood of P_1 included. The subsections 4.2 and 4.3 concern the determination of the unknown constant $C(N)$.

4.2 The forward equation

In this subsection we analyze the forward Fokker-Planck equation

$$Mp = 0 \tag{4.9}$$

with the elliptic operator M defined by (2.17). A solution $p(x)$ of this equation will be used in the divergence theorem in subsection 4.3. The function $p(x)$ we choose approximates, if appropriately scaled, the quasi-stationary distribution away from the boundary. With the quasi-stationary probability distribution we mean the distribution given the system has not reached the boundary of D.

The WKB-approximation. For the function $p(x)$ we use a WKB-approximation, see Ludwig (1975),

$$p(x_1, x_2) = w(x_1, x_2)\exp[-NQ(x_1, x_2)] \tag{4.10a}$$

for large values of N. The positive functions $Q(x_1, x_2)$ and $w(x_1, x_2)$ satisfy the following conditions in the stable equilibrium P_2 of the deterministic system:

$$Q(\bar{x}_1, \bar{x}_2) = 0, \tag{4.10b}$$

$$w(\bar{x}_1, \bar{x}_2) = 1, \tag{4.10c}$$

where $\bar{x}_1$ and $\bar{x}_2$ are the coordinates of P_2, see (2.7). Since P_2 is the most likely place for the state to be found, Q should have a minimum in P_2. Substituting the WKB-approximation into (4.9) and rearranging the terms, we obtain to leading order $O(N)$ the equation

$$\sum_{i=1}^{2} b_i(x)\frac{\partial Q}{\partial x_i} + \frac{1}{2}\sum_{i,j=1}^{2} a_{ij}(x)\frac{\partial Q}{\partial x_i}\frac{\partial Q}{\partial x_j} = 0, \tag{4.11}$$

which compares to the eikonal equation of the ray theory in optics. Next the terms of order $O(N^0)$ constitute the transport equation

$$\sum_{i=1}^{2} \frac{\partial}{\partial x_i}\{b_i(x)w\} + \sum_{i,j=1}^{2} [\frac{\partial Q}{\partial x_i}\frac{\partial}{\partial x_j}\{a_{ij}(x)w\} + \frac{1}{2}a_{ij}(x)w\frac{\partial^2 Q}{\partial x_i \partial x_j}] = 0. \tag{4.12}$$

The eikonal equation is a Hamilton-Jacobi equation and can be written in the form

$$H(x_1, x_2, p_1, p_2) \equiv \sum_{i=1}^{2} b_i p_i + \frac{1}{2}\sum_{i,j=1}^{2} a_{ij} p_i p_j = 0 \tag{4.13}$$

with

$$p_i = \frac{\partial Q}{\partial x_i} \tag{4.14}$$

and Hamiltonian H. The corresponding system of ordinary differential equations is (see, e.g., Courant and Hilbert (1962))

$$\frac{dx_i}{ds} = \frac{\partial H}{\partial p_i} = b_i + \sum_{j=1}^{2} a_{ij}p_j, \tag{4.15a}$$

$$\frac{dp_i}{ds} = -\frac{\partial H}{\partial x_i} = -\sum_{j=1}^{2} \frac{\partial b_j}{\partial x_i}p_j - \frac{1}{2}\sum_{j,k=1}^{2} \frac{\partial a_{jk}}{\partial x_i}p_j p_k, \tag{4.15b}$$

where s is a parameter along the characteristics. For the rate of change of Q with s we have

$$\frac{dQ}{ds} = - H(x, p) + \sum_{i=1}^{2} \frac{dx_i}{ds} p_i = \frac{1}{2} \sum_{i,j=1}^{2} a_{ij} p_i p_j \quad (\geq 0). \tag{4.15c}$$

The projections of the characteristics on the x_1,x_2-space are called rays. The Eqs. (4.15) are called the ray equations. From Eqs. (4.15a), (4.15b) it is seen that the state variables x and p do not depend on Q. The dynamical system (4.15a), (4.15b) for the state (x_1, x_2, p_1, p_2) has an equilibrium

$$(\frac{\nu + 1}{\kappa}, \frac{1}{\nu + 1} - \frac{1}{\kappa}, 0, 0). \tag{4.16}$$

The projection of this equilibrium on the x_1,x_2-space coincides with the stable equilibrium P_2 of the deterministic system. The rays can be interpreted as paths of maximum likelihood joining P_2 with points of the x_1,x_2-space, see Ludwig (1975). It is noted that two other equilibria of the dynamical system (4.15a), (4.15b) are given by

$$(1, 0, 0, 0) \quad \text{and} \quad (1, 0, 0, -\alpha) \tag{4.17a,b}$$

with α given by (3.7b). The projections of these equilibria on the x_1,x_2-space coincide with the saddle point P_1 of the deterministic system.

Local analysis near the stable equilibrium. Since the function $Q(x_1, x_2)$ has a minimum in the stable equilibrium P_2 of the deterministic system, its first order derivatives p_1 and p_2 vanish in P_2. Therefore, we can approximate $Q(x_1, x_2)$ in the neighbourhood of P_2 by the quadratic form

$$Q(x_1, x_2) \sim \frac{1}{2} \sum_{i,j=1}^{2} C_{ij}(x_i - \bar{x}_i)(x_j - \bar{x}_j), \tag{4.18}$$

see also (4.10b). In this formula $\bar{x}_1$ and $\bar{x}_2$ are the coordinates of P_2, the coefficients C_{ij} are the elements of a symmetric matrix C. Differentiation of (4.18) yields the approximation for p_i

$$p_i \sim \sum_{j=1}^{2} C_{ij}(x_j - \bar{x}_j). \tag{4.19}$$

Values for C_{ij} are obtained as follows. In the neighbourhood of P_2 the drift vector may be approximated by

$$b_i \sim \sum_{j=1}^{2} B_{ij}(x_j - \bar{x}_j), \tag{4.20a}$$

where the elements B_{ij} of the matrix B are given by

$$B_{ij} = \frac{\partial b_i}{\partial x_j}(\bar{x}_1, \bar{x}_2), \tag{4.20b}$$

and the elements a_{ij} of the diffusion matrix may be approximated by their values in P_2

$$a_{ij} \approx a_{ij}(\bar{x}_1, \bar{x}_2). \tag{4.21}$$

The $a_{ij}(\bar{x}_1, \bar{x}_2)$ are the elements of a symmetric matrix A. Substituting the approximations (4.19), (4.20) and (4.21) into the eikonal equation (4.11) and setting the coefficients of the products $(x_i - \bar{x}_i)(x_j - \bar{x}_j)$ equal to zero, we obtain the matrix equation

$$CAC + CB + B^{T}C = 0, \tag{4.22}$$

where we used the symmetry of A and C and introduced B^{T} denoting the transpose of matrix B. Left and right multiplication of (4.22) by the inverse $G = C^{-1}$ yields

$$A + BG + GB^{T} = 0. \tag{4.23}$$

This matrix equation can be seen as a system of linear equations in the elements of the matrix G. Solving this system and inverting G yields the elements C_{ij} of matrix C. It can easily be verified that these two operations can be carried out, indeed.

From this local analysis near P_2 initial values for Q, p_1 and p_2 may be obtained in case a numerical solution of the ray equations is desired. We note that, since p_1 and p_2 vanish in P_2, the rays cannot be chosen to emanate from P_2. Instead, the starting points of the rays (for $s = 0$) are chosen in a small neighbourhood of P_2. From Eq. (4.15c) it is seen, that the function Q is nondecreasing along each ray.

The boundary $x_2 = 0$. To determine the behaviour of $Q(x_1, x_2)$ in the neighbourhood of the boundary $x_2 = 0$, we substitute the expansion for small x_2

$$Q(x_1, x_2) = \tilde{Q}_0(x_1) + \tilde{Q}_1(x_1)x_2 + \tfrac{1}{2}\tilde{Q}_2(x_1)x_2^2 + ... \tag{4.24}$$

into the eikonal equation (4.11). Rearranging terms and setting the coefficient of x_2^0 equal to zero we obtain

$$\frac{d\tilde{Q}_0}{dx_1} = 2(1 - \frac{1}{x_1}) \tag{4.25}$$

with solution

$$\tilde{Q}_0(x_1) = 2x_1 - 2\ln x_1 + \text{constant}. \tag{4.26}$$

From these equations we infer that the only extremum of the function $Q(x_1, x_2)$ along the boundary $x_2 = 0$ is a minimum for $x_1 = 1$, that is in the saddle point of the deterministic system $P_1(1, 0)$. Therefore, along the x_1-axis the probability density function p, given by (4.10a), is peaked at $x_1 = 1$, as is expected from a physical point of view. Asymptotically for $N \to \infty$, exit at the boundary $x_2 = 0$ takes place near the point P_1. The solution is in agreement with the solution of a one-dimensional variant of our model ($x_2 = 0$) that can be calculated explicitly, using the positivity of the probability density function for $x_1 > 0$. We note, that substitution of (4.24) into (4.11) also admits the relation $d\tilde{Q}_0/dx_1 = 0$, yielding $\tilde{Q}_0(x_1) =$ constant. This solution does not lead to a probability density function that is in agreement with the one-dimensional variant and is, therefore, discarded.

Local analysis near the saddle point. To determine the behaviour of the WKB-approximation (4.10a) in the neighbourhood of the saddle point P_1, we substitute the expansion for small values of $(x_1 - 1)$ and x_2

$$Q(x_1, x_2) = Q_0 + Q_1(x_1 - 1) + Q_2 x_2 + \tfrac{1}{2}Q_3(x_1 - 1)^2 + ... \tag{4.27}$$

into the eikonal equation (4.11). Rearranging terms and setting the coefficients of the powers of $(x_1 - 1)$ and x_2 equal to zero, leads to

$$Q_1 = 0, \qquad Q_2 = -\alpha, \qquad Q_3 = 2, \tag{4.28}$$

with α given by (3.7b). This expansion for $Q(x_1, x_2)$ near P_1 is in accordance with the expansion (4.24) along the x_1-axis with $\tilde{Q}_0(x_1)$ given by (4.26), which can be seen by expanding (4.26) in a Taylor series around $x_1 = 1$

$$\tilde{Q}_0(x_1) = 2 + (x_1 - 1)^2 + ... + \text{constant}. \tag{4.29}$$

The expansion (4.27), (4.28) corresponds with the equilibrium $(1, 0, 0, -\alpha)$, see (4.17b), of the system (4.15a), (4.15b) of ray equations for the state (x_1, x_2, p_1, p_2). It is noted, that substitution of (4.27) into (4.11) also admits the values $Q_2 = 0$ and $Q_3 = 0$. The solution $Q_3 = 0$ is omitted, because it does not agree with $\tilde{Q}_0(x_1)$, see (4.29). The solution $Q_2 = 0$ corresponds with the equilibrium $(1, 0, 0, 0)$ of the system (4.15a), (4.15b), see (4.17a). In numerically solving this system for starting points in the neighbourhood of the stable equilibrium P_2, as described above, the equilibrium (4.17b) is attained, and not the equilibrium (4.17a). Therefore, the solution $Q_2 = 0$ is discarded. We note that the numerical solution of the ray equations yields the value of $Q_0 = Q(1, 0)$, that is undetermined by the substitution of (4.27) into (4.11).

We next study the transport function w near the saddle point $P_1(1, 0)$. Introduction of the stretched coordinates ξ and η, given by (3.5), yields for $N \to \infty$

$$\frac{\partial}{\partial\xi}(\xi p) - (\kappa - \nu - 1)\frac{\partial}{\partial\eta}(\eta p) + \frac{1}{2}\frac{\partial^2 p}{\partial\xi^2} + \frac{1}{2}(\kappa + \nu + 1)\frac{\partial^2}{\partial\eta^2}(\eta p) = 0, \tag{4.30}$$

which is satisfied by

$$p = \frac{\text{const.}}{\eta}\exp[-\xi^2 + \alpha\eta]. \tag{4.31}$$

The exponential function in this solution is in agreement with the leading order part of the WKB-approximation near P_1, which is given by (4.10a) with the expansion (4.27), (4.28) for $Q(x_1, x_2)$. This indicates a singular behaviour of the transport function w

$$w \sim \frac{1}{x_2}. \tag{4.32}$$

This is in accordance with the singularity found in the analysis by Roozen (1989). Substitution of the expansion for small values of $(x_1 - 1)$ and x_2

$$w(x_1, x_2) = \frac{w_0 + w_1(x_1 - 1) + w_2 x_2 + \frac{1}{2}w_3(x_1 - 1)^2 + \ldots}{x_2} \tag{4.33}$$

and (4.27), (4.28) into the transport equation (4.12), leads to recurrence relations for the coefficients w_i, leaving the constant w_0 undetermined. We conclude that near P_1 the behaviour of the WKB-approximation is given by

$$p(x_1, x_2) = \frac{k(N)}{x_2}\exp[-N\{(x_1 - 1)^2 - \alpha x_2\}] \tag{4.34}$$

with

$$k(N) = w_0 \exp[-NQ_0] \tag{4.35}$$

and α as in (3.7b). It is noted that w_0 and Q_0 are computed by integration of the eikonal and the transport equation along the ray that connects the stable equilibrium with the saddle point.

The boundary $x_1 = 0$. At the boundary $x_1 = 0$ the drift vector is pointing inward into the region D with normal component $b_1 = 1$, see (2.19), and the normal component of the diffusion vanishes, see (2.18). Therefore, the probability density function $p(x)$ is small near this boundary compared with the boundary $x_2 = 0$, where the normal components of both the drift vector and the diffusion vanish. In Roughgarden (1979) criteria have been described for a boundary classification for one-dimensional stochastic systems, originating from Feller (1952). Application of these criteria on our stochastic model for fixed values of x_2, indicates that the part of the boundary $x_1 = 0$ with $x_2 < (2N - 1)/\kappa$ is an entrance boundary. Such a boundary is unattainable from the interior of the domain, from which we conclude that $p = 0$ at this part of $x_1 = 0$. The other part of the boundary $x_1 = 0$ is at order $O(N)$ distance from the initial states of the stochastic system, which in view of the deterministic vector field leads to large values for Q at that part of the boundary. We note that application of the criteria for boundary classification at the boundary $x_2 = 0$ for fixed values of x_1 yields, indeed, that $x_2 = 0$ is an exit boundary.

4.3 The divergence theorem

In this subsection the constant $C(N)$ in the expected extinction time (4.1) is determined. Using the divergence theorem, the following integral relation can be derived, see Schuss (1980),

$$\int_{D_\delta} (pLT - TMp)\, dD_\delta =$$

$$= \int_{\partial D_\delta} \{ \frac{1}{2N} \sum_{i,j=1}^{2} n_j a_{ij} (p\frac{\partial T}{\partial x_i} - T\frac{\partial p}{\partial x_i}) + \sum_{j=1}^{2} n_j (b_j - \frac{1}{2N} \sum_{i=1}^{2} \frac{\partial a_{ij}}{\partial x_i}) pT \}\, dS \quad (4.36)$$

with n_j the components of the outer normal on the boundary ∂D_δ of the region D_δ. The region D_δ is taken slightly smaller than D:

$$D_\delta = \{(x_1, x_2) \mid x_1 > 0, x_2 > \delta\} \quad (4.37)$$

with

$$0 < \delta << \frac{1}{N} \quad (4.38)$$

(see Figure 3) to avoid evaluating the integrand at the right side of (4.36) at the x_1-axis, where p and its derivatives are singular, see (4.34). We first consider the integral at the right side of (4.36) for the boundary $x_2 = \delta$. The significant contribution to this integral comes from the neighbourhood of P_1. Using the expressions (4.34) for p and (4.5b) for T, and taking the limit $\delta \downarrow 0$, yields after some calculation the integral

$$-\tfrac{1}{2} k(N) C(N) \alpha \int_0^\infty (\kappa x_1 + \nu + 1) \exp[-N(x_1 - 1)^2]\, dx_1 \quad (4.39)$$

approximating asymptotically the right side of (4.36) for $N \to \infty$. Application of the method of Laplace (Erdélyi (1956)) on this integral and using (3.7b), yields the expression

$$-k(N) C(N) (\kappa - \nu - 1) \sqrt{\pi/N}\,. \quad (4.40)$$

The contribution of the boundary $x_1 = 0$ to the integral at the right side of (4.36) is negligible. The part of that boundary with $x_2 < (2N - 1)/\kappa$, where $p = 0$, gives a zero contribution, as can easily be verified. The contribution of the other part of the boundary $x_1 = 0$, where Q is large, is not significant compared with the contribution (4.40). Using (2.23a) and (4.9) the integral at the left side of (4.36) reduces to the integral of $-p$ over the region D_δ. The main contribution to this integral comes from the neighbourhood of the equilibrium P_2, where p is peaked. Using the method of Laplace for double integrals and the expressions (4.10), we obtain the result

$$-\frac{2\pi}{N\sqrt{\Delta}}, \tag{4.41}$$

where Δ is the determinant of the Hessian matrix of Q in P_2. Combining (4.40) and (4.41), and using the fact that w_0, Δ and $(\kappa - \nu - 1)$ are of order $O(1)$, we find that $C(N)$ is of the order

$$C(N) \sim \frac{1}{\sqrt{N}}\exp[Q_0 N]. \tag{4.42}$$

For the expected exit time T for initial states in the neighbourhood of the deterministic stable equilibrium we now obtain by (4.3b) the asymptotic expansion, up to terms of order $O(\ln N)$

$$\ln T(x) = Q_0 N - \tfrac{1}{2}\ln N. \tag{4.43}$$

This answers the question about the expected extinction time of the disease in case a major outbreak has occurred. The expansion (4.43) can be improved by adding an order $O(1)$ term containing w_0.

5. A comparison with stochastic simulations

In this section we compare results of stochastic simulations with the analytical approximations derived in the foregoing sections. The simulations have been carried out for the discrete stochastic model described in subsection 2.2, by following the path of the state (S, I) for initial states (S_0, I_0). For the parameters the values $\mu = 0.2$, $\tilde{\beta} = 0.8$ and $\gamma = 0.2$ have been used, corresponding with $\kappa = 4$ and $\nu = 1$.

In Tables 1 and 2 we give results for the probability that a major epidemic does not develop for small initial numbers of infectives. For initial states (S_0, I_0) an approximation U_{sim} for this probability has been found by carrying out simulation runs. For each starting point (S_0, I_0) 7500 runs have been made. In Table 1 the initial states (S_0, I_0) have been chosen in the neighbourhood of the saddle point $(N, 0)$, for different values of N. The values of I_0 vary from 1 to 5; $S_0 = N - I_0$. The simulation results U_{sim} are compared with the asymptotic approximation U_{as}, computed from (3.14) and (3.13), and the discrete approximation U_{da}, computed from (3.20). We observe that for sufficiently large N the simulation results are in good correspondence with the analytical approximations. In Table 2 the analytical approximations are compared with simulation results for initial states (S_0, I_0) near the S-axis, not only near the saddle point $(N, 0)$. The simulations have been carried out for $N = 104$ and $I_0 = 2$; the stable equilibrium of the corresponding deterministic model is given by $(S, I) = (52, 26)$. For the parameters used in these simulations, we see that the asymptotic approximation U_{aa} is in good accordance with the simulation results. For the discrete approximation U_{da}, this is only true for large values of S_0, as expected. We remark that the initial states (S_0, I_0) which are not in the neighbourhood of the saddle point, do not correspond with a population in which the disease has just been introduced.

In Figure 4 results are given for the expected extinction time of the disease for initial states in the deterministic stable equilibrium. For different values of N, an approximation T_{sim} for the expected extinction time has been found by carrying out simulation runs. For each starting point (2.9b) 2000 runs have been made. In order to compare the simulation results with the asymptotic expansion (4.43), values of $\ln T_{sim} + (\ln N)/2$ have been plotted as a function of N. For sufficiently large values of N, these points fit well with a line, as is expected from (4.43). The slope of the line fitted through these data points for $N > 60$ is 0.059. By numerically solving the ray equations (4.15) as follows, a value of $Q_0 = Q(1, 0)$ is obtained. A ray ending in the saddle point $P_1(1, 0)$ is found by a shooting method. Initial points (x_1, x_2) for the rays are chosen on a small circle around the stable equilibrium P_2. Corresponding initial values for p_1, p_2 and Q follow from the local analysis near P_2, as described in subsection 4.2. By systematically varying the initial points a ray is found ending in (close to) P_1. In this way the value 0.061 is obtained for Q_0, which is in good accordance with the simulation result.

Table 1. *The probability U that a major outbreak of the disease does not occur for small initial numbers of infectives I_0 in the neighbourhood of the deterministic saddle point $(N, 0)$. The values of U_{sim} have been obtained from* 7500 *runs for each initial state (S_0, I_0); $S_0 = N - I_0$. The analytical approximations U_{aa} and U_{da} have been computed from* (3.14), (3.13) *and* (3.20); μ = 0.2, β = 0.8 *and* γ = 0.2.

	$N = 60$			$N = 120$		
	U_{sim}	U_{aa}	U_{da}	U_{sim}	U_{aa}	U_{da}
$I_0 = 1$	.55	.52	.51	.51	.52	.50
$I_0 = 2$	.31	.27	.27	.27	.27	.26
$I_0 = 3$	.19	.15	.15	.15	.14	.13
$I_0 = 4$	.11	.08	.08	.08	.08	.07
$I_0 = 5$	.08	.05	.05	.05	.04	.04

Table 2. *The probability U that a major outbreak of the disease does not occur for $I_0 = 2$ and $N = 104$. The values of U_{sim} have been obtained from 7500 runs for each initial state (S_0, I_0). The analytical approximations U_{aa} and U_{da} have been computed from* (3.14), (3.13) *and* (3.20); $\mu = 0.2$, $\bar{\beta} = 0.8$ *and* $\gamma = 0.2$.

S_0	74	80	86	92	98	104	110	116	122
U_{sim}	.39	.37	.34	.32	.30	.26	.25	.24	.21
U_{aa}	.37	.35	.32	.30	.28	.26	.25	.23	.22
U_{da}	.49	.42	.37	.32	.28	.25	.22	.20	.18

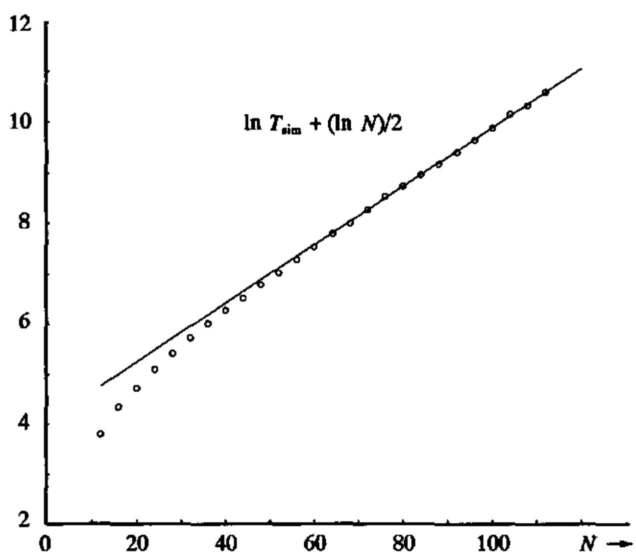


Figure 4. *The expected extinction time of the disease for initial states in the deterministic stable equilibrium. Values of* $\ln T_{sim} + (\ln N)/2$ *have been plotted as a function of N. The values of T_{sim} have been obtained from 2000 runs for each initial state. The plotted line has been fitted through the data points for $N > 60$;* $\mu = 0.2$, $\bar{\beta} = 0.8$ *and* $\gamma = 0.2$. *The slope of this line agrees with the value Q_0 from the asymptotic approximation.*

The disadvantage of simulations is that for each set of parameter values large scale simulation runs have to be made. If N is the only parameter that varies, we can use the theoretical results of the previous sections, because the expected extinction time is approximated by

$$\ln T = AN - \tfrac{1}{2}\ln N + B. \tag{5.1}$$

Thus, carrying out two sets of simulation runs for two different large values of N, we are in the position to estimate A and B without going through the rather complicated analysis of the ray equations.

6. Conclusions

In this paper we have studied a two-dimensional stochastic system modelling the spread of an infectious disease. In this model the population is divided in three classes: susceptibles (S), infectives (I) and removed (R). In our asymptotic analysis it is assumed that there is a large parameter N, the size of the population when the disease is absent.

By asymptotically solving a Fokker-Planck equation we have derived an approximation for the probability that a major outbreak of the disease will occur in case the number of infectives is small. For fixed values of N, this approximation is valid for a larger range of values of S than the approximation based on the discrete model. We have also obtained asymptotic approximations for the expected extinction time of the disease. The value for $Q_0 = Q(1, 0)$ in these expressions is obtained by numerically solving the so-called ray equations. Using initial conditions that follow from a local analysis near the deterministic stable equilibrium P_2, a ray is found that ends in the deterministic saddle point P_1. The asymptotic expression we have derived for ln T can be improved by including an order $O(1)$ term involving w_0, which is to be obtained by integrating the transport equation along the ray.

Because of the similarity of the present study with the analysis of Roozen (1989) for the closely related stochastic prey-predator equations, one would expect that the method and the result in the form of the asymptotic formula for the extinction time hold for a wide class of problems. However, for the case that both the inflow and outflow are deterministic in our epidemic model, we did not succeed in applying the divergence theorem in the present way, because of the resulting singular behaviour of the probability density function p near P_1. We remark that diffusion along the x_1-axis is absent in that case.

Finally, it is noted that, for the model analyzed in this paper, there is a good correspondence between the results obtained by stochastic simulations and the values of the asymptotic approximations we have derived for the probability that a major epidemic develops and the expected extinction time in case it has developed. The coefficients in the asymptotic formula for the extinction time can also be estimated from two sets of simulation runs for two different values of N. Then it is not necessary to solve the ray equations, see section 5.

Acknowledgements

We thank Professor Vincenzo Capasso for bringing to our attention the problem of this paper, and Bea van Wonderen-Tettelaar for advice on the use of a simulation procedure.

References

Courant, R. and D. Hilbert (1962), *Methods of Mathematical Physics*, Vol. 2, Wiley, New York.

Edelstein-Keshet, L. (1988), *Mathematical Models in Biology*, Random House, New York.

Erdélyi, A. (1956), *Asymptotic Expansions*, Dover, New York.

Feller, W. (1952), *The parabolic differential equations and the associated semigroups of transformations*, Ann. Math., 55, pp. 468-519.

Gardiner, C.W. (1983), *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, Springer-Verlag, Berlin.

Goel, N.S. and N. Richter-Dyn (1974), *Stochastic Models in Biology*, Academic Press, New York.

Karlin, S. (1966), *A First Course in Stochastic Processes*, Academic Press, New York.

Kendall, D.G. (1956), *Deterministic and stochastic epidemics in closed populations*, Proc. Symp. Math. Stat. Prob., 3rd, Berkeley, 4.

Kevorkian, J. and J.D. Cole (1981), *Perturbation Methods in Applied Mathematics*, Springer-Verlag, New York.

Ludwig, D. (1975), *Persistence of dynamical systems under random perturbations*, SIAM Rev., 17, pp. 605-640.

Roozen, H. (1989), *An asymptotic solution to a two-dimensional exit problem arising in population dynamics*, SIAM J. Appl. Math., 49, pp. 1793-1810.

Roughgarden, J. (1979), *Theory of Population Genetics and Evolutionary Ecology: an Introduction*, Macmillan, New York.

Schuss, Z. (1980), *Theory and Applications of Stochastic Differential Equations*, Wiley, New York.

Chapter 4

Stochastic epidemics: the probability of extinction of an infectious disease at the end of a major outbreak[1]

Abstract

The aim of this study is to derive an asymptotic expression for the probability that an infectious disease will disappear from a population at the end of a major outbreak ('fade-out'). The study deals with a stochastic *SIR*-model. Local asymptotic expansions are constructed for the deterministic trajectories of the corresponding deterministic system, in particular for the deterministic trajectory starting in the saddle point. The analytical expression for the probability of extinction is derived by asymptotically solving a boundary value problem based on the Fokker-Planck equation for the stochastic system. The asymptotic results are compared with results obtained by random walk simulations.

1. Introduction

In this paper we study a two-dimensional stochastic system modelling the spread of an infectious disease. In this model we consider a population that is divided in three classes: susceptibles, infectives and removed. The population is renewed at a constant rate. We consider the case where in the corresponding deterministic model a stable equilibrium with coexistence of susceptibles and infectives is possible. Then the disease may become endemic. In the stochastic model the disease can disappear from the population because of stochastic fluctuations. With probability one this will happen within a finite time. In Van Herwaarden and Grasman (1995) the following questions have been studied for this stochastic system. What is the probability that a major outbreak of the disease will

[1]Submitted for publication

occur upon the entry of one or a few infectives into the population? And, given that the disease has become endemic, what is the expected extinction time of the disease? In that paper asymptotic expressions have been derived answering these questions.

In the present paper we study the probability of an epidemic 'fade-out'. We are interested in answering the following question: given that a major outbreak of the disease occurs upon the entry of one or a few infectives into the population, what is the probability that the disease will disappear from the population at the end of the major outbreak? In the literature the problem of a stochastic fade-out is touched upon in, e.g., the classic work by Bartlett (1960); Bailey (1975), and Anderson and May (1979, 1991). The analysis in the present paper not only answers the question of the extinction probability of the disease at the end of a major outbreak following the introduction of the disease into a disease free population, but also shows how to determine the extinction probability for other initial situations.

Of particular importance in answering the question of epidemic fade-out is the rate at which the population is renewed. For a large renewal rate the disease will become endemic after a major outbreak with probability close to one. In case of a very small renewal rate extinction of the disease at the end of a major outbreak is almost certain. There is a range of renewal rate values for which the probability of extinction varies from zero to one. The study deals with this transitional case.

In section 2 the deterministic model is described. In section 3 we construct local asymptotic approximations for the trajectories of the deterministic system, in particular for the deterministic trajectory starting in the saddle point. In section 4 the stochastic model is presented, the corresponding Fokker-Planck (or Kolmogorov) equations are formulated and their use in an exit problem is described. In section 5 a parabolic boundary value problem is formulated and asymptotically solved, using the results of section 3. The solution yields an asymptotic approximation for the probability of extinction of the disease at the end of a major outbreak. In section 6 we compare results of numerical calculations for the asymptotic approximation with simulation results. In section 7 we indicate how to determine the extinction probability of the disease for other initial situations.

2. The deterministic model

The population in the epidemic model of this study is divided in three classes: susceptibles, infectives and removed, with sizes S, I and R, respectively. In the deterministic version of the model (see e.g. Edelstein-Keshet (1988)) the rate of transition from susceptibles to infectives is assumed to be proportional with S and I with rate constant β. The rate of transition from infectives to removed is proportional with I with rate constant γ, and renewal takes place with rate constant μ. The size of the total population is constant: $N = S + I + R$. A diagram of the model is shown in Figure 1. We assume that the dependence between β and N is given by

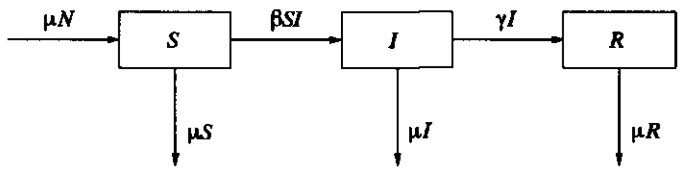


Figure 1. *Diagram of the epidemic model. The total population (N) is divided in susceptibles (S), infectives (I) and removed (R). Transition and renewal rate constants are $\beta = \tilde{\beta}/N$, γ and μ.*

$$\beta = \frac{\tilde{\beta}}{N} \tag{2.1}$$

with $\tilde{\beta}$ independent of N; $\tilde{\beta}$, γ and μ are (strictly) positive.

The deterministic system, therefore, satisfies the differential equations

$$\frac{dS}{dt} = \mu N - \frac{\tilde{\beta}}{N} SI - \mu S, \tag{2.2a}$$

$$\frac{dI}{dt} = \frac{\tilde{\beta}}{N} SI - \gamma I - \mu I, \tag{2.2b}$$

$$\frac{dR}{dt} = \gamma I - \mu R. \tag{2.2c}$$

By adding these equations it is easily seen that the total population N is constant. We can, therefore, eliminate one variable, say R, and restrict our attention to the Eqs. (2.2a) and (2.2b) in the two unknowns S and I. Introducing the scaled population sizes

$$x = \frac{S}{N}, \qquad y = \frac{I}{N} \tag{2.3}$$

we obtain

$$\frac{dx}{dt} = \mu - \tilde{\beta} xy - \mu x, \tag{2.4a}$$

$$\frac{dy}{dt} = (\tilde{\beta} x - \gamma - \mu) y. \tag{2.4b}$$

The system (2.4) has two equilibria:

$$P_1 = (1,\ 0), \qquad P_2 = (\frac{\gamma + \mu}{\bar{\beta}},\ \frac{\mu}{\gamma + \mu} - \frac{\mu}{\bar{\beta}}). \tag{2.5}$$

We are interested in values of the (strictly positive) parameters $\bar{\beta}$, γ and μ for which the critical point P_2 is situated in the first quadrant, that is, we assume

$$\bar{\beta} > \gamma + \mu. \tag{2.6}$$

By this assumption, P_1 is a saddle point and P_2 is a stable equilibrium. Depending on the values of $\bar{\beta}$, γ and μ, the stable equilibrium P_2 is a node, an improper node or a focus. The equilibrium P_2 is a focus if

$$\mu\bar{\beta}^2 - 4(\bar{\beta} - \gamma - \mu)(\gamma + \mu)^2 < 0. \tag{2.7}$$

At the boundary $x = 0$ the deterministic vector field enters the first quadrant, and at the boundary $y = 0$ the normal component of the deterministic vector field vanishes. We note that, once the number of infectives is not equal to zero, the state of the deterministic system will develop towards the stable equilibrium and the disease will not die out.

In this paper we are in particular interested in the development of the disease when one or a few infectives are introduced in a population that until then is free of the disease. In the deterministic system this development is given by the trajectory starting in the point (1, 0), assuming that N is sufficiently large. The aim of this paper is to determine the probability that in the related stochastic system the disease will die out at the end of a major outbreak following the introduction of one or a few infectives into the population. In the deterministic system we, therefore, intend to study the case that the trajectory starting in (1, 0) closely approaches the x-axis at the end of its first cycle. This case occurs for small values of the parameter μ.

Let us first consider the limit case $\mu = 0$, i.e., the Kermack and McKendrick model (see, e.g., Edelstein-Keshet (1988)), in which no renewal takes place. From Eqs. (2.4), that are also valid in the limit case $\mu = 0$, we can obtain an explicit expression for the trajectory of the Kermack and McKendrick model starting in (1, 0):

$$y = 1 - x + \frac{\gamma}{\bar{\beta}} \ln x. \tag{2.8}$$

This trajectory starts at (1, 0), attains a maximum value for $x = \gamma/\bar{\beta}$ and approaches again the x-axis on the interval $(0, \gamma/\bar{\beta})$. It is noted that in the Kermack and McKendrick model the vector field vanishes in all points of the x-axis. In the deterministic system with $0 < \mu << 1$ the trajectory starting in (1, 0) closely approaches the x-axis. In section 3 it will be seen that for small values of μ Eq. (2.8) is a first order approximation of this deterministic trajectory up to a distance of order $O(\mu)$ from the x-axis.

3. The deterministic trajectories

In this section we study the trajectories of the deterministic system introduced in section 2, in particular the trajectory starting in the saddle point P_1. We construct local asymptotic approximations for these trajectories. Our asymptotic analysis is based on the assumption that μ is a small parameter: $0 < \mu << 1$. The parameters $\tilde{\beta}$ and γ are assumed to be of order $O(1)$. For fixed positive values for $\tilde{\beta}$ and γ and sufficiently small values for μ the stable equilibrium P_2 is a focus, as follows from (2.7) and (2.6). It is noted that the distance of P_2 to the x-axis is of order $O(\mu)$.

By introduction of the coordinate

$$\bar{y} = \frac{y}{\mu} \tag{3.1}$$

and a new time scale t instead of μt the system (2.4) is transformed into

$$\frac{dx}{dt} = 1 - \tilde{\beta}x\bar{y} - x, \tag{3.2a}$$

$$\mu\frac{d\bar{y}}{dt} = (\tilde{\beta}x - \gamma - \mu)\bar{y}. \tag{3.2b}$$

We analyze the deterministic trajectories in the $(x, \bar{y})$-plane. In this plane the stable equilibrium P_2 is given by $P_2 = ((\gamma + \mu)/\tilde{\beta},\ 1/(\gamma + \mu) - 1/\tilde{\beta})$. Its distance to the x-axis is of order $O(1)$. We note that a Volterra-Lotka system in which one of the derivatives is multiplied by a small parameter, is studied by Grasman and Veling (1973). In that paper local asymptotic expansions for the closed trajectories are derived from the implicit solution, which is available, because the equivalent of the inflow term is absent in the Volterra-Lotka system. We derive local asymptotic approximations for the trajectories of (3.2) by substituting formal expansions in the differential equation for the trajectories

$$\frac{d\bar{y}}{dx} = \frac{(\tilde{\beta}x - \gamma - \mu)\bar{y}}{\mu(1 - \tilde{\beta}x\bar{y} - x)} \tag{3.3a}$$

or

$$\frac{dx}{d\bar{y}} = \frac{\mu(1 - \tilde{\beta}x\bar{y} - x)}{(\tilde{\beta}x - \gamma - \mu)\bar{y}}. \tag{3.3b}$$

The first two cycles of the trajectory starting in (1, 0) are divided in different segments. In the region $\bar{y} >> (1 - x)/(\tilde{\beta}x)$ we distinguish the segments A and E, in the region $0 < \bar{y} << (1 - x)/(\tilde{\beta}x)$ the segments C and G, and in the intermediate regions the segments B, D and F, see Figure 2. The solutions for these segments obtained by substitution of formal expansions into the differential equation, contain constants that will be denoted by A_i, B_i, ..., depending on the segment. We will determine these constants using matching and initial conditions.

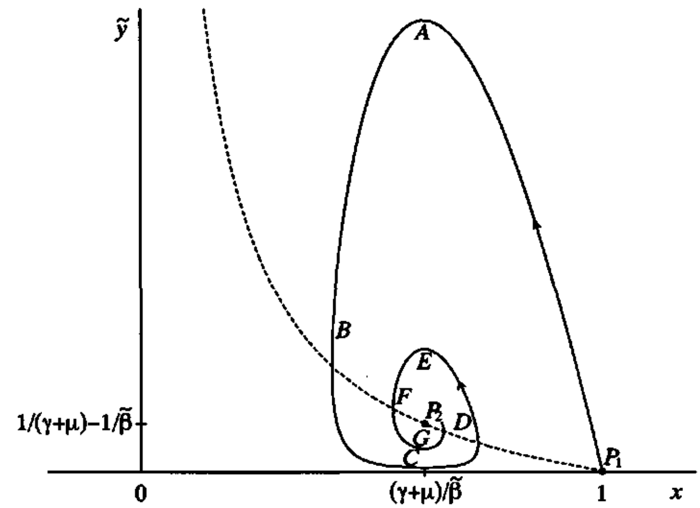


Figure 2. *The deterministic trajectory starting in* (1, 0) *is divided in segments A, B, ... The dashed line is the graph of* $\tilde{y} = (1 - x)/(\tilde{\beta}x)$, *where the tangents of the deterministic trajectory are vertical.*

Local expansions. For segments A and E substitution of the expansion

$$\tilde{y}(x) = \tilde{y}_1(x)\frac{1}{\mu} + \tilde{y}_2(x)\ln\mu + \tilde{y}_3(x) + ..., \qquad 0 < \mu << 1 \tag{3.4a}$$

into (3.3a) yields differential equations for the functions $\tilde{y}_i(x)$. For segment A we obtain to leading order the solution

$$\tilde{y}_1(x) = -x + \frac{\gamma}{\tilde{\beta}}\ln x + A_1 \tag{3.4b}$$

and for the higher order terms

$$\tilde{y}_2(x) = A_2 \tag{3.4c}$$

and

$$\tilde{y}_3(x) = \int_x^{\gamma/\tilde{\beta}} \frac{(1 - A_1)\tilde{\beta}s + \gamma(s - s\ln s - 1)}{\tilde{\beta}^2 s^2(-s + (\gamma/\tilde{\beta})\ln s + A_1)}\,ds + A_3. \tag{3.4d}$$

For segment E the solution can be written as (3.4) with the integration constants A_i replaced by E_i. We note that the $O(\ln \mu)$ term in the expansion (3.4a), that is not suggested by the differential equation, is needed for matching purposes.

For segments B, D and F we substitute the expansion

$$x(\tilde{y}) = x_1(\tilde{y}) + x_2(\tilde{y})\,\mu \ln \mu + x_3(\tilde{y})\,\mu + ..., \qquad 0 < \mu << 1 \tag{3.5a}$$

into (3.3b). Solving the resulting differential equations we obtain for segment B

$$x_1(\tilde{y}) = B_1, \tag{3.5b}$$

$$x_2(\tilde{y}) = B_2 \tag{3.5c}$$

and

$$x_3(\tilde{y}) = -\frac{\tilde{\beta} B_1}{\tilde{\beta} B_1 - \gamma}\tilde{y} + \frac{1 - B_1}{\tilde{\beta} B_1 - \gamma}\ln \tilde{y} + B_3. \tag{3.5d}$$

For the segments D and F the expansions can be written as (3.5) with the constants B_i replaced by D_i and F_i, respectively.

For the segments C and G we substitute into (3.3a) the expansion

$$\tilde{y}(x) = \exp[\hat{y}_1(x)/\mu + \hat{y}_2(x)\ln \mu + \hat{y}_3(x) + ...], \qquad 0 < \mu << 1. \tag{3.6a}$$

For segment C the solutions of the resulting differential equations can be written as

$$\hat{y}_1(x) = -\tilde{\beta} x - (\tilde{\beta} - \gamma)\ln(1 - x) + C_1, \tag{3.6b}$$

$$\hat{y}_2(x) = C_2, \tag{3.6c}$$

$$\hat{y}_3(x) = \ln(1 - x) + C_3. \tag{3.6d}$$

For segment G the expansion can be written as (3.6) with the integration constants C_i replaced by G_i.

Analysis of an integral. We first analyze the integral

$$I(x) = \int_x^{\gamma/\tilde{\beta}} \frac{(1 - \alpha)\tilde{\beta} s + \gamma(s - s\ln s - 1)}{\tilde{\beta}^2 s^2(-s + (\gamma/\tilde{\beta})\ln s + \alpha)}\,ds \tag{3.7}$$

that appears in the function $\tilde{y}_3(x)$ in (3.4d). For matching purposes we need the behaviour of $I(x)$ for $\alpha > \gamma/\tilde{\beta} - (\gamma/\tilde{\beta})\ln(\gamma/\tilde{\beta})$. For these values of α the function

$$f(s) = -s + (\gamma/\tilde{\beta})\ln s + \alpha \tag{3.8}$$

in the denominator of the integrand has two zeros x_l and x_r, satisfying

$$0 < x_l < \frac{\gamma}{\bar{\beta}} \quad \text{and} \quad \frac{\gamma}{\bar{\beta}} < x_r. \tag{3.9a,b}$$

The behaviour of $I(x)$ near x_l and x_r can easily be found using Taylor expansions. We obtain

$$I(x) \sim \frac{1 - x_l}{\bar{\beta}x_l}\ln(x - x_l) - \frac{1 - x_l}{\bar{\beta}x_l}\ln(\frac{\gamma}{\bar{\beta}} - x_l) + J(x_l) \quad \text{for} \quad x \downarrow x_l \tag{3.10a}$$

and, for $\alpha \neq 1$,

$$I(x) \sim \frac{1 - x_r}{\bar{\beta}x_r}\ln(x_r - x) - \frac{1 - x_r}{\bar{\beta}x_r}\ln(x_r - \frac{\gamma}{\bar{\beta}}) + J(x_r) \quad \text{for} \quad x \uparrow x_r \tag{3.10b}$$

with

$$J(b) = \int_b^{\gamma/\bar{\beta}} \frac{(1 - \alpha)\bar{\beta}s + \gamma(s - s\ln s - 1)}{\bar{\beta}^2 s^2(-s + (\gamma/\bar{\beta})\ln s + \alpha)} + \frac{1 - b}{\bar{\beta}b}\frac{1}{s - b}\,ds. \tag{3.10c}$$

For $\alpha = 1$ we have $x_r = 1$. In that case the integral $I(x)$ is convergent for $x = x_r$.
We now match the expansions for the segments obtained above.

Matching A and B. We match the expansion $\tilde{y}(x)$ for segment A given by (3.4) with the expansion $x(\tilde{y})$ for segment B given by (3.5), considering (3.4) as the inner expansion. Expressing (3.4) and (3.5) in the inner variable

$$\eta = \mu\tilde{y} \tag{3.11}$$

we obtain, respectively,

$$\eta = (-x + \frac{\gamma}{\bar{\beta}}\ln x + A_1) + A_2\mu\ln\mu + \tilde{y}_3(x)\mu + ... \tag{3.12}$$

and

$$x = B_1(1 + \xi_\eta) \tag{3.13a}$$

with

$$\xi_\eta = -\frac{\bar{\beta}}{\bar{\beta}B_1 - \gamma}\eta + \frac{B_2}{B_1}\mu\ln\mu + \frac{1 - B_1}{B_1(\bar{\beta}B_1 - \gamma)}\mu\ln\frac{\eta}{\mu} + \frac{B_3}{B_1}\mu + ... \tag{3.13b}$$

We match (3.4) and (3.5) for $\eta \downarrow 0$, $\mu \downarrow 0$, $\eta/(-\mu\ln\mu) \to \infty$ by requiring that substitution of (3.13) into (3.12) yields the identity. Using $\xi_\eta \downarrow 0$ the following matching condition is readily found

$$-B_1 + \frac{\gamma}{\bar{\beta}}\ln B_1 + A_1 = 0, \tag{3.14a}$$

so

$$B_1 = x_{lA}, \tag{3.14b}$$

that is the smaller zero of function $f(s)$, see (3.8), with $\alpha = A_1$. Then using the behaviour of $\tilde{y}_3(x)$ for $x \downarrow x_{lA}$, which follows from (3.10a), we obtain the matching conditions for the higher order terms

$$B_2 = \frac{(\tilde{\beta}A_2 - 1)x_{lA} + 1}{\tilde{\beta}x_{lA} - \gamma}, \tag{3.14c}$$

$$B_3 = \frac{1 - x_{lA}}{\tilde{\beta}x_{lA} - \gamma} \ln \frac{\tilde{\beta}^2 x_{lA}}{(\tilde{\beta}x_{lA} - \gamma)^2} + \frac{\tilde{\beta}x_{lA}}{\tilde{\beta}x_{lA} - \gamma}(A_3 + J(x_{lA})), \tag{3.14d}$$

with $J(x_{lA})$ given by (3.10c) with $\alpha = A_1$.

Matching B and C. We match the expansion $x(\tilde{y})$ for segment B given by (3.5) with the expansion $\tilde{y}(x)$ for segment C given by (3.6). Expressing (3.5) and (3.6) in the intermediate variable

$$\eta = \tilde{y}/\mu \tag{3.15}$$

we obtain, respectively,

$$x = B_1(1 + \xi_\eta) \tag{3.16a}$$

with

$$\xi_\eta = \frac{B_2}{B_1}\mu \ln \mu + \frac{1 - B_1}{B_1(\tilde{\beta}B_1 - \gamma)}\mu \ln(\eta\mu) + \frac{B_3}{B_1}\mu - \frac{\tilde{\beta}}{\tilde{\beta}B_1 - \gamma}\eta\mu^2 + ..., \tag{3.16b}$$

and

$$\eta = \frac{1}{\mu}\exp[(-\tilde{\beta}x - (\tilde{\beta} - \gamma)\ln(1 - x) + C_1)/\mu + C_2 \ln \mu + \ln(1 - x) + C_3]. \tag{3.17}$$

We match (3.5) and (3.6) by requiring that substitution of (3.16) in (3.17) yields the identity for $\eta = O(1)$, $\mu \downarrow 0$. Using $\xi_\eta \downarrow 0$ we obtain

$$C_1 = \tilde{\beta}B_1 + (\tilde{\beta} - \gamma)\ln(1 - B_1), \tag{3.18a}$$

$$C_2 = -\frac{\tilde{\beta}B_1 - \gamma}{1 - B_1}B_2, \tag{3.18b}$$

$$C_3 = -\frac{\tilde{\beta}B_1 - \gamma}{1 - B_1}B_3 - \ln(1 - B_1). \tag{3.18c}$$

Matching C and D. We match the expansion for segment C, given by (3.6), and the expansion for segment D, given by (3.5) with the constants B_i replaced by D_i, analogously to the matching of B and C. This yields the relations

$$\tilde{\beta}D_1 + (\tilde{\beta} - \gamma)\ln(1 - D_1) = C_1, \tag{3.19a}$$

$$D_2 = -\frac{1 - D_1}{\tilde{\beta}D_1 - \gamma}C_2, \tag{3.19b}$$

$$D_3 = -\frac{1 - D_1}{\tilde{\beta}D_1 - \gamma}(C_3 + \ln(1 - D_1)), \tag{3.19c}$$

determining the constants D_i when the constants C_i are known.

Matching D and E. The expansions for the segments D and E, given by (3.5) and (3.4) with the constants B_i and A_i replaced by D_i and E_i, respectively, are matched analogously to the segments B and A. Expressing the expansions in the inner variable $\eta = \mu\tilde{y}$ we obtain expressions analogous to (3.12) and (3.13). Now matching to leading order, using $\xi_\eta \uparrow 0$, yields the relation

$$E_1 = D_1 - \frac{\gamma}{\tilde{\beta}}\ln D_1. \tag{3.20a}$$

We find that D_1 is a zero of function $f(s)$, see (3.8), with $\alpha = E_1$. Writing

$$D_1 = x_{rE} \tag{3.20b}$$

we have $x \uparrow x_{rE}$. Using (3.10b) we obtain the matching relations for the higher order terms

$$E_2 = \frac{(\tilde{\beta}x_{rE} - \gamma)D_2 + x_{rE} - 1}{\tilde{\beta}x_{rE}}, \tag{3.20c}$$

$$E_3 = -\frac{1 - x_{rE}}{\tilde{\beta}x_{rE}}\ln\frac{\tilde{\beta}^2 x_{rE}}{(\tilde{\beta}x_{rE} - \gamma)^2} + \frac{\tilde{\beta}x_{rE} - \gamma}{\tilde{\beta}x_{rE}}D_3 - J(x_{rE}), \tag{3.20d}$$

with $J(x_{rE})$ given by (3.10c) with $\alpha = E_1$.

Matching other segments. The different types of matching needed to determine the deterministic trajectory starting in (1, 0) have now been worked out. The relations thus obtained can be used to complete the matching of the expansions for the other segments.

Matching of the expansions for the segments E and F, given by (3.4) and (3.5) with the constants A_i and B_i replaced by E_i and F_i, respectively, yields analogously to the matching of A and B the following relations

$$-F_1 + \frac{\gamma}{\tilde{\beta}} \ln F_1 + E_1 = 0, \tag{3.21a}$$

that is

$$F_1 = x_{IE}, \tag{3.21b}$$

the smaller zero of function $f(s)$, given by (3.8), with $\alpha = E_1$;

$$F_2 = \frac{(\tilde{\beta}E_2 - 1)x_{IE} + 1}{\tilde{\beta}x_{IE} - \gamma}, \tag{3.21c}$$

$$F_3 = \frac{1 - x_{IE}}{\tilde{\beta}x_{IE} - \gamma} \ln \frac{\tilde{\beta}^2 x_{IE}}{(\tilde{\beta}x_{IE} - \gamma)^2} + \frac{\tilde{\beta}x_{IE}}{\tilde{\beta}x_{IE} - \gamma}(E_3 + J(x_{IE})), \tag{3.21d}$$

with $J(x_{IE})$ given by (3.10c) with $\alpha = E_1$.

Matching of the expansions for the segments F and G, given by (3.5) and (3.6) with the constants B_i and C_i replaced by F_i and G_i, respectively, yields the relations

$$G_1 = \tilde{\beta}F_1 + (\tilde{\beta} - \gamma)\ln(1 - F_1), \tag{3.22a}$$

$$G_2 = -\frac{\tilde{\beta}F_1 - \gamma}{1 - F_1} F_2, \tag{3.22b}$$

$$G_3 = -\frac{\tilde{\beta}F_1 - \gamma}{1 - F_1} F_3 - \ln(1 - F_1). \tag{3.22c}$$

Initial condition. The constants A_1, A_2 and A_3 of the expansion (3.4) of segment A can be determined from the initial condition

$$\tilde{y}(1) = 0. \tag{3.23}$$

We obtain

$$A_1 = 1, \qquad A_2 = 0 \tag{3.24a,b}$$

and

$$A_3 = \int_{\gamma/\tilde{\beta}}^{1} \frac{\gamma(s - s\ln s - 1)}{\tilde{\beta}^2 s^2(-s + (\gamma/\tilde{\beta})\ln s + 1)}\, ds, \tag{3.24c}$$

where A_3 is a convergent integral. We note that in the original x, y-coordinates, see

(2.3), the leading order term of expansion (3.4) for segment A corresponds with expression (2.8) for the trajectory of the Kermack and McKendrick model starting in (1, 0).

We can combine the matching relations derived above. In particular, we obtain the following expression for the expansion of segment C:

$$\tilde{y}(x) = K \exp[(-\tilde{\beta}x - (\tilde{\beta} - \gamma)\ln(1 - x))/\mu + \ln(1 - x)] \tag{3.25a}$$

with

$$K = \frac{1}{\mu}\exp[(\tilde{\beta}x_{IA} + (\tilde{\beta} - \gamma)\ln(1 - x_{IA}))/\mu + C_3], \tag{3.25b}$$

where $x_{IA} \neq 1$ satisfies

$$-x_{IA} + \frac{\gamma}{\tilde{\beta}}\ln x_{IA} + 1 = 0 \tag{3.25c}$$

and

$$C_3 = -\ln\frac{-\tilde{\beta}x_{IA}}{\tilde{\beta}x_{IA} - \gamma} - \int_{x_{IA}}^{1} \frac{x_{IA}}{1 - x_{IA}}\frac{\gamma(s - s\ln s - 1)}{\tilde{\beta}s^2(-s + (\gamma/\tilde{\beta})\ln s + 1)} + \frac{1}{s - x_{IA}}\,ds. \tag{3.25d}$$

This expression will be used in section 5, when we determine an approximation for the probability that in the stochastic model the disease dies out at the end of a major outbreak following the introduction of one or a few infectives into the population.

4. The stochastic model and the Fokker-Planck equations

In this section a stochastic model is introduced corresponding with the deterministic model described in section 2. In this model we consider deterministic inflow, as above, but the outflow and other transitions are considered to be stochastic. To be precise, we assume that in the small time interval $(t, t + \Delta t)$ S decreases by one and I increases by one because of a transition from the susceptibles to the infectives with probability $\beta SI\Delta t$. Furthermore, in the small time interval of length Δt, I decreases by one because of a transition from the infectives to the removed with probability $\gamma I\Delta t$, and because of outflow of an individual the sizes of the classes S and I decrease independently by one with probability $\mu S\Delta t$ and $\mu I\Delta t$, respectively. The probability of more than one transition is of order $O((\Delta t)^2)$, and is neglected for small Δt. The deterministic inflow into the class of susceptibles in the time interval Δt is $\mu N\Delta t$ with N fixed. We remark that in this stochastic model the parameter N takes a slightly different role: it is the size of the population if the disease is absent. Introducing again the scaled population sizes (2.3), and using

(2.1), we obtain the following transition probabilities in the small time interval Δt:

$\tilde{\beta}Nxy\Delta t$	for the transition	$x \to x - 1/N$,	$y \to y + 1/N$,
$\gamma Ny\Delta t$	for the transition	$y \to y - 1/N$,	because of removal,
$\mu Nx\Delta t$	for the transition	$x \to x - 1/N$,	because of outflow,
$\mu Ny\Delta t$	for the transition	$y \to y - 1/N$,	because of outflow.

This yields for the first conditional moments of the changes of x and y over the time interval Δt

$$E(\Delta x) = \mu\Delta t - \frac{1}{N}\tilde{\beta}Nxy\Delta t - \frac{1}{N}\mu Nx\Delta t, \tag{4.1a}$$

$$E(\Delta y) = \frac{1}{N}\tilde{\beta}Nxy\Delta t - \frac{1}{N}\gamma Ny\Delta t - \frac{1}{N}\mu Ny\Delta t, \tag{4.1b}$$

and for the second moments

$$E((\Delta x)^2) = (\mu\Delta t)^2 + \frac{1}{N^2}\tilde{\beta}Nxy\Delta t + \frac{1}{N^2}\mu Nx\Delta t, \tag{4.2a}$$

$$E(\Delta x\Delta y) = -\frac{1}{N^2}\tilde{\beta}Nxy\Delta t, \tag{4.2b}$$

$$E((\Delta y)^2) = \frac{1}{N^2}\tilde{\beta}Nxy\Delta t + \frac{1}{N^2}\gamma Ny\Delta t + \frac{1}{N^2}\mu Ny\Delta t. \tag{4.2c}$$

We observe that the variances of Δx and Δy equal the second moments up to $O((\Delta t)^2)$. So, under the assumption of continuity, the stochastic jump process is approximated by the system of stochastic differential equations of Ito type

$$dx = (\mu - \tilde{\beta}xy - \mu x)dt - \sqrt{\tilde{\beta}xy/N}\,dW_1 - \sqrt{\mu x/N}\,dW_2, \tag{4.3a}$$

$$dy = (\tilde{\beta}xy - \gamma y - \mu y)dt + \sqrt{\tilde{\beta}xy/N}\,dW_1 - \sqrt{\gamma y/N}\,dW_3 - \sqrt{\mu y/N}\,dW_4 \tag{4.3b}$$

with dW_i the increments of the independent Wiener processes $W_i(t)$, $i = 1, \ldots, 4$, see Ludwig (1975).

From these equations we can obtain the forward Fokker-Planck (or forward Kolmogorov) equation, which is a differential equation for the probability density function $p(x, y, t)$ of finding the system in state (x, y) at time t. When we write the system (4.3) in the form

$$dz = b(z)dt + \sigma(z)dW(t) \tag{4.4}$$

with $z = (x, y)^T$ and $b(z) = (b_1(z), b_2(z))^T$ vectors in $\mathbb{R}^2$, $W(t)$ a 4-dimensional Wiener process and $\sigma(z)$ a 2×4 matrix, and define the 2×2 diffusion matrix $a(z)$ by

$$a_{ij} = \frac{1}{N}(\sigma\sigma^T)_{ij}, \tag{4.5}$$

then the function $p(z, t)$ satisfies the forward Fokker-Planck equation

$$\frac{\partial p}{\partial t} = - \sum_{i=1}^{2} \frac{\partial}{\partial z_i}\{b_i(z)p\} + \frac{1}{2N} \sum_{i,j=1}^{2} \frac{\partial^2}{\partial z_i \partial z_j}\{a_{ij}(z)p\}, \tag{4.6}$$

see, e.g., Gardiner (1983). Introducing again the new time scale t instead of μt and the coordinate $\bar{y} = y/\mu$, see (3.1), we obtain the forward Fokker-Planck equation

$$\frac{\partial p}{\partial t} = Mp \equiv - \frac{\partial}{\partial x}\{(1 - \bar{\beta}x\bar{y} - x)p\} - \frac{1}{\mu}\frac{\partial}{\partial \bar{y}}\{(\bar{\beta}x - \gamma - \mu)\bar{y}p\} + \tag{4.7}$$

$$+ \frac{1}{2N}[\frac{\partial^2}{\partial x^2}\{(\bar{\beta}\bar{y} + 1)xp\} - \frac{2}{\mu}\frac{\partial^2}{\partial x \partial \bar{y}}(\bar{\beta}x\bar{y}p) + \frac{1}{\mu^2}\frac{\partial^2}{\partial \bar{y}^2}\{(\bar{\beta}x + \gamma + \mu)\bar{y}p\}].$$

Related to the forward operator M in (4.7) is the backward operator L, which is the formal adjoint of M:

$$L \equiv (1 - \bar{\beta}x\bar{y} - x)\frac{\partial\,\cdot}{\partial x} + \frac{1}{\mu}(\bar{\beta}x - \gamma - \mu)\bar{y}\frac{\partial\,\cdot}{\partial \bar{y}} + \tag{4.8}$$

$$+ \frac{1}{2N}[(\bar{\beta}\bar{y} + 1)x\frac{\partial^2\,\cdot}{\partial x^2} - \frac{2}{\mu}\bar{\beta}x\bar{y}\frac{\partial^2\,\cdot}{\partial x \partial \bar{y}} + \frac{1}{\mu^2}(\bar{\beta}x + \gamma + \mu)\bar{y}\frac{\partial^2\,\cdot}{\partial \bar{y}^2}].$$

It is noted that the coefficients of the first order derivatives correspond with the deterministic vector field (3.2). The operator L plays an important role in exit problems. Let us consider a domain Ω with the boundary $\partial\Omega$ composed of two disjoint parts $\partial\Omega_0$ and $\partial\Omega_1$. We are interested in the probability $u(z)$ that the state of the system reaches $\partial\Omega$ the first time at the part $\partial\Omega_1$, if starting in $z \in \Omega$. This function $u(z)$ satisfies the Dirichlet problem

$$Lu = 0 \quad \text{in} \quad \Omega, \tag{4.9a}$$

$$u = 0 \quad \text{at} \quad \partial\Omega_0, \qquad u = 1 \quad \text{at} \quad \partial\Omega_1, \tag{4.9b}$$

see Schuss (1980). Eqs. (4.9) will be used in section 5 to formulate a boundary value problem for the probability that the disease will die out after a major outbreak has occurred.

5. The probability of extinction after a major outbreak

In the stochastic model the disease can disappear from the population because of stochastic fluctuations. In this section we study the probability that the disease will die out at the end of a major outbreak. In particular, we are interested in the probability of fade-out after a major outbreak following the introduction of one or a few infectives into a disease free population. In the development of the disease several stages can be distinguished. In the initial stage after the introduction of one or a few infectives it is possible that the disease directly dies out. In x, $\tilde{y}$-coordinates this stage corresponds with the state of the stochastic system remaining in the neighbourhood of the deterministic saddle point (1, 0), where diffusion plays an essential role. The probability of extinction of the disease in this stage of the process is studied in Van Herwaarden and Grasman (1995). In case the disease has survived this initial period it is likely that a major outbreak of the disease takes place. This part of the process is dominated by the deterministic vector field. For small values of μ the deterministic trajectory starting at (1, 0) closely approaches the x-axis at the end of its first cycle, see section 3. Here diffusion again plays an important role and extinction of the infection can occur at this stage of the process. We will formulate and solve a boundary value problem for the probability that the disease dies out during this part of the process. From the solution we obtain an expression for the probability of extinction of the disease at the end of a major outbreak. In our asymptotic analysis we assume that $0 < 1/\sqrt{N} << \mu << 1$.

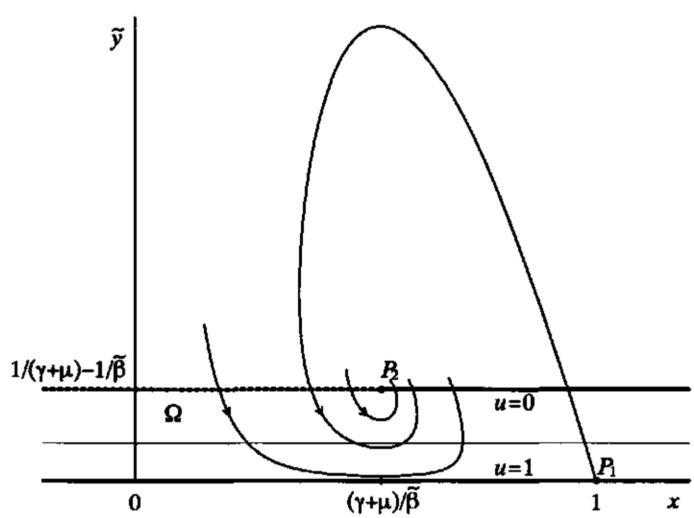


Figure 3. *Domain Ω for the boundary value problem (5.3), divided in boundary layer and outer region. The solution approximates the probability that the disease dies out before a major outbreak occurs.*

Boundary value problem. In the x, $\tilde{y}$-plane we consider the domain

$$\Omega = \{(x, \tilde{y}) \mid 0 < \tilde{y} < \frac{1}{\gamma + \mu} - \frac{1}{\tilde{\beta}}\}, \tag{5.1}$$

see Figure 3. On Ω we can approximate Lu, see (4.8), by

$$\tilde{L}u \equiv (1 - \tilde{\beta}x\tilde{y} - x)\frac{\partial u}{\partial x} + \frac{1}{\mu}(\tilde{\beta}x - \gamma - \mu)\tilde{y}\frac{\partial u}{\partial \tilde{y}} + \frac{1}{2N\mu^2}(\tilde{\beta}x + \gamma + \mu)\tilde{y}\frac{\partial^2 u}{\partial \tilde{y}^2}. \tag{5.2}$$

We now consider the parabolic boundary value problem

$$\tilde{L}u = 0 \quad \text{in} \quad \Omega, \tag{5.3a}$$

$$u = 1 \quad \text{for} \quad \tilde{y} = 0, \qquad u = 0 \quad \text{for} \quad \tilde{y} = \frac{1}{\gamma + \mu} - \frac{1}{\tilde{\beta}}, \quad x > \frac{\gamma + \mu}{\tilde{\beta}}. \tag{5.3b}$$

The solution $u(x, \tilde{y})$ of this problem approximates the probability that the state of the system reaches the boundary of Ω the first time at $\tilde{y} = 0$, if starting in $(x, \tilde{y}) \in \Omega$. We asymptotically solve (5.3) for $0 < x < 1$, where the solution $u(x, \tilde{y})$ yields an approximation for the probability that the disease dies out before a major outbreak takes place. We are in particular interested in the value of $u(x, \tilde{y})$ on the deterministic trajectory starting in (1, 0), at the part where it enters Ω at the end of its first cycle, that is where segment B (see Figure 2) enters Ω. We note that the stochastic state is expected to enter Ω after a major outbreak along this trajectory. This value of u can be used as an approximation for the probability of extinction of the disease directly at the end of the major outbreak.

Solution of the diffusion equation. Boundary layer analysis (see Kevorkian and Cole (1981)) reveals the existence of a boundary layer of width $O(1/(\mu N))$ along the x-axis. In the domain Ω outside the boundary layer $\tilde{L}u = 0$ reduces to

$$(1 - \tilde{\beta}x\tilde{y} - x)\frac{\partial u}{\partial x} + \frac{1}{\mu}(\tilde{\beta}x - \gamma - \mu)\tilde{y}\frac{\partial u}{\partial \tilde{y}} = 0, \tag{5.4}$$

which is satisfied by u = constant along the segments of the deterministic trajectories contained in this outer region. In particular, $u = 0$ along the segments of the deterministic trajectories that meet the part of the boundary of Ω with boundary condition $u = 0$.

Inside the boundary layer we introduce the local coordinate

$$\eta = \mu N\tilde{y}. \tag{5.5}$$

Substitution of (5.5) in (5.3) yields for $N \to \infty$

$$(1 - x)\frac{\partial u}{\partial x} + \frac{1}{\mu}(\tilde{\beta}x - \gamma - \mu)\eta\frac{\partial u}{\partial \eta} + \frac{1}{2\mu}(\tilde{\beta}x + \gamma + \mu)\eta\frac{\partial^2 u}{\partial \eta^2} = 0 \quad (5.6a)$$

with the boundary condition

$$u(x, \eta) = 1 \quad \text{for} \quad \eta = 0. \quad (5.6b)$$

Matching the solution with the outer solution $u = 0$ along the deterministic trajectories, which have the form (3.6), leads to the matching condition

$$u(x, \eta) = 0 \quad \text{for} \quad \eta \exp[(\tilde{\beta}x + (\tilde{\beta} - \gamma)\ln(1 - x))/\mu - \ln(1 - x)] = \text{const},$$

$$\eta \to \infty \quad \text{and} \quad x > \frac{\gamma + \mu}{\tilde{\beta}}. \quad (5.6c)$$

Introduction of new coordinates

$$\tau = \frac{1}{N\mu^2}\int_x^1 (\tilde{\beta}s + \gamma + \mu)\exp[(\tilde{\beta}s + (\tilde{\beta} - \gamma)\ln(1 - s))/\mu - 2\ln(1 - s)]\,ds, \quad (5.7a)$$

$$\zeta = \frac{2}{N\mu}\eta\exp[(\tilde{\beta}x + (\tilde{\beta} - \gamma)\ln(1 - x))/\mu - \ln(1 - x)] \quad (5.7b)$$

transforms (5.6) into the initial value problem on the domain $\{(\tau, \zeta) \mid \tau > 0, \zeta > 0\}$

$$\frac{\partial u}{\partial \tau} = \zeta\frac{\partial^2 u}{\partial \zeta^2}, \quad (5.8a)$$

$$u(0, \zeta) = 0, \qquad u(\tau, 0) = 1 \quad (5.8b,c)$$

with solution

$$u(\tau, \zeta) = \exp[-\zeta/\tau]. \quad (5.9)$$

Bringing this solution in x, η-coordinates we obtain

$$u(x, \eta) = \exp[-2\mu\eta\Phi(x)] \quad (5.10a)$$

with

$$\Phi(x) = \frac{\exp[(\tilde{\beta}x + (\tilde{\beta} - \gamma)\ln(1 - x))/\mu - \ln(1 - x)]}{\int_x^1 (\tilde{\beta}s + \gamma + \mu)(1 - s)^{(\tilde{\beta} - \gamma - 2\mu)/\mu}\exp[\tilde{\beta}s/\mu]\,ds}. \quad (5.10b)$$

As mentioned above we are in particular interested in the value of u where the deterministic trajectory starting at (1, 0) enters Ω at the end of its first cycle, that is where segment B of the deterministic trajectory (see Figure 2) enters Ω. This value is given by the (constant) value of u along the part of the deterministic trajectory contained in the outer region. It is obtained by matching the boundary layer solution (5.10) with the outer solution u = constant along this deterministic trajectory. Inside the boundary layer the deterministic trajectory is given by

$$\eta \exp[(\tilde{\beta}x + (\tilde{\beta} - \gamma)\ln(1 - x))/\mu - \ln(1 - x)] = KN\mu, \tag{5.11}$$

see (3.25). Matching along this trajectory for $\eta \to \infty$, $x < (\gamma + \mu)/\tilde{\beta}$ yields for the outer solution the value

$$u = \exp[-2KN\mu^2/\int_{-\infty}^{1} (\tilde{\beta}s + \gamma + \mu)(1 - s)^{(\tilde{\beta} - \gamma - 2\mu)/\mu} e^{\tilde{\beta}s/\mu} ds]. \tag{5.12}$$

The integral in (5.12) can be expressed in a gamma function. We thus obtain the following approximation for the probability of extinction of the disease at the end of a major outbreak following the introduction of one or a few infectives into the population, given that a major outbreak occurs,

$$u = \exp[-KN\mu^2(\tilde{\beta}/\mu)^{(\tilde{\beta} - \gamma - \mu)/\mu} e^{-\tilde{\beta}/\mu}/\{(\gamma + \mu)\Gamma((\tilde{\beta} - \gamma - \mu)/\mu)\}] \tag{5.13}$$

with K given by (3.25b,c,d).

6. A comparison with numerical results

In this section some numerical results are presented for the models described above. In particular, the analytical approximation (5.13) for the probability of fade-out at the end of the first epidemic cycle is compared with results of stochastic simulations. For the parameters we use the values $\tilde{\beta} = 1$, $\gamma = 0.8$, $\mu = 0.01$ and $N = 50000$. The initial numbers of infectives and susceptibles in the population are $I(0) = 10$ and $S(0) = 49990$, respectively.

Deterministic trajectories. By numerically solving the deterministic system (2.2a,b) with the above initial values we obtain the trajectory depicted in Figure 4a. This deterministic trajectory starts close to the saddle point P_1. In Figure 4b the number of infectives I in the deterministic system is given as a function of time t. The periods with small numbers of infectives are large compared with the periods needed to complete the epidemic peaks.

Stochastic simulation. For the values of the parameters given above we have carried out simulations for the discrete stochastic model of section 4 by following the path of the state (S, I), starting in $(S(0), I(0))$. For this initial state 10000 simulation runs have been made. For each run the extinction time T_{ext} of the disease has been determined. Results are given in the histogram of Figure 4c.

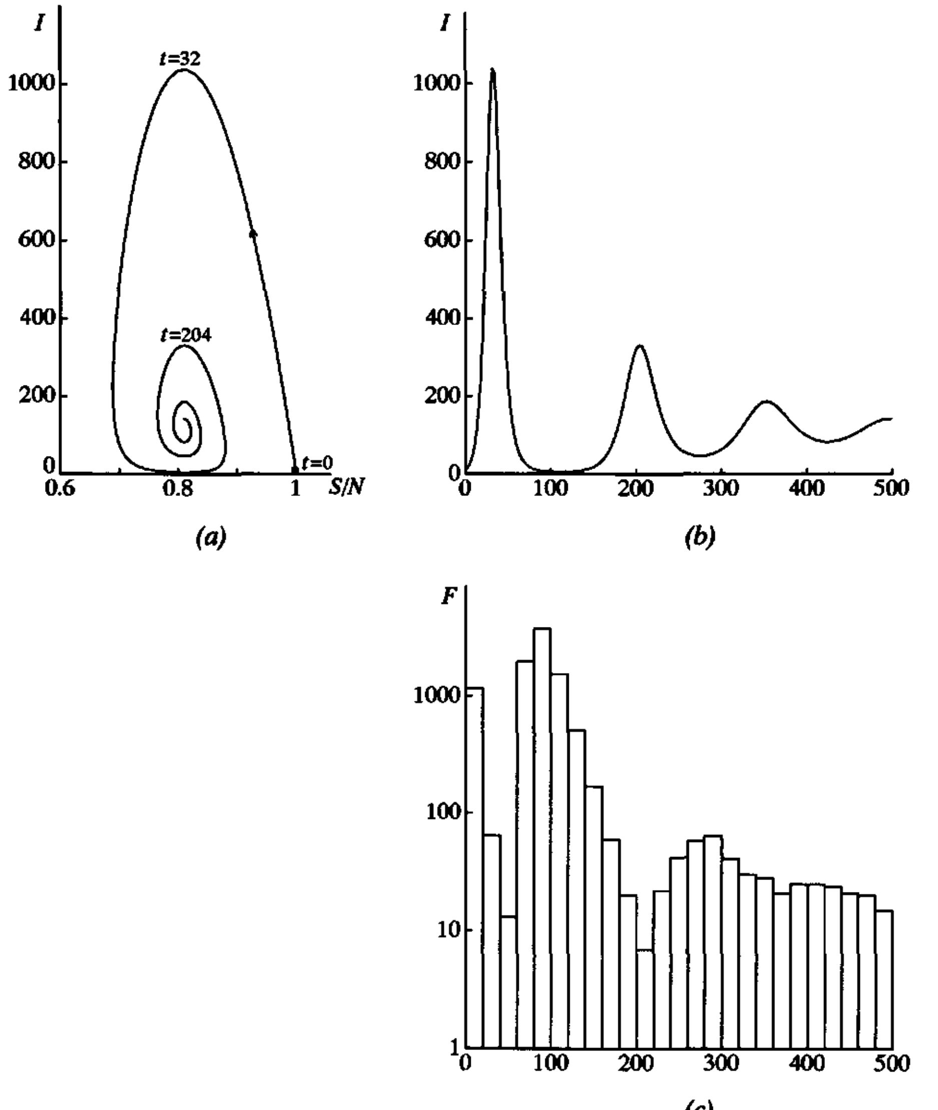


Figure 4. Occurrence of epidemic peaks in the deterministic system compared with extinction of the disease in the stochastic system for the initial state $I(0) = 10$, $S(0) = 49990$; $\mu = 0.01$, $\tilde{\beta} = 1$, $\gamma = 0.8$, $N = 50000$. The total number of simulation runs is 10000. (a) Deterministic trajectory starting near P_1. (b) Number of infectives from deterministic model. (c) Number of simulation runs F with extinction in the specified time intervals. Not shown are 276 simulation runs with $T_{ext} > 500$.

Probability that a major outbreak does not occur. Extinction of the disease can take place in the initial period after its introduction into the population. Analytical approximations for this extinction probability are given by Van Herwaarden and Grasman (1995). Using the expression given in that paper (cf. Goel and Richter-Dyn (1974), Kendall (1956))

$$U(S(0), I(0)) \sim \left(\frac{\gamma + \mu}{\beta S(0)/N}\right)^{I(0)}, \tag{6.1}$$

we obtain the probability of extinction in the initial stage

$$\Pr\{T_{ext} < 32\} = 0.1218 \qquad (0.1215). \tag{6.2}$$

The value between brackets is obtained from the stochastic simulations. In the histogram a peak can be distinguished corresponding with extinction in the initial period.

Probability of epidemic fade-out. If a major outbreak takes place, the disease can die out at the end of the outbreak. Using expression (5.13) we obtain the (conditional) probability of extinction

$$\Pr\{32 < T_{ext} < 204\} = 0.89 \qquad (0.92). \tag{6.3}$$

The value between brackets, obtained from the stochastic simulations, is in good accordance with the analytical result. We note that in the histogram a peak can be distinguished corresponding with extinction in this period, as well as a (smaller) peak corresponding with fade-out at the end of the second cycle.

7. Extinction probability for other initial states

Until now we have mainly dealt with a population in which the infection has just been introduced. We may also be interested in the extinction probability of the disease for other initial situations, e.g., when the disease has survived the first epidemic cycle. In that case the state is expected to arrive in the neighbourhood of segment D, see Figure 2. We will show how to determine the extinction probability if the state is found in the neighbourhood of that segment.

Say, the state has arrived in $(x_0, \bar{y}_0)$ near D. We first determine local expansions for the deterministic trajectory starting in $(x_0, \bar{y}_0)$. In particular we need the expansion near the x-axis. Proceeding along the lines of section 3, we divide the trajectory in segments $\tilde{D}$, $\tilde{E}$, $\tilde{F}$, $\tilde{G}$, with $(x_0, \bar{y}_0) \in \tilde{D}$, segment $\tilde{E}$ in the region $\bar{y} >> (1 - x)/(\bar{\beta}x)$, segment $\tilde{G}$ in the region $0 < \bar{y} << (1 - x)/(\bar{\beta}x)$, and the segments $\tilde{D}$ and $\tilde{F}$ in the intermediate regions. The expansion for segment $\tilde{D}$ is given by

$$x(\bar{y}) = \bar{D}_1 + \bar{D}_2 \mu \ln \mu + \left(- \frac{\bar{\beta}\bar{D}_1}{\bar{\beta}\bar{D}_1 - \gamma}\bar{y} + \frac{1 - \bar{D}_1}{\bar{\beta}\bar{D}_1 - \gamma} \ln \bar{y} + \bar{D}_3\right)\mu + ..., \tag{7.1}$$

as follows from (3.5). By substitution of $(x_0, \bar{y}_0)$ in (3.3b) and in the derivative $dx/d\bar{y}$ of (7.1) the constant $\bar{D}_1$ can be determined. Then using (7.1) we can obtain $\bar{D}_2 \ln \mu + \bar{D}_3$. The expansions for the segments $\bar{E}$, $\bar{F}$ and $\bar{G}$ are given by (3.4), (3.5) and (3.6) with the constants A_i, B_i and C_i replaced by $\bar{E}_i$, $\bar{F}_i$ and $\bar{G}_i$, respectively. The matching relations between the constants are given by (3.20), (3.21) and (3.22) with the appropriate renaming of the constants: $\bar{D}_i$, $\bar{E}_i$, $\bar{F}_i$ and $\bar{G}_i$ for D_i, E_i, F_i and G_i; $x_{r\bar{E}}$ and $x_{l\bar{E}}$ for x_{rE} and x_{lE}, respectively. Inspection of these relations and of the form of the expansion (3.6) for segment $\bar{G}$ shows that knowledge of $\bar{D}_1$ and $\bar{D}_2 \ln \mu + \bar{D}_3$ suffices to determine the expansion for segment $\bar{G}$. We obtain the following result for the expansion for segment $\bar{G}$

$$\bar{y}(x) = \bar{K} \exp[(-\bar{\beta}x - (\bar{\beta} - \gamma)\ln(1 - x))/\mu + \ln(1 - x)] \tag{7.2a}$$

with

$$\bar{K} = \exp[(\bar{\beta}x_{l\bar{E}} + (\bar{\beta} - \gamma)\ln(1 - x_{l\bar{E}}))/\mu + \bar{G}_2 \ln \mu + \bar{G}_3], \tag{7.2b}$$

where

$$x_{l\bar{E}} \neq x_{r\bar{E}} = \bar{D}_1 \tag{7.2c}$$

satisfies

$$-x_{l\bar{E}} + \frac{\gamma}{\bar{\beta}} \ln x_{l\bar{E}} + x_{r\bar{E}} - \frac{\gamma}{\bar{\beta}} \ln x_{r\bar{E}} = 0, \tag{7.2d}$$

$$\bar{G}_2 = - \frac{x_{l\bar{E}}(\bar{\beta}x_{r\bar{E}} - \gamma)}{x_{r\bar{E}}(1 - x_{l\bar{E}})}\bar{D}_2 + \frac{x_{l\bar{E}} - x_{r\bar{E}}}{x_{r\bar{E}}(1 - x_{l\bar{E}})}, \tag{7.2e}$$

$$\begin{aligned}\bar{G}_3 = &- \ln \frac{\bar{\beta}^2 x_{l\bar{E}}(1 - x_{l\bar{E}})}{(\bar{\beta}x_{l\bar{E}} - \gamma)^2} - \frac{x_{l\bar{E}}(\bar{\beta}x_{r\bar{E}} - \gamma)}{x_{r\bar{E}}(1 - x_{l\bar{E}})}\bar{D}_3 + \\ &+ \frac{x_{l\bar{E}}(1 - x_{r\bar{E}})}{x_{r\bar{E}}(1 - x_{l\bar{E}})} \ln \frac{\bar{\beta}^2 x_{r\bar{E}}}{(\bar{\beta}x_{r\bar{E}} - \gamma)^2} + \frac{\bar{\beta}x_{l\bar{E}}}{1 - x_{l\bar{E}}}(J(x_{r\bar{E}}) - J(x_{l\bar{E}}))\end{aligned} \tag{7.2f}$$

with $J(x_{r\bar{E}})$ and $J(x_{l\bar{E}})$ given by (3.10c) for $\alpha = x_{r\bar{E}} - (\gamma/\bar{\beta})\ln x_{r\bar{E}}$. We can now determine the probability of extinction of the disease at the end of the cycle starting in $(x_0, \bar{y}_0)$. We note that (7.2a) has the same form as (3.25a) for the deterministic trajectory starting near P_1. Proceeding along the lines of section 5 we obtain the approximation (5.13) with K replaced by $\bar{K}$, where $\bar{K}$ is given by (7.2b-f).

For other initial states the extinction probability can be determined in a similar way. By monitoring the state of the system this can be used to forecast if the disease will die out at the end of the following cycle.

8. Conclusions

In this paper we have studied a two-dimensional stochastic system modelling the spread of an infectious disease. In this model the population is divided in three classes: susceptibles, infectives and removed. We have been dealing with an approximation of the probability of an epidemic fade-out. For that purpose we have considered a small renewal rate of the population, which causes the susceptible population to be restored only slowly after a major outbreak. Extinction of the disease by stochastic fluctuations is probable at that stage of the process. The case studied in this paper with the range of renewal rate values such that the extinction probability varies from close to zero to close to one, can be seen as a transitional case. It connects the Kermack and McKendrick model with the deterministic *SIR*-model with renewal. Due to the risk of fade-out, the model is inherently stochastic.

For the probability of extinction of the disease at the end of a major outbreak following the introduction of one or a few infectives into the population we have derived an analytical approximation, see Eq. (5.13). In deriving this expression several approximations have been made. It is expected that the inaccuracies caused by the asymptotic approximation of the deterministic trajectories are negligible compared with the inaccuracies introduced by the diffusion approximation of the Fokker-Planck equation in section 5. In computing the extinction probability we need the value of x_{I4}, which has to be calculated by numerical methods. We remark that x_{I4} represents the final fraction of susceptibles in the corresponding deterministic Kermack and McKendrick model (renewal rate $\mu = 0$).

Finally, it is noted that there is a good correspondence between the results obtained by stochastic simulations and the asymptotic approximation derived for the probability of fade-out at the end of the first epidemic cycle.

Acknowledgements

I want to thank Professor Johan Grasman for the very stimulating discussions and useful remarks on the subject treated in this paper.

References

Anderson, R.M. and R.M. May (1979), *Population biology of infectious diseases*, Part I, Nature, 280, pp. 361-367.

Anderson, R.M. and R.M. May (1991), *Infectious Diseases of Humans*, Oxford University Press, Oxford.

Bailey, N.T.J. (1975), *The Mathematical Theory of Infectious Diseases and its Applications*, 2nd edn, Griffin, London.

Bartlett, M.S. (1960), *Stochastic Population Models in Ecology and Epidemiology*, Methuen, London.

Edelstein-Keshet, L. (1988), *Mathematical Models in Biology*, Random House, New York.

Gardiner, C.W. (1983), *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, Springer-Verlag, Berlin.

Goel, N.S. and N. Richter-Dyn (1974), *Stochastic Models in Biology*, Academic Press, New York.

Grasman, J. and E. Veling (1973), *An asymptotic formula for the period of a Volterra-Lotka system*, Math. Biosc., 18, pp. 185-189.

Kendall, D.G. (1956), *Deterministic and stochastic epidemics in closed populations*, Proc. Symp. Math. Stat. Prob., 3rd, Berkeley, 4.

Kevorkian, J. and J.D. Cole (1981), *Perturbation Methods in Applied Mathematics*, Springer-Verlag, New York.

Ludwig, D. (1975), *Persistence of dynamical systems under random perturbations*, SIAM Rev., 17, pp. 605-640.

Schuss, Z. (1980), *Theory and Applications of Stochastic Differential Equations*, Wiley, New York.

Van Herwaarden, O.A. and J. Grasman (1995), *Stochastic epidemics: major outbreaks and the duration of the endemic period*, J. Math. Biol., 33, pp. 581-601.

Chapter 5

The expected exit time of unexpected exits

Abstract

A study is made of the expected exit time from an interval for a one-dimensional stochastic dynamical system at precisely that boundary where exit is not likely. For three fundamental cases the conditional expected exit time is determined by asymptotically solving Dirichlet problems based on the Fokker-Planck equation for the stochastic system for small values of the diffusion parameter. The asymptotic approximations reveal interesting features that remain hidden in the exact solutions. As the usual method of matched asymptotic expansions with local solutions of boundary layer type fails, use has been made of WKB-expansions. The analytical expressions are compared with simulation results and with numerical values for the exact solutions.

1. Introduction

In this study we consider stochastic dynamical systems with small stochastic fluctuations. These fluctuations may drive the system against the deterministic flow and considerably influence its long time behaviour. Because of the stochasticity, the system can leave the domain of attraction of a stable steady state. This exit problem has been the subject of many investigations, see e.g. Naeh et al. (1990) and the references mentioned in that paper. Because of stochastic fluctuations, the system can also leave a domain in the state space at a part of the boundary that can not be reached in the corresponding deterministic system. In this study we are interested in the (conditional) expected exit time at precisely that part of the boundary of the domain where exit is not likely. This problem came up in studying the expected arrival time of pollution at a well in a dispersive groundwater flow, see Van Herwaarden (1994), for starting points outside the region of advective flow towards the well. This problem is also considered in the appendix of Mangel (1979), where the deterministic system of interest contains two stable steady states and a saddle point. Mangel focuses his attention on the bifurcation behaviour as two or three of the steady states coalesce.

In the present paper we determine asymptotic solutions for the expected (first) exit time at one of the boundaries for three fundamental one-dimensional cases. The asymptotic solutions reveal essential features of the expected exit time that remain hidden in the exact solutions. Moreover, the methods we present may be extended to higher dimensional problems where exact solutions are not available. We describe the stochastic process in terms of a Fokker-Planck (or Kolmogorov) equation. In section 2 the equations and the boundary conditions for the exit problem are formulated for exit of a stochastic process from the interval $D = \{x \in \mathbb{R} \mid 0 < x < 1\}$ at the boundary $x = 0$. In the next sections we determine asymptotic solutions for the (conditional) expected exit time at $x = 0$. In section 3 the underlying deterministic flow is directed towards $x = 1$, so in the stochastic process the boundary $x = 0$ is the unlikely exit point for initial states away from this boundary. An asymptotic formula for the expected exit time is obtained that is simple and intriguing. In sections 4 and 5 the deterministic flow vanishes in one point of D. The deterministic steady state in section 4 is unstable. In section 5 we consider a stable steady state in D that is chosen such that the boundary $x = 0$ is the unlikely point of exit for initial states away from this boundary. In section 6 the asymptotic expressions are compared with numerical results obtained from random walk simulations or from the exact solutions.

2. The Fokker-Planck equation

We consider a stochastic dynamical system on the domain

$$D = \{x \in \mathbb{R} \mid 0 < x < 1\}. \tag{2.1}$$

Let $p(x, t)$ be the probability density function to find the state of the system at a point x at time t, if initially it is in $\bar{x} \in D$. The function $p(x, t)$ satisfies the forward Fokker-Planck (or forward Kolmogorov) equation

$$\frac{\partial p}{\partial t} = Mp, \qquad p(x, 0) = \delta(x - \bar{x}), \tag{2.2a}$$

$$M = -\frac{\partial}{\partial x}(b(x)\,.\,) + \tfrac{1}{2}\varepsilon^2 \frac{\partial^2}{\partial x^2}(a(x)\,.\,), \tag{2.2b}$$

where $b(x)$ and $\varepsilon^2 a(x)$ are the drift and diffusion coefficients, respectively. The parameter ε is small, $0 < \varepsilon << 1$, indicating that the stochastic fluctuations are small. We suppose that $a(x)$ is a positive function. The system can also be described by a stochastic differential equation of Ito type

$$dx = b(x)dt + \varepsilon\sqrt{a(x)}\,dW, \tag{2.3a}$$

$$x(0) = \tilde{x} \tag{2.3b}$$

with $W(t)$ a Wiener process, see Gardiner (1983). A discretised version of the stochastic differential equation can be used to generate simulations for the stochastic process, see section 6.

Related to the forward operator M is the backward operator

$$L = b(x)\frac{\partial\,\cdot}{\partial x} + \tfrac{1}{2}\varepsilon^2 a(x)\frac{\partial^2\,\cdot}{\partial x^2}, \tag{2.4}$$

which is the formal adjoint of M. It plays an important role in exit problems. We use L to formulate boundary value problems for the exit of the system through a particular end of the interval D. Say the system starts in $x \in D$, what is the probability that it reaches the boundary ∂D the first time in $x = 0$? Let $u(x)$ denote this probability. It can be shown that $u(x)$ satisfies

$$Lu = b(x)\frac{du}{dx} + \tfrac{1}{2}\varepsilon^2 a(x)\frac{d^2u}{dx^2} = 0 \quad \text{in} \quad D, \tag{2.5a}$$

$$u(0) = 1, \quad u(1) = 0. \tag{2.5b}$$

For an initial state $x \in D$ we are in particular interested in the expected (first) exit time $T_1(x)$ from D with the condition that exit takes place at $x = 0$. It can be derived, see Gardiner (1983), that

$$T_1(x) = T(x)/u(x) \tag{2.6}$$

with the function $T(x)$ satisfying the boundary value problem

$$LT = b(x)\frac{dT}{dx} + \tfrac{1}{2}\varepsilon^2 a(x)\frac{d^2T}{dx^2} = -u \quad \text{in} \quad D, \tag{2.7a}$$

$$T(0) = 0, \quad T(1) = 0. \tag{2.7b}$$

These boundary value problems can be extended to higher dimensional state spaces. It is noted that in Mangel (1979) the denominator $u(x)$ in (2.6) is not present.

3. Flow without steady states

In this section we consider a stochastic system on the domain D given by (2.1). We suppose that the deterministic flow

$$\frac{dx}{dt} = b(x) \tag{3.1}$$

is a positive function independent of ε:

$$b(x) > 0 \quad \text{for} \quad x \in D, \tag{3.2}$$

so the flow is pointing towards $x = 1$ and has no equilibria on D. Because the fluctuations we consider are small, the boundary point $x = 1$ instead of $x = 0$ is most likely reached for starting points away from $x = 0$. The solutions $u(x)$ and $T_1(x)$ of Eqs. (2.5) - (2.7) represent the probability of exit and the (conditional) expected exit time, respectively, at the unlikely exit point $x = 0$ for starting points $x \in D$. We are interested in deriving an asymptotic approximation for $T_1(x)$ for small values of ε.

3.1 An asymptotic expansion of the exact solution

In this subsection we derive an asymptotic approximation for $T_1(x)$ by expanding the exact solution with the method of Laplace (Erdélyi (1956)). Eqs. (2.5a) and (2.7a) are first order differential equations in du/dx and dT/dx, respectively, and can be solved exactly. Defining

$$I(x) = \int_0^x \frac{2b(s)}{a(s)}\, ds, \tag{3.3}$$

the solution $u(x)$ of the boundary value problem (2.5) can be written as

$$u(x) = \frac{1}{R}\int_x^1 \exp[-I(s)/\varepsilon^2]\, ds \tag{3.4a}$$

with the constant R given by

$$R = \int_0^1 \exp[-I(s)/\varepsilon^2]\, ds. \tag{3.4b}$$

For small values of ε the main contribution to the integral in (3.4a) comes from a neighbourhood of the minimum of $I(s)$, where the integrand is peaked. For the deterministic flow $b(x)$ considered in this section, this minimum is found in the end point $s = x$ of the integration interval. Applying Laplace's method, by expanding $I(s)$ in a Taylor series near $s = x$, the integral in (3.4a) is approximated. Treating the integral (3.4b) for R in the same way, we obtain for x away from 1

$$u(x) \sim \frac{b(0)}{a(0)}\frac{a(x)}{b(x)} \exp[-I(x)/\varepsilon^2]. \tag{3.5}$$

The solution of (2.7) is given by

$$T(x) = \frac{2}{\varepsilon^2}\int_x^1 \exp[-I(s)/\varepsilon^2]\int_0^s \frac{u(t)}{a(t)}\exp[I(t)/\varepsilon^2]\,dt\,ds +$$

$$-\frac{1}{R}\frac{2}{\varepsilon^2}\int_x^1 \exp[-I(s)/\varepsilon^2]\,ds\int_0^1 \exp[-I(s)/\varepsilon^2]\int_0^s \frac{u(t)}{a(t)}\exp[I(t)/\varepsilon^2]\,dt\,ds. \quad (3.6)$$

After some manipulations $T(x)$ can be written in a form that is more convenient to expand for the present flow field $b(x)$:

$$T(x) = \frac{1}{R}\frac{2}{\varepsilon^2}\int_x^1\int_0^1 f(s, r)\exp[-(I(s) + I(r))/\varepsilon^2]\,dr\,ds \quad (3.7a)$$

with

$$f(s, r) = \int_r^s \frac{u(t)}{a(t)}\exp[I(t)/\varepsilon^2]\,dt. \quad (3.7b)$$

The exponential function in (3.7a) is peaked in $(s, r) = (x, 0)$, whereas $f(s, r)$ behaves regularly, see Eq. (3.5). Therefore, the largest contribution to the double integral in $T(x)$ comes from a vicinity of $(x, 0)$. Applying Laplace's method for double integrals and using the approximation for R, we find

$$T(x) \sim \frac{b(0)}{a(0)}\frac{a(x)}{b(x)}\int_0^x \frac{1}{b(s)}\,ds\ \exp[-I(x)/\varepsilon^2], \quad (3.8)$$

and, using $T_1(x) = T(x)/u(x)$ and (3.5),

$$T_1(x) \sim \int_0^x \frac{1}{b(s)}\,ds \quad (3.9)$$

for x away from 1. This simple result for the expected exit time at $x = 0$ for a starting point $x \in D$, is very intriguing. We note that the integral in (3.9) also represents the deterministic travel time from $x = 0$ to the point $x \in D$, as can easily be derived from (3.1). So the first order approximation for the expected exit time for diffusion against the flow equals the travel time with the deterministic flow in opposite direction. We are not familiar with an explanation for this relation from a physical point of view. It is seen that, if we let the deterministic flow increase, the expected exit time for diffusion against the flow decreases. This apparent contradiction is taken away by considering that T_1 is a conditional expected exit time. For an increasing deterministic flow, less particles reach the boundary $x = 0$. An interpretation is that, if we let the deterministic flow increase, the "average particle" that reaches the boundary point $x = 0$ has to travel against the flow faster in order to be able to reach that boundary at all.

We note that the result (3.9) is in agreement with the following result that can be derived for a stochastic process with constant drift μ and variance σ^2 and one absorbing barrier at $a > 0$. For the initial state $x = 0$ the moment generating function of the first passage time at a is given by

$$g^*(s) = \int_0^\infty g(t)\exp[-st]\,dt = \exp[a(\mu - \sqrt{\mu^2 + 2s\sigma^2})/\sigma^2], \tag{3.10}$$

see Cox and Miller (1965), with $g(t)$ the (possibly defective) probability density function of the first passage time. Expression (3.10) is valid both for drift towards the barrier ($\mu > 0$) and for drift away from the barrier ($\mu < 0$). The conditional expected exit time T_a at the barrier a is given by $-d(\ln g^*)/ds$, evaluated for $s = 0$, and can easily be calculated. The result is

$$T_a = \frac{a}{|\mu|}. \tag{3.11}$$

Expression (3.11) also equals the deterministic travel time from $x = 0$ to $x = a$ with drift $\mu > 0$ or vice versa from $x = a$ to $x = 0$ with drift $\mu < 0$. So we see that in this case with one absorbing barrier the conditional expected exit time for diffusion against the flow exactly equals the travel time with the deterministic flow in opposite direction. From (3.9) we conclude that adding a second absorbing barrier can be seen as the introduction of a perturbation that does not affect the first order approximation for the expected exit time for diffusion against the flow. We note that Cox and Miller (1965) state the result (3.11) explicitly only for $\mu > 0$. Therefore, the remarkable fact that the result also holds for diffusion against the flow has remained unnoticed. Moreover, we note that the first order approximation (3.9) for $T_1(x)$ is independent of the diffusion coefficient $a(x)$.

3.2 An asymptotic treatment of the differential equations

In this subsection we derive an asymptotic approximation for T_1 without using the exact solution. We consider the boundary value problem (2.7). Substituting $T_1(x) = T(x)/u(x)$, see (2.6), into (2.7a) and using (2.5a), we obtain the differential equation for $T_1(x)$

$$(b(x) + \varepsilon^2 a(x)\frac{u'(x)}{u(x)})\frac{dT_1}{dx} + \tfrac{1}{2}\varepsilon^2 a(x)\frac{d^2T_1}{dx^2} = -1. \tag{3.12a}$$

This equation must be supplemented by boundary conditions. It is seen from the boundary conditions (2.5b) for $u(x)$ that we impose sufficient conditions by requiring

$$T_1(0) = 0, \quad T_1(1) < \infty. \tag{3.12b}$$

In our asymptotic treatment of (3.12) we use an approximation for $u(x)$ that is obtained by asymptotically solving (2.5). We note that solving (2.5) by the familiar method of matched asymptotic expansions with regular expansions on subdomains (see, e.g., Eckhaus (1973)), yields an approximation for $u(x)$ that is only sufficiently accurate in a boundary layer near $x = 0$ to be used for the determination of $T_1(x)$. Therefore, we use a WKB-expansion for $u(x)$. Substitution of the WKB-approximation

$$u(x) = w(x)\exp[-Q(x)/\varepsilon^2] \tag{3.13}$$

into (2.5a) yields to leading order $O(\varepsilon^{-2})$

$$w(x)Q'(x)(-b(x) + \tfrac{1}{2}a(x)Q'(x)) = 0 \tag{3.14}$$

and to order $O(\varepsilon^0)$

$$(b(x) - a(x)Q'(x))w'(x) - \tfrac{1}{2}a(x)Q''(x)w(x) = 0. \tag{3.15}$$

Eqs. (3.14) and (3.15) are satisfied by $Q(x) = I(x)$ and $w(x) = a(x)/b(x)$. Using the left boundary condition $u(0) = 1$ we obtain the WKB-approximation

$$u(x) = \frac{b(0)}{a(0)}\frac{a(x)}{b(x)}\exp[-I(x)/\varepsilon^2], \tag{3.16}$$

which equals the approximation (3.5) found by expanding the exact solution. We note that the right boundary condition $u(1) = 0$ is satisfied with an exponentially small error.

Substitution of (3.16) into (3.12a) yields

$$\left(-b(x) + \varepsilon^2\frac{a'(x)b(x) - a(x)b'(x)}{b(x)}\right)\frac{dT_1}{dx} + \tfrac{1}{2}\varepsilon^2 a(x)\frac{d^2T_1}{dx^2} = -1, \tag{3.17}$$

valid away from the right boundary point $x = 1$. For the present flow field $b(x)$ it is possible to find a regular expansion in this part of the domain

$$T_1(x) = \sum_{n=0}^{\infty} \varepsilon^{2n} T_{1(n)}(x). \tag{3.18}$$

Substitution of this expansion in (3.17) yields the differential equation for the order $O(1)$ term

$$-b(x)\frac{dT_{1(0)}}{dx} = -1. \tag{3.19}$$

Using the boundary condition $T_{1(0)}(0) = 0$, see (3.12b), we obtain the first order approximation

$$T_{1(0)}(x) = \int_0^x \frac{1}{b(s)}\,ds, \tag{3.20}$$

that is the same approximation as we have found above by asymptotically expanding the exact solution, see (3.9). Higher order terms of the regular expansion can be successively obtained in the same way. It is noted that the expression (3.20) can also be used as a first order approximation for $T_1(x)$ in the neighbourhood of $x = 1$.

4. Flow with an unstable steady state

We again consider a one-dimensional stochastic system on the domain D given by (2.1). In this section we assume that the corresponding deterministic flow $b(x)$ has one unstable equilibrium x_e in D:

$$b(x_e) = 0, \quad b(x) < 0 \quad \text{for} \quad 0 \le x < x_e, \quad b(x) > 0 \quad \text{for} \quad x_e < x \le 1. \qquad (4.1)$$

As before, the solutions $u(x)$ and $T_1(x)$ of Eqs. (2.5) - (2.7) represent the probability of exit and the (conditional) expected exit time, respectively, at the boundary $x = 0$ for starting points $x \in D$. We are in particular interested in the behaviour of $T_1(x)$ for starting points $x_e < x < 1$, where the deterministic flow points to $x = 1$ and exit at $x = 0$ is unlikely. As in section 3, the exact solutions for $u(x)$ and $T_1(x)$ follow from the expressions (3.4) and (3.6). For the present flow field $b(x)$ it is rather complicated to derive an asymptotic approximation of $T_1(x)$ by expanding the exact solution. Therefore, we proceed along the lines of subsection 3.2. We asymptotically solve (2.5) for $u(x)$ and, using this solution, determine a solution for $T_1(x)$ by asymptotically solving the boundary value problem (3.12) for $T_1(x)$.

4.1 The probability of exit

For the flow field considered in this section it is expected that $u(x)$ will change rapidly from about 1 to 0 in a small region near the unstable steady state x_e. So a boundary layer is expected to be present at this place. An examination of different stretchings of the coordinate shows the presence of a boundary layer of width $O(\varepsilon)$ around x_e. This boundary layer divides D in three subdomains: the boundary layer D_e, the region D_ε^- at the left of D_e, and the region D_ε^+ at the right of D_e. In particular for starting points $x \in D_\varepsilon^+$ the approximation for $u(x)$ should be sufficiently accurate to be used for the determination of $T_1(x)$. The expression for $u(x)$ obtained by the familiar method of matched asymptotic expansions is not accurate enough. Therefore, following Cook and Eckhaus (1973), we use WKB-expansions for $u(x)$ on subdomains.

We first consider the regions D_ε^- and D_ε^+. Substitution of the WKB-approximation $u(x) = w(x)\exp[-Q(x)/\varepsilon^2]$, see (3.13), into (2.5a) again yields the equations of the type (3.14) and (3.15). We note that these equations admit the solutions

$$Q(x) = J(x) \quad \text{and} \quad w(x) = a(x)/b(x), \qquad (4.2a)$$

$$Q(x) = \text{constant} \quad \text{and} \quad w(x) = \text{constant}, \qquad (4.2b)$$

with $J(x)$ defined by

$$J(x) = \int_{x_e}^{x} \frac{2b(s)}{a(s)}\, ds. \tag{4.3}$$

Taking linear combinations of WKB-approximations we find

$$u(x) = B_1 + B_2 \frac{a(x)}{b(x)} \exp[-J(x)/\varepsilon^2] \quad \text{in} \quad D_e^-, \tag{4.4a}$$

$$u(x) = B_3 + B_4 \frac{a(x)}{b(x)} \exp[-J(x)/\varepsilon^2] \quad \text{in} \quad D_e^+. \tag{4.4b}$$

To find the expansion of $u(x)$ valid in the boundary layer D_e, we introduce the appropriate stretched variable

$$\xi = (x - x_e)/\varepsilon. \tag{4.5}$$

Substitution into (2.5a) yields for $\varepsilon \to 0$

$$b'(x_e)\xi \frac{du}{d\xi} + \frac{1}{2} a(x_e) \frac{d^2u}{d\xi^2} = 0, \tag{4.6}$$

which has the solution

$$u(\xi) = B_5 + B_6 \int_0^{\xi} \exp[-c_e v^2]\, dv \quad \text{with} \quad c_e = b'(x_e)/a(x_e). \tag{4.7a,b}$$

We note that for the present flow field $c_e > 0$.

The constants B_i are determined by matching the solutions (4.4a) and (4.4b) with (4.7), and by applying the boundary conditions (2.5b). Using the asymptotic expansion of the complementary error function, see Abramowitz and Stegun (1965), we find that the inner solution $u(\xi)$ behaves like

$$B_5 \pm \frac{1}{2} B_6 \sqrt{\pi/c_e} - \frac{B_6}{2c_e\xi} \exp[-c_e\xi^2] \quad \text{for} \quad \xi \to \pm\infty. \tag{4.8}$$

Expressing the outer solution (4.4a) in the inner variable ξ and expanding for $\xi \to -\infty$, we see that the outer solution behaves like

$$B_1 + \frac{B_2}{c_e \varepsilon \xi} \exp[-c_e\xi^2] \quad \text{for} \quad \xi \to -\infty. \tag{4.9}$$

Matching (4.4a) and (4.7), using (4.8) and (4.9), yields

$$B_1 = B_5 - \frac{1}{2} B_6 \sqrt{\pi/c_e}\,, \qquad B_2/\varepsilon = -B_6/2. \tag{4.10a,b}$$

By matching (4.4b) and (4.7) we obtain in an analogous manner the relations

$$B_3 = B_5 + \tfrac{1}{2} B_6 \sqrt{\pi/c_e}\,, \qquad B_4/\varepsilon = -B_6/2, \tag{4.11a,b}$$

and by applying the boundary conditions (2.5b)

$$B_1 + B_2 \frac{a(0)}{b(0)} \exp[-J(0)/\varepsilon^2] = 1, \qquad B_3 + B_4 \frac{a(1)}{b(1)} \exp[-J(1)/\varepsilon^2] = 0. \tag{4.12a,b}$$

Solving Eqs. (4.10) - (4.12) yields the constants B_i. The expressions (4.4) and (4.7) thus obtained for $u(x)$ can be simplified by using appropriate approximations. The following asymptotic approximation is then found

$$u(x) = 1 + \tfrac{1}{2}\varepsilon \sqrt{c_e/\pi}\, \frac{a(x)}{b(x)} \exp[-J(x)/\varepsilon^2] \quad \text{in} \quad D_e^-, \tag{4.13a}$$

$$u(x) = \tfrac{1}{2} - \sqrt{c_e/\pi} \int_0^{(x - x_e)/\varepsilon} \exp[-c_e v^2]\, dv \quad \text{in} \quad D_e, \tag{4.13b}$$

$$u(x) = \tfrac{1}{2}\varepsilon \sqrt{c_e/\pi}\, \Big(\frac{a(x)}{b(x)} \exp[-J(x)/\varepsilon^2] - \frac{a(1)}{b(1)} \exp[-J(1)/\varepsilon^2] \Big) \quad \text{in} \quad D_e^+. \tag{4.13c}$$

It is seen from these approximations that $u(x)$ is exponentially small in D_e^+ and equals 1 with an exponentially small deviation in D_e^-.

4.2 The expected exit time

We now asymptotically solve the boundary value problem (3.12) for $T_1(x)$. Using Eqs. (4.13) we determine solutions of (3.12a) valid on subdomains. The unknown constants are obtained by matching the solutions and using the boundary conditions (3.12b).

The region D_e^-. On D_e^- the differential equation (3.12a) is approximated by

$$b(x)\frac{dT_1}{dx} + \tfrac{1}{2}\varepsilon^2 a(x) \frac{d^2T_1}{dx^2} = -1, \tag{4.14}$$

as follows from substitution of (4.13a). Taking $\varepsilon = 0$ yields the reduced equation. Solving this equation and using the boundary condition $T_1(0) = 0$, see (3.12b), we obtain the first term of a regular expansion

$$T_1(x) = - \int_0^x \frac{1}{b(s)}\, ds. \tag{4.15}$$

We note that this approximation equals the travel time with the deterministic flow, as is expected for starting points in this part of the domain. For matching purposes we need the behaviour of expression (4.15) near the boundary layer D_e. It is given by

$$T_1(x) \sim -\frac{1}{b'(x_e)}\ln(x_e - x) + T_1^- \tag{4.16a}$$

with

$$T_1^- = \frac{1}{b'(x_e)}\ln x_e - \int_0^{x_e} \frac{1}{b(s)} - \frac{1}{b'(x_e)(s - x_e)}\,ds. \tag{4.16b}$$

The region D_e^+. In D_e^+ we substitute (4.13c) into (3.12a). For x away from 1 we find the approximation

$$(-b(x) + \varepsilon^2\frac{a'(x)b(x) - a(x)b'(x)}{b(x)})\frac{dT_1}{dx} + \tfrac{1}{2}\varepsilon^2 a(x)\frac{d^2T_1}{dx^2} = -1, \tag{4.17}$$

equal to Eq. (3.17) for the flow field of section 3. Taking $\varepsilon = 0$ and solving the equation we obtain

$$T_1(x) = -\int_x^1 \frac{1}{b(s)}\,ds + P, \tag{4.18}$$

with the constant P to be determined by matching. For this purpose the behaviour of (4.18) near the boundary layer D_e is needed. It is given by

$$T_1(x) \sim \frac{1}{b'(x_e)}\ln(x - x_e) + T_1^+ + P \tag{4.19a}$$

with

$$T_1^+ = -\frac{1}{b'(x_e)}\ln(1 - x_e) - \int_{x_e}^1 \frac{1}{b(s)} - \frac{1}{b'(x_e)(s - x_e)}\,ds. \tag{4.19b}$$

For the neighbourhood of $x = 1$ it can be verified that the differential equation in the appropriate local variable is satisfied by T_1 = constant. Taking this constant equal to P for matching purposes, we can conclude that near $x = 1$ the first order approximation for $T_1(x)$ is contained in expression (4.18).

The boundary layer region D_e. We now consider the region D_e. Substitution into (3.12a) of the approximation (4.13b) for $u(x)$ and the boundary layer variable ξ given by (4.5), yields for $\varepsilon \to 0$

$$(b'(x_e)\xi - a(x_e)\exp[-c_e\xi^2] \,/ \int_\xi^\infty \exp[-c_e v^2]\,dv)\frac{dT_1}{d\xi} + \tfrac{1}{2}a(x_e)\frac{d^2T_1}{d\xi^2} = -1. \tag{4.20}$$

This first order differential equation in $dT_1/d\xi$ can be solved. Its solution is written as

$$T_1(\xi) = \frac{2}{a(x_e)} \int_\xi^\infty \exp[c_e t^2] (\int_t^\infty \exp[-c_e v^2]\, dv)^2\, dt \;/ \int_\xi^\infty \exp[-c_e v^2]\, dv \; +$$

$$+ \frac{2}{a(x_e)} \int_0^\xi \exp[c_e t^2] \int_t^\infty \exp[-c_e v^2]\, dv\, dt + k_1 / \int_\xi^\infty \exp[-c_e v^2]\, dv + k_2, \quad (4.21)$$

where the constants k_1 and k_2 follow from matching with the outer solutions (4.16) and (4.19). The form (4.21) is convenient for matching with (4.19). We investigate the behaviour of (4.21) for $\xi \to \infty$. It can be shown that the first term at the right side of (4.21) tends to 0 for $\xi \to \infty$. For the second term at the right side of (4.21) it can be shown that it behaves like

$$\frac{1}{b'(x_e)} \ln \xi + \frac{1}{b'(x_e)} (\beta + \tfrac{1}{2} \ln c_e) \quad \text{for} \quad \xi \to \infty \quad (4.22a)$$

with

$$\beta = \lim_{\zeta \to \infty} 2 \int_0^\zeta \exp[\tau^2] \int_\tau^\infty \exp[-\phi^2]\, d\phi\, d\tau - \ln \zeta. \quad (4.22b)$$

The value of β can be determined numerically, $\beta = .982$. The third term at the right side of (4.21) behaves like

$$2k_1 c_e \xi \exp[c_e \xi^2] \quad \text{for} \quad \xi \to \infty. \quad (4.23)$$

Expressing the outer solution (4.19a) in the inner variable $\xi = (x - x_e)/\varepsilon$ and expanding for $\xi \to \infty$, it is seen that it behaves like

$$\frac{1}{b'(x_e)} \ln \xi + \frac{1}{b'(x_e)} \ln \varepsilon + T_1^+ + P \quad \text{for} \quad \xi \to \infty. \quad (4.24)$$

From (4.23) we see that the third term at the right side of (4.21), which grows exponentially for $\xi \to \infty$, can not be matched for $k_1 \neq 0$. Comparing (4.22a) and (4.24) it is seen that the second term in (4.21) shows the needed asymptotic behaviour for $\xi \to \infty$. We are led to the matching conditions

$$k_1 = 0, \quad (4.25a)$$

$$\frac{1}{b'(x_e)} (\beta + \tfrac{1}{2} \ln c_e) + k_2 = \frac{1}{b'(x_e)} \ln \varepsilon + T_1^+ + P. \quad (4.25b)$$

In order to match (4.21) with the outer solution (4.16) on D_e^-, we rewrite (4.21). After some manipulations and using (4.25a), we can write

$$T_1(\xi) = \frac{2}{a(x_e)} \int_{\xi}^{\infty} \exp[-c_e s^2] \int_{0}^{s} \exp[c_e t^2] \int_{t}^{\infty} \exp[-c_e v^2]\, dv\, dt\, ds \Big/ \int_{\xi}^{\infty} \exp[-c_e v^2]\, dv + k_2. \tag{4.26}$$

Now one can show that $T_1(\xi)$ behaves like

$$-\frac{1}{b'(x_e)}\ln(-\xi) - \frac{1}{b'(x_e)}(\beta - 4\gamma/\sqrt{\pi} + \tfrac{1}{2}\ln c_e) + k_2 \quad \text{for} \quad \xi \to -\infty \tag{4.27a}$$

with

$$\gamma = \int_0^{\infty} \exp[-\sigma^2] \int_0^{\sigma} \exp[\tau^2] \int_{\tau}^{\infty} \exp[-\phi^2]\, d\phi\, d\tau\, d\sigma. \tag{4.27b}$$

The value of γ can be determined numerically, $\gamma = .307$.

Substituting the inner variable ξ into the outer solution (4.16a) and expanding it for $\xi \to -\infty$, we find that the outer solution behaves like

$$-\frac{1}{b'(x_e)}\ln(-\xi) - \frac{1}{b'(x_e)}\ln\varepsilon + T_1^- \quad \text{for} \quad \xi \to -\infty. \tag{4.28}$$

Now matching (4.27a) and (4.28) yields

$$k_2 = -\frac{1}{b'(x_e)}\ln\varepsilon + \frac{1}{b'(x_e)}(\beta - 4\gamma/\sqrt{\pi} + \tfrac{1}{2}\ln c_e) + T_1^-. \tag{4.29}$$

We note that, using (4.25b) and (4.29), we can write the constant P in (4.18) as

$$P = -\frac{2}{b'(x_e)}\ln\varepsilon + \frac{2}{b'(x_e)}(\beta - 2\gamma/\sqrt{\pi} + \tfrac{1}{2}\ln c_e) + T_1^- - T_1^+. \tag{4.30}$$

We can summarize the results for $T_1(x)$. In D_e^- we have the approximation (4.15), that is the deterministic travel time. In D_e we have the boundary layer solution $T_1(\xi)$ given by (4.26), with ξ given by (4.5) and k_2 by (4.29). And in the region D_e^+, that we are in particular interested in, $T_1(x)$ is approximated by (4.18), with P given by (4.30). We note that it is possible to construct a composite expansion T_{comp} to avoid the difficulty of deciding whether a starting point is within or outside the boundary layer, see Van Dyke (1975):

$$T_{comp} = T_{outer} + T_{bound} - T_{match}. \tag{4.31}$$

Here T_{bound} is given by (4.26). For $x < x_e$ the function T_{outer} is given by (4.15) and T_{match} by (4.28), and for $x > x_e$ we have T_{outer} given by (4.18) and T_{match} by (4.24).

5. Flow with a stable steady state

In this section we study as third fundamental example a stochastic dynamical system with one stable deterministic equilibrium. The stochastic system is given on the interval D defined by (2.1). We assume that the deterministic flow $b(x)$ has a stable equilibrium x_e in D:

$$b(x_e) = 0, \quad b(x) > 0 \quad \text{for} \quad 0 \le x < x_e, \quad b(x) < 0 \quad \text{for} \quad x_e < x \le 1, \tag{5.1a}$$

and that the flow and the diffusion coefficient satisfy

$$\int_0^1 \frac{2b(x)}{a(x)}\,dx > 0. \tag{5.1b}$$

The corresponding deterministic system will always approach the equilibrium x_e without reaching one of the boundary points $x = 0$ or $x = 1$. Because of the diffusion, the stochastic system can leave the domain of attraction of the stable steady state and reach the boundary. With probability one this will happen in a finite time. Condition (5.1b) implies that for an initial state away from $x = 0$ the boundary $x = 1$ is most likely reached before $x = 0$. So for starting points x away from $x = 0$ the solutions $u(x)$ and $T_1(x)$ of Eqs. (2.5) - (2.7) are the probability of exit and the (conditional) expected exit time, respectively, at the unlikely point of exit $x = 0$. We are interested in an asymptotic approximation of $T_1(x)$ for small values of ε. We note that the exact solution for $T_1(x)$ is known. It is given by $T_1(x) = T(x)/u(x)$ where $u(x)$ and $T(x)$ can be written as (3.4) and (3.6), respectively, but it gives little insight in the behaviour of $T_1(x)$. To find the asymptotic expansion we proceed along the lines of section 4. We asymptotically solve (2.5) for $u(x)$ and, using the solution, asymptotically solve the boundary value problem (3.12) for $T_1(x)$.

5.1 The probability of exit

We consider the boundary value problem (2.5) for $u(x)$. Following Cook and Eckhaus (1973), we again use WKB-approximations on subdomains. The same subdomains as in section 4 are distinguished. Around the equilibrium x_e, where the coefficient of du/dx in (2.5a) vanishes, we have the region D_e of width $O(\varepsilon)$. At the left and at the right of D_e we have the regions D_e^- and D_e^+, respectively.

First we find an approximation for $u(x)$ in the regions D_e^- and D_e^+ by substituting the WKB-approximation $u(x) = w(x)\exp[-Q(x)/\varepsilon^2]$ into (2.5a). In the same way as in subsection 4.1 we obtain approximations for $u(x)$ that can be written as

$$u(x) = A_1 + A_2 \frac{a(x)}{b(x)} \exp[-K(x)/\varepsilon^2] \quad \text{in} \quad D_e^-, \tag{5.2a}$$

$$u(x) = A_3 + A_4 \frac{a(x)}{b(x)} \exp[-K(x)/\varepsilon^2] \quad \text{in} \quad D_e^+, \tag{5.2b}$$

with the integral $K(x)$ defined by

$$K(x) = -\int_x^1 \frac{2b(s)}{a(s)}\, ds. \tag{5.3}$$

In the region D_e we substitute the appropriate stretched variable $\xi = (x - x_e)/\varepsilon$ and let $\varepsilon \to 0$. The resulting equation is of the form (4.6) and has the solution

$$u(\xi) = A_5 + A_6 \int_0^\xi \exp[-c_e v^2]\, dv \qquad \text{with} \qquad c_e = b'(x_e)/a(x_e). \tag{5.4a,b}$$

For the present flow field we have $c_e < 0$. For matching purposes we need the behaviour of (5.4) for $\xi \to \pm\infty$. From the asymptotic expansion of Dawson's integral, see Spanier and Oldham (1987), it is seen that $u(\xi)$ behaves like

$$A_5 - \frac{A_6}{2c_e\xi} \exp[-c_e\xi^2] \quad \text{for} \quad \xi \to \pm\infty. \tag{5.5}$$

In the same way as in subsection 4.1 we can now determine the constants A_i by matching (5.2a) and (5.2b) with (5.4), and by using the boundary conditions (2.5b). We find the following asymptotic approximation

$$u(x) = -\frac{b(0)}{a(0)} \exp[K(0)/\varepsilon^2] \left(\frac{a(1)}{b(1)} - \frac{a(x)}{b(x)} \exp[-K(x)/\varepsilon^2]\right) \ \text{in} \ D_e^- \ \text{and} \ D_e^+, \tag{5.6a}$$

$$u(x) = -\frac{b(0)}{a(0)} \exp[K(0)/\varepsilon^2] \left(\frac{a(1)}{b(1)} + \frac{2}{\varepsilon} \exp[-K(x_e)/\varepsilon^2] \int_0^{(x-x_e)/\varepsilon} \exp[-c_e v^2]\, dv\right) \ \text{in} \ D_e. \tag{5.6b}$$

We note that the approximation for $u(x)$ that can be found by solving (2.5) with the familiar method of matched asymptotic expansions and by using the divergence theorem (see Schuss (1980)), is not sufficiently accurate to determine T_1.

5.2 The expected exit time

Using the solution for $u(x)$, we asymptotically solve the boundary value problem (3.12) for $T_1(x)$. We first consider the coefficient $b(x) + \varepsilon^2 a(x)u'(x)/u(x)$ of $dT_1(x)/dx$ in (3.12a). Because of the term $\exp[-K(x)/\varepsilon^2]$ in the expression (5.6a) for $u(x)$, the coefficient vanishes in an order $O(\varepsilon^2)$ neighbourhood of the point x_0 with $0 < x_0 < x_e$ characterized by

$$\int_{x_0}^{1} \frac{2b(x)}{a(x)}\,dx = 0. \tag{5.7}$$

This can be seen from substitution of (5.6a) into the coefficient and considering that $\exp[-K(x)/\varepsilon^2]$ changes from exponentially large to exponentially small in an order $O(\varepsilon^2)$ neighbourhood of x_0. This region around x_0 divides the domain D in three subdomains: the order $O(\varepsilon^2)$ neighbourhood D_0 of x_0, the region D_0^- at the left of D_0, and the region D_0^+ at the right of D_0, see Figure 1. We will determine solutions of (3.12a) valid on these subdomains. It is assumed that $T_1(x)$ is of the form

$$T_1(x) = C(\varepsilon)\tau(x) \tag{5.8}$$

with $C(\varepsilon)$ exponentially large.

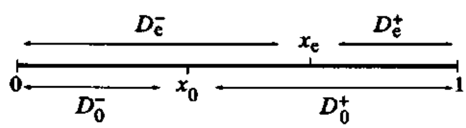


Figure 1. *Subdomains for the solutions of $u(x)$ and $T_1(x)$ for a flow with a stable steady state x_e.*

5.2.1 Solutions on subdomains

We first consider the region D_0^+. Substitution of (5.6a), (5.6b) and (5.8) in (3.12a) yields for x away from 1 the approximation for small values of ε

$$\left(b(x) + 2a(x)\frac{b(1)}{a(1)}\exp[-K(x)/\varepsilon^2]\right)\frac{d\tau}{dx} + \tfrac{1}{2}\varepsilon^2 a(x)\frac{d^2\tau}{dx^2} = -\frac{1}{C(\varepsilon)}. \tag{5.9}$$

Taking the limit $\varepsilon \to 0$ yields the reduced equation

$$b(x)\frac{d\tau}{dx} = 0. \tag{5.10}$$

This equation is satisfied by a constant that we can take 1 (any other value can be taken up in $C(\varepsilon)$):

$$\tau(x) = 1. \tag{5.11}$$

It can be verified that this solution is also valid near $x = 1$, by asymptotically solving the differential equation (3.12a) in the appropriate local variable.

In the region D_0^- we obtain by substitution of (5.6a) and (5.8) in (3.12a) the approximating differential equation

$$\left(-b(x) + \varepsilon^2 \frac{a'(x)b(x) - a(x)b'(x)}{b(x)}\right)\frac{d\tau}{dx} + \tfrac{1}{2}\varepsilon^2 a(x)\frac{d^2\tau}{dx^2} = -\frac{1}{C(\varepsilon)}. \tag{5.12}$$

Taking the limit $\varepsilon \to 0$ we find

$$-b(x)\frac{d\tau}{dx} = 0, \tag{5.13}$$

with the solution satisfying the boundary condition $T_1(0) = 0$, see (3.12b),

$$\tau(x) = 0. \tag{5.14}$$

We note that in subsection 5.2.3 we will refine this solution.

We now consider the region D_0. Substitution of (5.6a) and (5.8) into (3.12a) yields, after introduction of the local variable

$$\rho = (x - x_0)/\varepsilon^2 \tag{5.15}$$

and taking the limit $\varepsilon \to 0$, the differential equation

$$c_0\left(1 - 2\left(1 - c_0\frac{a(1)}{b(1)}\exp[2c_0\rho]\right)^{-1}\right)\frac{d\tau}{d\rho} + \tfrac{1}{2}\frac{d^2\tau}{d\rho^2} = 0 \tag{5.16a}$$

with the constant

$$c_0 = b(x_0)/a(x_0) > 0. \tag{5.16b}$$

The solution of this equation is

$$\tau(\rho) = d_1 + d_2\left(1 - c_0\frac{a(1)}{b(1)}\exp[2c_0\rho]\right)^{-1}. \tag{5.17}$$

Matching (5.17) with the solutions (5.11) on D_0^+ and (5.14) on D_0^- yields the constants

$$d_1 = 1, \quad d_2 = -1, \tag{5.18a,b}$$

respectively. The constant $C(\varepsilon)$ is yet unknown. Subsection 5.2.2 concerns the determination of this constant.

5.2.2 The constant $C(\varepsilon)$

To determine the constant $C(\varepsilon)$, we first solve the forward equation

$$Mp = 0 \tag{5.19}$$

with the operator M defined by (2.2b). A solution $p(x)$ of this equation will be used for the determination of $C(\varepsilon)$. The function $p(x)$ describes, if appropriately scaled, the quasi-stationary distribution. With this quasi-stationary probability distribution we mean the distribution given the system has not reached the boundary of the domain D.

We seek a solution of (5.19) in the form of a WKB-approximation

$$p(x) = r(x)\exp[-S(x)/\varepsilon^2] \tag{5.20a}$$

with the conditions in the equilibrium point x_e

$$S(x_e) = 0, \quad r(x_e) = 1. \tag{5.20b,c}$$

Since x_e is the most likely place for the state to be found, $S(x)$ should have a minimum in x_e. Substituting the WKB-approximation into (5.19) and rearranging terms, we obtain to order $O(\varepsilon^0)$ two equations, that can be solved for the functions $S(x)$ and $r(x)$. We obtain

$$p(x) = \frac{a(x_e)}{a(x)}\exp[(K(x) - K(x_e))/\varepsilon^2]. \tag{5.21}$$

This function $p(x)$ will be used in the following relation that can easily be derived

$$\int_0^1 (pLT - TMp)\,dx = \left[bpT + \tfrac{1}{2}\varepsilon^2(apT' - ap'T - a'pT)\right]_0^1 \tag{5.22}$$

(divergence theorem). By (2.7a) and (5.19) the left side reduces to an integral that, after substitution of (5.6) and (5.21), can be evaluated with the method of Laplace. The result is

$$\frac{b(0)}{a(0)}\frac{a(1)}{b(1)}\sqrt{-\pi/c_e}\;\varepsilon\exp[K(0)/\varepsilon^2]. \tag{5.23}$$

The right side of (5.22) can be simplified and expressed in $T_1(x)$ with (2.7b), (2.6) and (2.5b). Then using the results for $u(x)$, $T_1(x)$ and $p(x)$ derived in this section, we obtain the approximation for the right side

$$-a(x_e)\frac{b(0)}{a(0)}C(\varepsilon)\exp[(K(0) - K(x_e))/\varepsilon^2]. \tag{5.24}$$

Combining (5.23) and (5.24) we find

$$C(\varepsilon) = -\frac{a(1)}{b(1)}\sqrt{-\pi/(b'(x_e)a(x_e))}\;\varepsilon\exp[K(x_e)/\varepsilon^2]. \tag{5.25}$$

Summarizing we have thus far obtained the following asymptotic results for $T_1(x)$. In the region D_0^- we have the approximation $T_1(x) = 0$, see (5.14). In the region D_0^+ we have obtained $T_1(x) = C(\varepsilon)$ with $C(\varepsilon)$ given by (5.25), while in the boundary layer D_0 we have the approximation $T_1(\rho)$ given by (5.8), (5.17), (5.18).

We remark that the exact value for $T_1(1)$, which is determined by the relation $T_1(x) = T(x)/u(x)$ with $u(x)$ and $T(x)$ written as (3.4) and (3.6), respectively, can be evaluated for small values of ε using Laplace's method. The result thus obtained is equal to the value (5.25) for $C(\varepsilon)$. In the following subsection the approximation for $T_1(x)$ in D_0^- will be refined.

5.2.3 A refinement of the solution in D_0^-

We are not content with the very rude approximation $T_1(x) = 0$ we have obtained thus far for the expected exit time at $x = 0$ for starting points in the region D_0^-. This solution will now be improved. The solution (5.17), (5.18a) for $\tau(\rho)$ in the region D_0 behaves like

$$1 + d_2(1 + c_0 \frac{a(1)}{b(1)} \exp[2c_0\rho]) \quad \text{for} \quad \rho \to -\infty. \tag{5.26}$$

This suggests an approximation for $\tau(x)$ in the region D_0^- in the form of a WKB-expansion

$$\tau(x) = v(x) \exp[-P(x)/\varepsilon^2]. \tag{5.27}$$

Substitution in the differential equation (5.12) for $\tau(x)$ and comparison with the value (5.25) of $C(\varepsilon)$ show that, as long as

$$P(x) < K(x_e), \tag{5.28}$$

the differential equation can be seen as a homogeneous equation for small values of ε. Under the assumption (5.28) we obtain in the way of subsection 4.1 the linear combination of WKB-approximations

$$\tau(x) = d_3 + d_4 \frac{b(x)}{a(x)} \exp[K(x)/\varepsilon^2]. \tag{5.29}$$

To match this solution with (5.26) we express (5.29) in the inner variable $\rho = (x - x_0)/\varepsilon^2$ and expand it for $\rho \to -\infty$. We find that this outer solution behaves like

$$d_3 + d_4 c_0 \exp[2c_0\rho] \quad \text{for} \quad \rho \to -\infty. \tag{5.30}$$

So matching (5.30) and (5.26) yields two relations from which d_3 and d_4 can be expressed in d_2. We obtain the WKB-solution

$$\tau(x) = 1 + d_2 + d_2 \frac{a(1)}{b(1)} \frac{b(x)}{a(x)} \exp[K(x)/\varepsilon^2]. \tag{5.31}$$

Now we can distinguish two cases.

Case 1. If

$$\int_0^{x_0} \frac{2b(x)}{a(x)}\,dx < \int_{x_0}^{x_\varepsilon} \frac{2b(x)}{a(x)}\,dx, \tag{5.32}$$

then $-K(0) < K(x_\varepsilon)$ and assumption (5.28) is satisfied for $x \in D_0^-$. In that case (5.31) is valid for $x \in D_0^-$ and the boundary condition $T_1(0) = 0$, see (3.12b), determines d_2. Carrying out appropriate approximations we find that the solution for $\tau(\rho)$ in D_0 remains unchanged and is given by (5.17), (5.18), and that in D_0^- we have the improved approximation

$$\tau(x) = \frac{a(1)}{b(1)}\Big(\frac{b(0)}{a(0)}\exp[K(0)/\varepsilon^2] - \frac{b(x)}{a(x)}\exp[K(x)/\varepsilon^2]\Big). \tag{5.33}$$

Case 2. If

$$\int_0^{x_0} \frac{2b(x)}{a(x)}\,dx \geq \int_{x_0}^{x_\varepsilon} \frac{2b(x)}{a(x)}\,dx, \tag{5.34}$$

then there is a point $0 \leq x_1 < x_0$ characterized by

$$\int_{x_1}^{x_0} \frac{2b(x)}{a(x)}\,dx = \int_{x_0}^{x_\varepsilon} \frac{2b(x)}{a(x)}\,dx. \tag{5.35}$$

In that case assumption (5.28) is satisfied for $x_1 < x < x_0$ and solution (5.31) is still valid for $x \in D_0^-$ with $x > x_1$. For $x \leq x_1$ we asymptotically solve the inhomogeneous differential equation, that in D_0^- is given by

$$\Big(-b(x) + \varepsilon^2\frac{a'(x)b(x) - a(x)b'(x)}{b(x)}\Big)\frac{dT_1}{dx} + \frac{1}{2}\varepsilon^2 a(x)\frac{d^2T_1}{dx^2} = -1, \tag{5.36}$$

see (5.12). Taking $\varepsilon = 0$ and using the boundary condition $T_1(0) = 0$, see (3.12b), we find for $x \leq x_1$ the first order approximation

$$T_1(x) = \int_0^x \frac{1}{b(s)}\,ds, \tag{5.37}$$

in correspondence with expression (3.20) for the flow field without steady states. This solution can be matched with (5.31) in x_1. We obtain

$$\int_0^{x_1} \frac{1}{b(s)}\,ds = C(\varepsilon)(1 + d_2). \tag{5.38}$$

The asymptotic solution in D_0 remains unchanged and is given by (5.17) and (5.18), whereas the solution in D_0^- for $x > x_1$ is approximated by

$$T_1(x) = \int_0^{x_1} \frac{1}{b(s)}\, ds - C(\varepsilon)\frac{a(1)}{b(1)}\frac{b(x)}{a(x)} \exp[K(x)/\varepsilon^2]. \tag{5.39}$$

We can summarize the asymptotic results we have obtained for $T_1(x)$ in this section. In the region D_0^+ we have found $T_1(x) = C(\varepsilon)$ with $C(\varepsilon)$ given by (5.25). In the boundary layer D_0 we have the approximation $T_1(\rho)$ given by (5.8), (5.17), (5.18). In the region D_0^- we have in case 1 the approximation (5.8), (5.33), and in case 2 we have in D_0^- the approximation (5.39) for $x > x_1$ and (5.37) for $x \leq x_1$.

5.2.4 A comparison of exit times at different boundaries

Above we have obtained results for the conditional expected exit time $T_1(x)$ at the unlikely exit point $x = 0$ for starting points $x \in D$. It is interesting to compare these results with the conditional expected exit time $T_0(x)$ at the other boundary point $x = 1$. The function $T_0(x)$ is determined by

$$T_0(x) = T(x)/(1 - u(x)) \tag{5.40}$$

with $T(x)$ satisfying the boundary value problem

$$LT = -(1 - u) \quad \text{in} \quad D \tag{5.41a}$$

$$T(0) = 0, \quad T(1) = 0. \tag{5.41b}$$

Here $u(x)$ denotes the exit probability at $x = 0$, determined by the boundary value problem (2.5). The solution for $u(x)$ is given by (5.6). In the same way as we obtained a solution for $T_1(x)$ we can determine an asymptotic solution for $T_0(x)$. A boundary layer of width $O(\varepsilon^2)$ is found near $x = 1$. Outside this boundary layer we obtain

$$T_0(x) = C(\varepsilon), \tag{5.42a}$$

inside this boundary layer

$$T_0(x) = C(\varepsilon)(1 - \exp[-2b(1)(x - 1)/(a(1)\varepsilon^2)]). \tag{5.42b}$$

Using the divergence theorem we obtain the same value (5.25) for $C(\varepsilon)$ as we found for the expected exit time $T_1(x)$ at the unlikely exit point $x = 0$. So for starting points $x \in D_0^+$, away from $x = 1$, we have the interesting result that the (conditional) expected exit times at the unlikely exit point $x = 0$ and the most probable exit point $x = 1$ are equal in first order approximation. An interpretation is as follows. For a starting point $x \in D_0^+$, away from $x = 1$, we expect the system to approach first the stable equilibrium and to remain in the neighbourhood of it for a long time. Large excursions from the equilibrium take place with small probabilities. Exit at the boundary occurs during such an excursion. Now it is expected that the exit time at

$x = 0$ equals the exit time at $x = 1$ in first order approximation, because if, say, the exit time at $x = 0$ were considerably larger exit would already have occurred at $x = 1$.

We note that for the determination of $T_0(x)$ it is not necessary to use a WKB-approximation for $u(x)$ as obtained in subsection 5.1, but that an approximation of $u(x)$ found by solving (2.5) with the familiar method of matched asymptotic expansions and using the divergence theorem is sufficiently accurate.

The conditional expected exit times can also be compared with the expected exit time $T(x)$ at the boundary ∂D regardless of the point where exit takes place. It immediately follows from the information in section 2 that for a starting point $x \in D$ the expected exit time $T(x)$ from D where we do not distinguish between exit at $x = 0$ or at $x = 1$, satisfies

$$LT = -1 \quad \text{in} \quad D, \tag{5.43a}$$

$$T(0) = 0, \quad T(1) = 0. \tag{5.43b}$$

Solving (5.43) we find boundary layers of width order $O(\varepsilon^2)$ near $x = 0$ and $x = 1$. In the boundary layer near $x = 0$ we obtain

$$T(x) = C(\varepsilon)(1 - \exp[-2b(0)x/(a(0)\varepsilon^2)]), \tag{5.44a}$$

outside the boundary layers

$$T(x) = C(\varepsilon), \tag{5.44b}$$

and inside the boundary layer near $x = 1$

$$T(x) = C(\varepsilon)(1 - \exp[-2b(1)(x - 1)/(a(1)\varepsilon^2)]). \tag{5.44c}$$

The divergence theorem again yields the value (5.25) for $C(\varepsilon)$. We can use this result to examine the asymptotic approximations we have derived in this section by verifying the relation

$$T_1 u + T_0(1 - u) = T. \tag{5.45}$$

By substitution in the left side of (5.45) of the asymptotic solutions derived for $u(x)$, $T_1(x)$ and $T_0(x)$ the approximation (5.44) for $T(x)$ is obtained, indeed, with an exponentially small relative error.

6. A comparison with numerical results

In this section the analytical approximations derived in the foregoing sections are compared with numerical results. First we present results of stochastic simulations for the flow without steady states. The stochastic differential equation (2.3a) can be approximated by the stochastic difference equation

$$\Delta x = b(x)\Delta t + \varepsilon \sqrt{a(x)}\, \xi(t) \sqrt{\Delta t}\ , \tag{6.1}$$

where $\xi(t)$ is normally distributed with zero mean and unit variance, see Gardiner (1983). Using (6.1) the stochastic process can be simulated. In Tables 1 and 2 we give simulation results for the one-dimensional flow of section 3. By carrying out simulation runs for starting points $0 < x < 1$, we have found approximations u_{sim} and T_{1sim} for the probability that exit occurs at $x = 0$ and for the conditional expected exit time at this boundary point, respectively. These simulation results are compared with the asymptotic approximations u_{as} and T_{1as}, computed from (3.5) and (3.9), or (3.20), respectively. In Table 1 we are interested in the x-dependence of the asymptotic approximations. For each starting point 25000 runs have been made; $\varepsilon^2 = .1$, $b(x) = .5$ and $a(x) = 1$. In Table 2 we attempt to investigate the independence of $a(x)$ of the first order approximation (3.9) for the expected exit time $T_1(x)$. For each (constant) function $a(x)$ 40000 runs have been made, with starting point $x = .13$; $\varepsilon^2 = .05$ and $b(x) = .5 + .5x$. It is seen that the asymptotic approximations are in good accordance with the simulation results.

Table 1. *The probability u of exit and the expected exit time T_1 at the boundary point $x = 0$ for a flow without steady states. The values of u_{sim} and T_{1sim} have been obtained from $N = 25000$ runs at each starting point x with $0 < x < 1$;* $\varepsilon^2 = .1$, $b(x) = .5$ and $a(x) = 1$. *The asymptotic approximations u_{as} and T_{1as} have been computed from* (3.5) *and* (3.9).

x	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50	.55
u_{sim}	.57	.35	.21	.13	.077	.047	.028	.017	.009	.007	.003
u_{as}	.61	.37	.22	.14	.082	.050	.030	.018	.011	.007	.004
T_{1sim}	.11	.21	.30	.41	.50	.62	.71	.85	1.0	1.1	1.3
T_{1as}	.10	.20	.30	.40	.50	.60	.70	.80	.9	1.0	1.1

Table 2. *The probability u of exit and the expected exit time T_1 at the boundary point $x = 0$ for a flow without steady states. For each (constant) function a the values of u_{sim} and T_{1sim} have been obtained from $N = 40000$ runs with starting point $x = .13$; $\varepsilon^2 = .05$ and $b(x) = .5 + .5x$. The asymptotic approximations u_{as} and T_{1as} have been computed from* (3.5) *and* (3.9).

a	.67	1.0	1.5
u_{sim}	.013	.052	.13
u_{as}	.014	.056	.14
T_{1sim}	.23	.23	.22
T_{1as}	.24	.24	.24

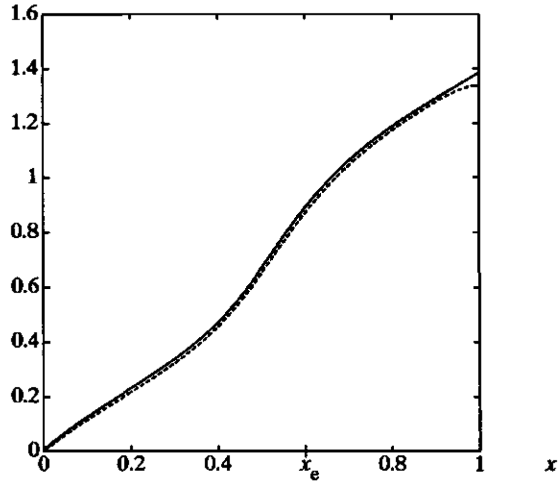


Figure 2. *The expected exit time at the boundary point $x = 0$ for a flow with an unstable steady state x_e. The composite expansion T_{comp} (—), calculated from* (4.31), *is compared with the exact solution T_1 (– –), given by* (2.6), (3.4) *and* (3.6); $\varepsilon^2 = .04$, $b(x) = -\sin((5x + 1)\pi/4)$ *and* $a(x) = 1$.

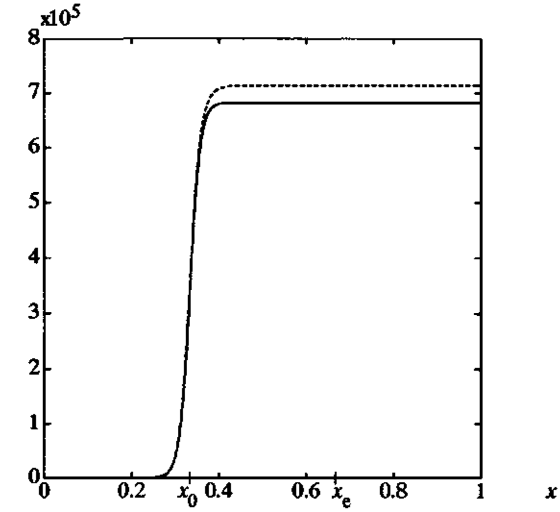


Figure 3. *The expected exit time at the boundary point $x = 0$ for a flow with a stable steady state x_e. The asymptotic approximation for T_1 (—) obtained in subsection 5.2, is compared with the exact solution T_1 (– –), given by* (2.6), (3.4) *and* (3.6); $\varepsilon^2 = .006$, $b(x) = .5 - .75x$ *and* $a(x) = 1$.

The asymptotic approximations we have derived for the conditional expected exit time at $x = 0$ for the flow problems in sections 4 and 5 are compared with the exact solutions. The exact solutions for $T_1(x)$ are given by $T_1(x) = T(x)/u(x)$, see (2.6), where $u(x)$ and $T(x)$ are of the type (3.4) and (3.6), respectively. In Figure 2 we have plotted the composite expansion T_{comp}, given by (4.31), and the exact solution for $T_1(x)$ for the flow field $b(x) = -\sin((5x + 1)\pi/4)$, which has the unstable steady state $x_e = .6$; $\varepsilon^2 = .04$ and $a(x) = 1$. The deviation near $x = 1$ is caused by higher order terms. In Figure 3 we compare the asymptotic approximation obtained for $T_1(x)$ in subsection 5.2 with the exact solution for the flow $b(x) = .5 - .75x$; $\varepsilon^2 = .006$ and $a(x) = 1$. This flow field has the stable equilibrium $x_e = .67$; $x_0 = .33$ and $x_1 = .20$. The difference between the functions for $x > x_0$ diminishes for $\varepsilon \to 0$, see the remark in subsection 5.2.2.

7. Conclusions

In this paper we have studied the (conditional) expected exit time from a domain in a one-dimensional stochastic dynamical system at the unlikely exit boundary. For three fundamental cases we have determined asymptotic solutions for the expected exit time for small values of the diffusion parameter by solving boundary value problems based on the Fokker-Planck equation for the stochastic system. The asymptotic solutions reveal the essential features of the expected exit time, whereas the exact solutions give little insight in the behaviour. Moreover, no numerical problems are involved in calculating the asymptotic solutions. For the flow without steady states we have found the interesting result that the expected exit time for diffusion against the flow equals in first order approximation the travel time with the deterministic flow in opposite direction. This type of solution is met again on a subdomain in the example of a flow with a stable steady state. In that example the solution for the expected exit time at the unlikely exit point exhibits a boundary layer at the rather unexpected place x_0 determined by the relation (5.7). An other interesting feature of this example is that for starting points in the subdomain containing the stable equilibrium the conditional expected exit times at the unlikely exit point and at the most probable exit point are equal in first order approximation.

We have determined the approximations by asymptotically solving the differential equations. For the determination of the expected exit time at the unlikely exit boundary an accurate approximation is needed for the probability of exit, which is very small. The usual method of matched asymptotic expansions with regular expansions on subdomains does not provide us with such an approximation. Therefore, we have made use of WKB-expansions on subdomains. In the example of flow with a stable steady state the divergence theorem has been used to determine the exponentially large value of the expected exit time for starting points in the subdomain containing the stable equilibrium. The method may be extended to higher dimensional problems. Finally, we note that the analytical results we have derived are in good correspondence with simulation results for the example without steady states and with numerical values of the exact solutions for the examples with a stable and an unstable equilibrium.

Acknowledgements

I wish to thank Johan Grasman for the discussions concerning the subject treated in this paper and for his remarks on the text. I also thank Albert Otten for discussions on the expected exit time in the example of flow without steady states, and Maarten de Gee for numerical advice.

References

Abramowitz, M. and I. Stegun (1965), *Handbook of Mathematical Functions*, Dover, New York.

Cook, L.P. and W. Eckhaus (1973), *Resonance in a boundary value problem of singular perturbation type*, Stud. in Appl. Math., 52, pp.129-139.

Cox, D.R. and H.D. Miller (1965), *The Theory of Stochastic Processes*, Redwood Press, Trowbridge, Wiltshire.

Eckhaus, W. (1973), *Matched Asymptotic Expansions and Singular Perturbations*, North Holland, Amsterdam.

Erdélyi, A. (1956), *Asymptotic Expansions*, Dover, New York.

Gardiner, C.W. (1983), *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, Springer-Verlag, Berlin.

Mangel, M. (1979), *Small fluctuations in systems with multiple steady states*, SIAM J. Appl. Math., 36, pp. 544-572.

Naeh, T., M.M. Klosek, B.J. Matkowsky and Z. Schuss (1990), *A direct approach to the exit problem*, SIAM J. Appl. Math., 50, pp. 595-627.

Schuss, Z. (1980), *Theory and Applications of Stochastic Differential Equations*, Wiley, New York.

Spanier, J. and K.B. Oldham (1987), *An Atlas of Functions*, Hemisphere Publishing Corporation, Washington.

Van Dyke, M. (1975), *Perturbation Methods in Fluid Mechanics*, The Parabolic Press, Stanford.

Van Herwaarden, O.A. (1994), *Spread of pollution by dispersive groundwater flow*, SIAM J. Appl. Math., 54, pp. 26-41.

Samenvatting

Analyse van onverwachte uittredingen met de Fokker-Planck-vergelijking

Vele verschijnselen in bijvoorbeeld de biologie of de natuurkunde kunnen worden gemodelleerd als een dynamisch systeem. Dikwijls vormen toevalsfluctuaties een essentieel deel van deze verschijnselen. Het is dan wenselijk stochastische dynamische systemen te beschouwen. Stochastische systemen laten gedrag toe dat is uitgesloten in het corresponderende deterministische systeem. Zo kan de rand (of een gedeelte van de rand) van een domein in de toestandsruimte bereikbaar zijn in het stochastische systeem, maar onbereikbaar in het deterministische systeem. We bestuderen in dit proefschrift stochastische dynamische systemen waarin de toevalsfluctuaties klein zijn.

We gebruiken voor de stochastische systemen in dit proefschrift een diffusiebenadering. De systemen worden beschreven met een Fokker-Planck-vergelijking. De eerste-orde afgeleiden in deze differentiaalvergelijking corresponderen met de deterministische beweging, de tweede-orde afgeleiden met de (kleine) stochastische fluctuaties. Met behulp van de (achterwaartse) Fokker-Planck-vergelijking worden uittredingsproblemen geformuleerd. Hierbij beschouwen we in het bijzonder uittreding uit een domein bij (een gedeelte van) een rand, waar in het deterministische systeem geen uittreding kan plaatshebben. Voor deze 'onverwachte uittredingen' bepalen we de uittredingskans of de verwachte uittredingstijd. Hiertoe lossen we het bijbehorende Dirichlet-probleem asymptotisch op, gebruik makende van singuliere storingsrekening, met de diffusieparameter als kleine parameter. We vergelijken de asymptotische oplossingen met resultaten van random walk simulaties.

In hoofdstuk 2 bestuderen we het transport van vervuiling in grondwater. Hierbij is het niet voldoende alleen advectief transport te beschouwen. We dienen ook rekening te houden met macroscopische dispersie. Als gevolg van dispersie kan vervuiling een gebied bereiken, dat onbereikbaar is voor de advectieve stroming. Zo kan vervuiling worden opgepompt in een drinkwaterput, terwijl deze vervuiling in het grondwater is terechtgekomen buiten het advectieve aanstroomgebied van de put. Voor een put in een willekeurige achtergrondstroming bepalen we asymptotische uitdrukkingen voor de kans dat verontreiniging de put bereikt en voor de verwachte aankomsttijd in de put. Omdat dispersie aanzienlijk minder bijdraagt aan de verplaatsing van de verontreiniging dan advectie, gaan we uit van het advectieve stroompatroon en nemen de invloed van de dispersie mee in een grenslaag langs de separatrix in de richting van het stagnatiepunt. Voor het advectieve aanstroomgebied van de put construeren we voor de verwachte aankomsttijd in de put bovendien een composiete oplossing, die zowel binnen als buiten de grenslaag geldig is. Van het Dirichlet-probleem voor de verwachte uittredingstijd bij een specifiek deel van de rand van een domein in een meer-dimensionale toestandsruimte wordt in dit hoofdstuk een afleiding gegeven.

In de hoofdstukken 3 en 4 beschouwen we een geheel ander toepassingsgebied van deze wiskundige analyse van uittredingsproblemen. Het betreft een model uit de epidemiologie voor de verspreiding van een infectieziekte. We beschouwen een populatie, bestaande uit vatbaren, besmettelijken en herstelden, die wordt vernieuwd met constante snelheid (vervangingsgraad). We bestuderen het geval waarbij in het corresponderende deterministische systeem de ziekte endemisch wordt. In het stochastische systeem kan de ziekte uitsterven als gevolg van de stochastische fluctuaties. Wanneer de ziekte in de populatie wordt geïntroduceerd, heeft er in het stochastische model niet noodzakelijk een grote uitbraak plaats. In hoofdstuk 3 bestuderen we de kans dat de ziekte uitsterft voordat een grote uitbraak plaatsvindt. We bepalen voor deze kans een asymptotische uitdrukking voor grote populatieomvang. Bovendien bestuderen we in dit hoofdstuk de verwachte uitsterftijd van de ziekte, wanneer deze endemisch is geworden. De asymptotische oplossing van het bijbehorende Dirichlet-probleem bevat een onbekende constante, die we bepalen door de WKB-methode toe te passen op de voorwaartse Fokker-Planck-vergelijking, de stralenvergelijkingen numeriek op te lossen en de divergentiestelling toe te passen. Wanneer in dit epidemiologische model de vervangingsgraad klein is, herstelt de vatbare populatie zich slechts langzaam na een grote uitbraak van de ziekte. In deze fase is uitsterven van de ziekte zeer wel mogelijk. In hoofdstuk 4 bepalen we een asymptotische uitdrukking voor de kans dat de ziekte uitsterft na een grote uitbraak volgende op de introductie van de ziekte in de populatie.

Hoofdstuk 5 van dit proefschrift houdt zich bezig met een (vooralsnog) meer theoretisch aspect van uittredingsproblemen, namelijk het probleem van de 'eigenwijze deeltjes'. We stootten hierop bij het bepalen van de verwachte aankomsttijd van vervuiling in een put in hoofdstuk 2. Voor vervuiling die in het grondwater is terechtgekomen buiten het advectieve aanstroomgebied van de put, beperkt de door ons afgeleide oplossing zich tot stortplaatsen in de grenslaag langs de separatrix. Buiten deze grenslaag zijn meer verfijnde methoden nodig voor het bepalen van de (kleine) kans dat vervuiling de put bereikt en de verwachte aankomsttijd. Hoofdstuk 5 is een eerste stap in het onderzoek van dergelijke problemen. Voor een interval in een een-dimensionale toestandsruimte bestuderen we uittreding bij juist die rand waar uittreding het minst waarschijnlijk is. Voor drie fundamentele gevallen bepalen we de verwachte uittredingstijd, namelijk voor, respectievelijk, een interval zonder evenwicht, een interval met een instabiel evenwicht en een interval met een stabiel evenwicht. We merken op dat de afgeleide asymptotische uitdrukkingen meer inzicht geven in de verwachte uittredingstijd van de 'eigenwijze deeltjes' dan de exacte oplossingen.

Curriculum Vitae

Onno Arjen van Herwaarden werd op 16 juni 1956 te Apeldoorn geboren. Hij bezocht het Stedelijk Gymnasium te Breda van 1967 tot 1973. Na een verblijf van een jaar in de Verenigde Staten ging hij in 1974 studeren aan de Rijksuniversiteit Utrecht. In 1981 behaalde hij het doctoraal examen wiskunde met bijvak natuurkunde. Hij studeerde af in de projectieve meetkunde bij prof. dr. G.J. Schellekens. Tijdens zijn studie behaalde hij de onderwijsbevoegdheden wiskunde en natuurkunde. Van 1982 tot 1989 was hij part-time medewerker van de Vakgroep Wiskunde van de Landbouwuniversiteit Wageningen met een onderwijstaak. Tevens was hij gedurende deze periode als wiskunde-docent verbonden aan de Gemeentelijke Hogere Technische Avondschool (later Hogeschool Utrecht, Sector Elektrotechniek en Werktuigbouwkunde) te Utrecht. Vanaf september 1989 is hij full-time universitair docent bij de Vakgroep Wiskunde van de Landbouwuniversiteit, met naast onderwijs en onderzoek o.a. de coördinatie van het propaedeuse wiskunde-onderwijs als taak. Het vanaf september 1989 onder leiding van prof. dr. ir. J. Grasman verrichte onderzoek heeft geleid tot het voor u liggende proefschrift.