



ELSEVIER

Physica A 310 (2002) 347–363

PHYSICA A

www.elsevier.com/locate/physa

# Acceleration of one-dimensional mixing by discontinuous mappings

Peter Ashwin<sup>a,\*</sup>, Matthew Nicol<sup>b</sup>, Norman Kirkby<sup>c</sup>

<sup>a</sup>*School of Mathematical Sciences, Laver Building, University of Exeter, Exeter EX4 4QE, UK*

<sup>b</sup>*Department of Mathematics and Statistics, University of Surrey, Guildford GU2 7XH, UK*

<sup>c</sup>*Department of Chemical and Process Engineering, University of Surrey, Guildford GU2 7XH, UK*

Received 3 September 2001; received in revised form 6 December 2001

---

## Abstract

The paper considers a simple class of models for mixing of a passive tracer into a bulk material that is essentially one dimensional. We examine the relative rates of mixing due to diffusion, stretch and fold operations and permutation of sections of the sample. In particular we show how a combination of diffusion with permutation of sections of the sample ('chopping and shuffling') can achieve a faster rate of mixing than pure diffusion. This is done by numerical approximation of eigenvalues of certain linear operators. © 2002 Published by Elsevier Science B.V.

*PACS:* 83.50.Xa; 05.45

*Keywords:* Mixing rate; Discontinuous map; Permutation

---

## 1. Introduction: mixing in one dimension

Mixing processes in nature and technology are ubiquitous and vitally important, and can occur through a variety of physical mechanisms. For example, mixing can occur purely by diffusion at a molecular scale or it can occur by stretching and folding actions at a lengthscale comparable to the system; this can ultimately homogenize the system down to the molecular scale. Mixing via pure diffusion is rarely adequate and so (turbulent) convection is used to increase the rate of mixing over pure diffusion especially for relatively inviscid fluids. It is known that steady flows can only be chaotic in dimensions greater than two and in any case turbulence is not necessary; there has been a considerable amount of work showing that even steady or time periodic flows

---

\* Corresponding author.

E-mail address: p.ashwin@ex.ac.uk (P. Ashwin).

may have very good mixing properties, essentially due to deterministic chaos in the Lagrangian particle paths; see Refs. [1–3].

In cases where the substance to be mixed is a polymer or a paste, convection cannot be used so easily; the energetic cost of deformation may be too great or have other undesirable side-effects. Industrial processes often use static mixers in the form of vanes fixed within pipes to achieve mixing in such cases; see for example Refs. [4,5] for a range of numerical and experimental approaches to this. These static mixers use a mechanism that has been to a large extent ignored in mathematical models of mixing viscous materials, namely the effect of discontinuous ‘chop and rearrange’ transformations on the mixing (see, however, Refs. [6–9] although these all include some sort of material stretching).

In this paper we investigate and compare some possible mixing methods for materials that are essentially homogeneous in two of their dimensions, i.e., where the mixing only needs to occur in one dimension. These models are idealized but nonetheless they suggest some interesting effects. Most significantly we show how the use of two processes, diffusion and permutation alternately, can give rise to a faster asymptotic rate of mixing than either on its own, provided that the permutation is carefully chosen.

A novelty of our work is that we study mixing by composing a standard mixing process (diffusion) with a process that does not mix at all (permutation). Nonetheless, this combination of diffusion and permutation can mix better than either. We study the resulting speedup of mixing numerically for permutations of the domain cut into 3–5 pieces of equal length to find the permutations that give optimal speedup for each of these.

### 1.1. One-dimensional mixing

We consider mixing of a passive tracer in a block of material that is essentially one dimensional by application of a sequence of deterministic (i.e., non-random) operations to the block. For convenience, we always parameterize the domain describing this block by a non-dimensional distance  $x$  ranging from 0 to 1 and assume that we commence with a distribution  $p(x, 0)$  of some passive tracer in the domain. Let the distribution at time  $t$  be  $p(x, t) \geq 0$  and we normalize the total quantity so that

$$\int_{x=0}^1 p(x, 0) dx = 1.$$

We are only interested in operations that preserve the total quantity of tracer. For simplicity we consider only discrete time mixing, and so consider values of  $t = 0, 1, 2, \dots$ . We assume that the initial distribution  $p(x, 0)$  can be differentiated arbitrarily many times in  $x$  (we say  $p(x, 0)$  is *smooth* in  $x$ ). This assumption may be justified by the fact that diffusion acting over an arbitrarily small time will regularize any initial discontinuities in the concentration. However, a mixing operation such as permutation may introduce discontinuities into the concentration and the resulting concentration may only be smooth on a finite number of pieces (we say the concentration function is *piecewise smooth*) thus we must consider the larger function class of piecewise smooth functions on the interval which we denote  $\mathcal{X} = C_p^\infty(0, 1)$ . When mixing is complete

and the concentration is uniform we have  $\bar{p}(x)=1$  for all  $x \in [0, 1]$ ; we refer to this as the *uniform state*. A good mixing protocol will ensure that  $p(x, t)$  converges (within a reasonable error) to  $\bar{p}$  quickly.

### 1.2. Mixing in different norms

There are many ways of quantifying how far  $p(x, t)$  is from uniform; which one is appropriate depends on the application considered. For example, if one is trying to disinfect a food substance it is important that all locations are smoothly covered during the mixing process. In such a scenario the  $L^\infty$  norm is perhaps the most useful

$$\|p(x, t) - 1\|_\infty = \text{ess sup}_{x \in [0, 1]} |p(x, t) - 1|,$$

where

$$\text{ess sup}_{x \in [0, 1]} |p(x, t) - 1| = \inf \{M : |p(x, t) - 1| \leq M \text{ a.e. } x \in [0, 1]\}.$$

In the setting of piecewise smooth functions the above definition is equivalent to  $\|p(x, t) - 1\|_\infty = \inf \{M : |p(x, t) - 1| \leq M \text{ except for finitely many } x \in [0, 1]\}$ .

If instead we are trying to use the mixing process to achieve mixing over most of the domain as fast as possible then a more appropriate criterion would be the  $L^q$  norm

$$\|p(x, t)\|_q = \left( \int_0^1 |p(x, t) - 1|^q dx \right)^{1/q}.$$

Which  $q \geq 1$  we choose governs how much weight is given to unmixed ‘outliers’. The space  $L^2$  is a Hilbert space and this structure simplifies many calculations.

A mixing protocol may mix at different rates depending on which norm we choose to consider. We note, however, if  $0 < p \leq p'$  then  $\|\cdot\|_p \leq \|\cdot\|_{p'} \leq \|\cdot\|_\infty$ , so that as one would expect it is most difficult to achieve good mixing in the  $L^\infty$  norm. In this paper, mainly because it simplifies calculations, we will consider mixing either using  $\|\cdot\|_\infty$  or  $\|\cdot\|_2$ . For practical purposes one would like to know how many mixing operations are required to ensure that the concentration has been mixed to within, say, 5% of uniform; we refer this to the *time to 95% mixing*,  $t_{95}$  and defined to be the smallest  $t_{95} > 0$  such that

$$\|p(x, t) - 1\|_\alpha \leq 0.05$$

for all  $t > t_{95}$  where  $\alpha = 1, 2, 3, \dots$ .

We will attempt to estimate  $t_{95}$  as a function of the initial distribution, the norm chosen and the mixing protocol.

## 2. Methods of mixing

We now characterize some mixing processes either generated by diffusion or by point operations (in other words maps of  $[0, 1]$  to itself) as detailed in Section 2.2.

## 2.1. Mixing by diffusion

We consider diffusion occurring over a timestep of 1 unit with a normalized diffusivity rate  $D > 0$ . For simplicity we suppose that the domain is periodic during a diffusion step but one could similarly consider different boundary conditions (for example Neumann/no-flux) as long as they do not allow concentration to leak out of or into the domain. Although the following calculation is routine we emphasize that one can explicitly solve the diffusion equation:

$$p_t = D p_{xx} \quad (1)$$

over one unit time in the following way. Assume that  $p(x, 0) \in \mathcal{X}$  has Fourier expansion

$$p(x, 0) = \sum_{n=-\infty}^{\infty} a_n e^{2\pi i n x},$$

(assumed to be real, so  $a_{-n} = a_n^*$ , where  $*$  denotes complex conjugation) where, as usual, we can compute  $a_n = \int p(x, 0) e^{-2\pi i n x} dx$ <sup>1</sup> (note that  $a_0 = \int p(x, 0) dx = 1$ ) so that

$$p(x, t) = \sum_{n=-\infty}^{\infty} a_n e^{2\pi i n x} e^{-4\pi^2 n^2 D t}.$$

We can thus write one time step of diffusion as the operator

$$P_{\text{diff}} : \mathcal{X} \rightarrow \mathcal{X}$$

which is explicitly

$$(P_{\text{diff}}(p))(x) = \sum_n e^{2\pi i n x} \left( e^{-4\pi^2 n^2 D} \int_{y=0}^1 p(y) e^{-2\pi i n y} dy \right).$$

Note that all Fourier modes decay at an exponential rate. The slowest rate of decay corresponds to the longest wavelength and equals  $e^{-4\pi^2 D}$ . We may make a crude estimate of the decay in the  $L^2$  norm as follows:

$$\begin{aligned} \|p(x, t) - 1\|_2^2 &= \sum_{n \neq 0} |a_n|^2 e^{-8\pi^2 n^2 D t} \\ &\leq e^{-8\pi^2 D t} \sum_{n \in \mathbb{Z}, n \neq 0} |a_n|^2 = \|p(x, 0) - 1\|_2^2 e^{-8\pi^2 D t}. \end{aligned}$$

Thus  $\|p(x, t) - 1\|_2 \leq \|p(x, 0) - 1\|_2 e^{-4\pi^2 D t}$  which means that a lower bound on the time to 95% mixing for  $\|\cdot\|_2$  can be found as

$$t_{95} = \frac{\log(20\|p(x, 0) - 1\|_2)}{4\pi^2 D}.$$

<sup>1</sup> NB integrals with respect to  $x$  where no limits are marked are assumed to be definite integrals from  $x = 0$  to 1.

## 2.2. Mixing by point operations

Several mixing operations  $T$  that we consider will deform the material; we call these *point operations*. They can be considered as mappings

$$T : [0, 1] \rightarrow [0, 1]$$

such that the map is surjective, i.e., such that  $T([0, 1]) = [0, 1]$ , however we do not require the map to be invertible. We assume that the map is piecewise smooth. For any such map  $T$  we can derive an associated map on  $P(T) : \mathcal{X} \rightarrow \mathcal{X}$ ,  $p \rightarrow P(T)p$  called the *Perron–Frobenius operator* [10] in the dynamics literature

$$(P(T)p)(x) = \sum_{y=T^{-1}(x)} \frac{p(y)}{|T'(y)|},$$

where  $T'(x)$  denotes the derivative  $dT/dx$ .

In fact  $P(T)$  is a linear operator that takes a density function  $p \in \mathcal{X}$  and maps it to another density function  $P(T)p \in \mathcal{X}$ . This operator expresses how a mapping of the domain changes a concentration on that domain.

### 2.2.1. Stretching and folding

The first point operation we consider is an  $m$ -fold stretch and fold; the analysis of this map has an extensive literature (see Ref. [10]) and such maps are considered a canonical example of a mixing transformation. In light of our scenario we will consider what the action of this map implies for physical mixing of a passive scalar.

We stretch out the interval  $[0, 1]$  to be  $m$ -times its length, cut them into  $m$  equal sized sections and superimpose them on each other. This can be represented as a transformation

$$T(x) = mx(\text{mod } 1),$$

where  $m \geq 2$  is an integer. Note that since  $m$  is an integer  $T$  is a smooth mapping.  $T : [0, 1] \rightarrow [0, 1]$  is an expanding map of the interval and much is known about its mixing properties. In particular,  $T'(x) = m > 1$  and it has been shown [10] that the Perron–Frobenius operator  $P(T)$  has a unique fixed point that is the uniform density,  $\bar{p}(x) = 1$ . Moreover, for any smooth normalized initial density  $p(x, 0)$  the iterates  $(P(T)^n)p(x, 0)$  approach the uniform density at an exponential rate determined by the maximum derivative of  $p(x)$  and  $m$ . We write

$$p(x, n) = (P(T)^n)p(x, 0),$$

and note that under the assumption  $p(x, 0)$  is  $C^1$  one can show

$$\|p(x, n) - 1\|_\infty \leq \frac{\|p_x(x, 0)\|_\infty}{m^n},$$

where  $p_x(x, 0)$  is the partial derivative in the  $x$  direction. Note that the Perron–Frobenius operator does not increase the maximal spatial derivative i.e.,

$$\|p_x(x, n)\|_\infty \leq \|p_x(x, 0)\|_\infty.$$

Using this we can obtain a prediction of the time to 95% mixing (i.e.,  $\|p(x, n) - 1\|_\infty < \frac{1}{20}$ ) for  $m$ -fold stretch and fold, namely

$$t_{95} = \frac{\log(20\|p_x(x, 0)\|_\infty)}{\log m}.$$

### 2.2.2. Permutation of cells

A second class of point operations that we consider consist of operations that partition the domain into a finite number of equal sized pieces ('cells') and then rearrange the cells. We refer to such an operation as a *permutation of cells*. More precisely, suppose that the interval is divided into  $N$  subintervals (called 'cells')

$$[0, 1/N), [2/N, 3/N), \dots, [(N-1)/N, 1).$$

We consider a permutation  $\sigma \in S_N$  where  $S_N$  is the group of all permutations of the symbols  $\{1, \dots, N\}$ . This defines a mapping on  $[0, 1]$  defined by

$$T_\sigma(x) = x + \frac{\sigma(i) - i}{N}$$

whenever  $x \in [(i-1)/N, i/N)$ . The permutation acts on  $p(x)$  as  $P(T_\sigma) \circ p(x) = p(T_\sigma^{-1}x)$ .

Note that after applying  $T_\sigma$  a finite number of times we must return to the original density:

$$P(T_\sigma)^K p(x) = p(x).$$

This is because each  $\sigma$  is of finite order. Therefore permutation of cells cannot lead to mixing on its own. Moreover each point in the domain must return to its original location in a periodic manner. In other words if  $p(x, t) = P(T_\sigma)^t p(x, 0)$  then  $\|p(x, t) - 1\| \not\rightarrow 0$  for any norm we choose whenever the initial state is not uniform:  $\|p(x, 0) - 1\| \neq 0$ .

## 3. Mixing via permutation of cells

In this section we examine permutation of cells in more detail. Although such operations cannot mix on their own, they can considerably accelerate mixing by diffusion. We consider an initial  $p(x) \in \mathcal{X}$  in terms of its Fourier expansion

$$p(x) = \sum_{-\infty}^{\infty} a_n e^{2\pi i n x}$$

with  $a_n = \int p(x) e^{-2\pi i n x} dx$  (sums without limits in this section are assumed to run from  $-\infty$  to  $\infty$ ). Under permutation this initial distribution is transformed to a distribution  $q(x)$  where

$$q(x) = \sum_n b_n e^{2\pi i n x}.$$

We may calculate

$$b_n = \int_0^1 p(T_\sigma^{-1}(x)) e^{-2\pi i n x} dx$$

$$\begin{aligned} &= \int_0^1 p(x) e^{-2\pi i n T_\sigma(x)} dx \\ &= \sum_k a_k \int e^{2\pi i k x - 2\pi i n T_\sigma(x)} dx. \end{aligned}$$

Thus the transformation is a linear transformation of the complex Fourier coefficients:

$$b_n = \sum_{k=-\infty}^{\infty} \tau_{nk} a_k,$$

where for any  $k$  we have  $\sum_n |\tau_{nk}|^2 = 1$  (and thus  $|\tau_{nk}| \leq 1$ ). We can compute

$$\begin{aligned} \tau_{nk} &= \int e^{2\pi i k x - 2\pi i n T_\sigma(x)} dx \\ &= \sum_{l=1}^N \int_{(l-1)/N}^{l/N} e^{2\pi i k x - 2\pi i n x - (2\pi i n/N)(\sigma(l)-l)} dx \\ &= \sum_{l=1}^N e^{(-2\pi i n/N)(\sigma(l)-l)} \int_{(l-1)/N}^{l/N} e^{2\pi i (k-n)x} dx. \end{aligned}$$

Define the primitive  $N$ th root of unity as follows:

$$w = e^{-2\pi i/N}$$

and then the diagonal terms can be written as

$$\tau_{nn} = \frac{1}{N} \sum_{l=1}^N w^{n(\sigma(l)-l)}. \quad (2)$$

When  $k \neq n$  we obtain the off-diagonal terms:

$$\tau_{nk} = \frac{(1 - w^{(k-n)})}{2\pi i(k-n)} \sum_{l=1}^N w^{n\sigma(l)-kl}. \quad (3)$$

We now consider some example permutations and the structure of the associated transformation  $(\tau_{nk})$ .

### 3.1. Some example permutations

There are a number of ways of representing permutations; in simple cases for  $n$  cells we will write them as a formula that gives an invertible map of the integers  $\{1, \dots, n\}$  to themselves. More generally we will use *disjoint cycle notation*. This is a way of uniquely representing the permutation as a number of cycles of possibly different lengths such that each symbol appears in at most one cycle. For example  $(13)(245)$  is the permutation that interchanges the 1st and 3rd places while cyclically permuting those in the 2nd, 4th and 5th places. The empty brackets  $()$  indicate the trivial (identity) permutation. For more details, see any introductory text on group theory, for example Ref. [11].

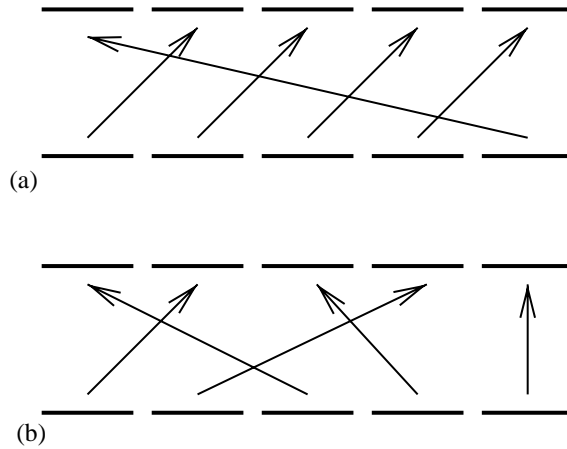


Fig. 1. Schematic diagram showing the action of (a) primary rotation and (b) interleaving of  $N = 5$  equal sized cells on the one-dimensional domain.

### 3.1.1. Simple rotation

A simple rotation of the cells can be expressed as the permutation

$$\sigma(j) = j + m \pmod{N},$$

where  $m$  gives the shift in cells; this corresponds to the map  $T(x) = x + m/N \pmod{1}$  on the domain; see Fig. 1(a). In this case we can compute

$$\tau_{nm} = w^{nm}, \quad \tau_{nk} = 0 \text{ for } n \neq k.$$

Thus the transformation is diagonal and moreover  $|\tau_{nn}| = 1$  for all  $n$ . This means that the only effect of the simple rotation is to shift the phase of the modes but no energy is exchanged between the modes.

### 3.1.2. Secondary rotation

Suppose that  $N = ab$  is a product of two integers  $a > 1$  and  $b > 1$ . Let  $1 \leq r < a$ . We consider a permutation  $\sigma$  which within each block of  $a$  cells performs a rotation by  $r$ , i.e., given any  $0 \leq k < b$  and  $1 \leq l < a$  we define

$$\sigma(ka + l) = ka + (l + r \pmod{a}).$$

This transformation can transport energy between modes and will typically introduce discontinuities to  $p(x)$ ; however, the entries in  $\tau_{nk}$  are more difficult to compute and we do not have explicit formulae for all  $(a, b)$ .

### 3.1.3. Interleaving

Now we consider a permutation we refer to as ‘interleaving’ because of its action on cells. We split the cells into two groups of almost equal size and then reassemble by interleaving them; see Fig. 1(b).



Assume  $N$  is odd and define  $\sigma(l) = 2l \pmod{N}$  where mod takes the representative in  $\{1, \dots, N\}$ . We calculate

$$\tau_{nn} = \frac{1}{N} \sum_{l=1}^N w^{nl}. \quad (4)$$

Thus  $\tau_{nn} = 0$  if  $N$  is not a factor of  $n$  and 1 if  $N$  is a factor of  $n$ . It is possible to calculate

$$\tau_{nk} = \frac{(1 - w^{(k-n)})}{2\pi i(k-n)} \sum_{l=1}^N w^{2nl-kl} \quad (5)$$

which is equal to zero if  $N$  is not a factor of  $(2n - k)$  and  $N(1 - w)/2\pi i(k - n)$  otherwise. Thus

$$|\tau_{nk}| = \begin{cases} 0 & \text{if } N \text{ is not a factor of } (2n - k), \\ \leq 1 & \text{if it is a factor.} \end{cases}$$

**Remark 1.** In all such permutation operations one can check that the matrix with entries  $(\tau_{nk})$  is unitary in the sense that

$$\sum_n \tau_{nk} \tau_{nl}^* = \delta_{kl}$$

and the sum of the moduli of entries in each row and column equals 1 *i.e.*, for all  $k, n$ ,  $\sum_{j=-\infty}^{\infty} |\tau_{nj}|^2 = \sum_{j=-\infty}^{\infty} |\tau_{jk}|^2 = 1$ .

#### 4. Mixing using permute-diffuse

We now consider mixing via multi-stage mixing protocols, namely when members of a sequence of maps are applied one after the other. Namely, we consider  $P_{\text{diff}} \circ P(T_\sigma)$  where  $T_\sigma$  is some permutation of cells (given two operators  $A$  and  $B$  the composition is written  $A \circ B$ ). If the action of  $P(T_\sigma)$  introduces discontinuities to a continuous function then in doing so it will send energy into modes  $e^{2\pi i k x}$  with  $k$  high. In this way it can increase the rate of mixing over that by pure diffusion. How much this is achieved and how effectively depends on the chosen permutation. We first consider some generalities before re-examining the examples discussed in Section 3 in turn.

Note that  $P_{\text{diff}}$  acts on Fourier coefficients as a diagonal operator where the  $n$ th entry is  $e^{-4\pi^2 n^2 D}$ . Hence if we commence with a concentration density

$$p(x) = \sum_n a_n e^{2\pi i n x}$$

after one timestep we have

$$q(x) = P_{\text{diff}}(P(T_\sigma)p)(x) = \sum_{n=-\infty}^{\infty} \left( \sum_{k=-\infty}^{\infty} e^{-4\pi^2 n^2 D} \tau_{nk} a_k \right) e^{2\pi i n x}.$$

In terms of the Fourier coefficients  $b_n$  of  $q(x) = \sum b_n e^{2\pi i n x}$  we have

$$b_n = \sum_k e^{-4\pi^2 n^2 D} \tau_{nk} a_k = \sum_k m_{nk} a_k,$$

where we define  $m_{nk} = \rho^{n^2} \tau_{nk}$  and  $\rho = e^{-4\pi^2 D}$ .

Let us now consider the specific case of the interleaving permutation with  $N = 3$ . The minimal rate of mixing for pure diffusion is given by

$$\|p(x, n) - 1\| \leq C \rho^n.$$

This mixing rate is governed by  $\rho$ , the largest eigenvalue of the diagonal matrix with entries  $d_{ij} = \rho^{i^2} \delta_{ij}$  (excluding the eigenvalue 1).

The largest eigenvalue  $\rho$  (excluding the eigenvalue 1) gives the *asymptotic rate of mixing*. If  $p(x, 0)$  is a typical concentration with components in all modes i.e.,  $a_n \neq 0$  for each  $n$ , then  $\rho$  is the smallest number such that there exists a constant  $C$  with the property

$$\|p(x, n) - 1\| \leq C \rho^n$$

for all  $n \geq 0$ . Unfortunately estimates of the constant  $C$  are in general difficult to obtain, and the value of the smallest such constant  $C$  will depend upon both the mixing protocol and the initial concentration. In this paper we will *assume* that the value of the smallest such constant  $C$  is uniform over the range of initial concentrations and protocols that we examine.

The standing assumption of this paper will be that the best rate of mixing is achieved when the leading non-unit eigenvalue of the  $L^2$  operator associated to the multi-stage mixing protocol is minimized. To simplify notation we will call the leading non-unit eigenvalue simply the *leading eigenvalue*.

We approximate the Fourier expansion using the finite number of modes  $\{a_n : -N_t \leq n \leq N_t\}$  and compute the truncated matrices  $\{d_{ij}\}$ ,  $\{\tau_{ij}\}$  and  $\{m_{ij}\}$  above as a function of the diffusion parameter  $\rho = e^{-4\pi^2 D}$ . The (typically complex) eigenvalues of the truncated matrices were found using Maple,  $\lambda_d^i$  being the  $i$ th eigenvalue of  $\{d_{ij}\}$  and  $\lambda_{pd}^i$  being the  $i$ th eigenvalue of  $\{m_{ij}\}$ . Table 1(a) shows the first five eigenvalues for  $\rho = 0.5$  computed by this method to six decimal places. Observe that the combination of permute-diffuse gives rise to a leading eigenvalue (eigenvalue with maximum modulus) that is appreciably smaller than the pure diffusion case. Table 1(b) illustrates the same thing for the case of stronger diffusion,  $\rho = 0.1$ .

For comparison, Table 1(c) compares the results for the Fourier expansion (FE) with a finite difference (FD) computation, using Crank–Nicolson with  $N_t$  timesteps and  $N_s$  gridpoints in space, for each diffusive step. The two methods can be seen to produce similar results for the leading eigenvalue in this case  $\rho = 0.5$  ( $D = 0.017557$  in (1)). For the finite difference calculations, the eigenvalues of the associated linear transformation were computed using Maple. The leading eigenvector (slowest decaying mode) is shown in Fig. 2 computed by the FD computation.

Table 1

Absolute values of the eigenvalues of the pure diffusion operator  $\lambda_d^i$  and the permute-diffuse operator  $\lambda_{pd}^i$  ordered by decreasing size for the cases (a)  $\rho = 0.5$  and (b)  $\rho = 0.1$  from the truncated Fourier expansion (FE) calculations. Observe the unity eigenvalues corresponding to the uniform state and that the eigenvalues (except for 1) come in pairs corresponding to the real and imaginary parts of complex modes. Note that  $|\lambda_d^{2i-1}| = |\lambda_d^{2i}| = \rho^i$ . These eigenvalues were computed using a Fourier truncation  $N_t = 30$ . (c) Compares the leading eigenvalues for the FE with finite-difference (FD) calculations for grid-size  $(N_s, N_t)$  (See text)

(a)		$i = 0$	$i = 1$	$i = 2$	$i = 3$	$i = 4$
	$ \lambda_d^i $	1.0	0.5	0.5	0.0625	0.0625
	$ \lambda_{pd}^i $	1.0	0.279769	0.279769	0.085946	0.085946
(b)	$i$	0	1	2	3	4
	$ \lambda_d^i $	1.0	0.1	0.1	0.001	0.001
	$ \lambda_{pd}^i $	1.0	0.041514	0.041514	0.000185	0.000185
(c)		Fourier	Finite difference			
		$N_t = 30$	$(N_s, N_t) = (90, 25)$		$(N_s, N_t) = (45, 10)$	
	$ \lambda_d^1 $	0.5	0.5001		0.5005	
	$ \lambda_{pd}^1 $	0.2797	0.2805		0.2813	

One can similarly find the ratio of leading eigenvalues as a function of  $\rho$ ; Fig. 3 plots

$$R = \left| \frac{\lambda_{pd}^1}{\lambda_d^1} \right|$$

from the FE calculations, on varying  $\rho$  between 0 and 1. Note that (a)  $R < 1$  for all  $0 < \rho < 1$ ; this means that permute-diffuse mixes asymptotically better than pure diffusion for all diffusion rates; (b) near  $\rho=1$  we see that  $R$  is proportionate to  $\rho$ , suggesting that for slow diffusion we have a speed-up such that one step of permute-diffuse is almost as good as two steps of pure diffusion; (c) It appears that  $R \geq R_0 = 0.4135$  and  $R$  approaches this value as  $\rho \rightarrow 0$ . This implies that we cannot speed up the decay of the  $L^2$  norm better than by this ratio at every step.

**Remark 2.** The limiting value  $R_0$  seems to correspond to  $3\sqrt{3}/4\pi = 0.4134966717$ . Note that there is one term of order  $\rho$  in  $\{m_{ij}\}$  which has coefficient  $3\sqrt{3}/2\pi$  which suggests that the limit can be understood by analytical methods.

*Implications for time to 95% mixing.* Recall that there will be a constant  $C \geq 0$  (depending on the mixing protocol) such that for all initial concentrations  $p(x, 0)$ ,

$$\|p(x, n) - 1\| < \|p(x, 0) - 1\| C |\lambda^1|^n,$$

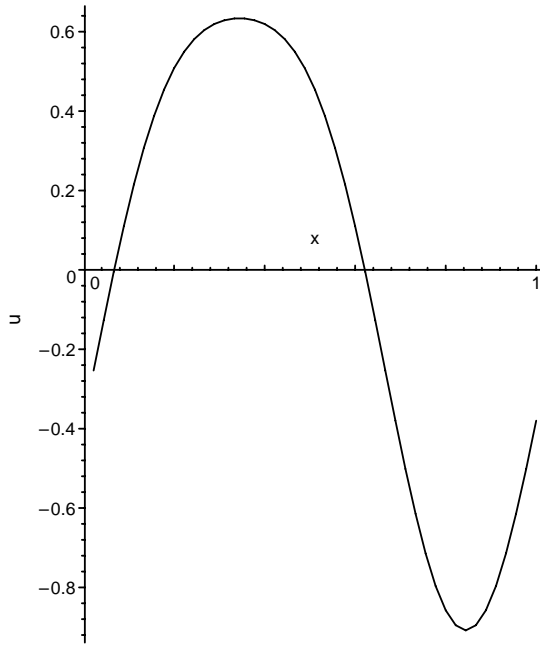


Fig. 2. The slowest decaying eigenvector corresponding to the leading eigenvalue, computed using finite difference with  $(N_s, N_t) = (45, 10)$  for  $N = 3$  interleaving followed by diffusion with  $\rho = 0.5$ . Observe the asymmetry; it is clearly not a pure Fourier mode. This profile shifted by one-third of a period will also decay at the same rate, interchanging with the same profile shifted by two-thirds of a period.

where  $p(x, n) = P^n p(x, 0)$  and  $|\lambda^1|$  is the leading eigenvalue of  $P$ . Thus the time to 95% mixing will be at most

$$t_{95} = \frac{\log(20C\|p(x, 0) - 1\|)}{-\log|\lambda^1|}.$$

Hence, under the assumption that  $C$  is a constant over mixing protocols for the same initial distribution  $t_{95}$  will be approximately decreased by a factor given by the ratio of the logarithm of the leading eigenvalues for the mixing protocols. For example, for  $N = 3$  and  $\rho = 0.5$  we estimate that 95% mixing using permute-diffuse will happen in 0.54 the time it would take to achieve 95% mixing using pure diffusion. Fig. 4 shows an example of computed decay measured as  $L = \ln(\|p(x, n) - 1\|_2)$  with  $n$  for an initial (discontinuous) concentration  $p(x, 0) = 0.9548 + 0.5 \sin(1.7\pi x)$ . The pure diffusion decay (dashed line) occurs with slope at most  $\ln(\rho)$  whereas the permute-diffuse decay (solid line) has asymptotic slope  $\ln(\lambda_{pd}^1)$  as expected.

The above considerations suggest some rather interesting questions:

- Q1. For a given number of cells  $N$  and diffusion constant  $\rho$ , which permutation leads to the optimum speedup, i.e., what is the  $\sigma \in S_N$  such that

$$P_{\text{diff}} \circ P(T_\sigma)$$

has leading eigenvalue with minimal modulus?

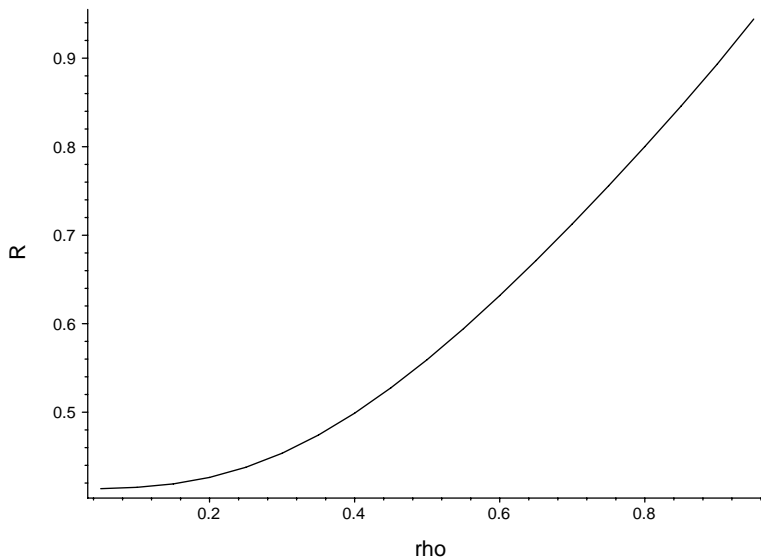


Fig. 3. The rate  $R$  of increase of mixing as a function of  $\rho$  for the interleaving with  $N = 3$ . Note that this lies below  $R=1$  at all values and so the transformation leads to faster asymptotic mixing than pure diffusion for all diffusion rates. This was computed using the truncation  $N_t = 30$ . Note the approach to  $R = 0.4135$  as  $\rho \rightarrow 0$ .

Q2. Is this optimal permutation independent of the diffusion constant  $\rho$ ?

The  $\sigma$  obtained in answer to Q1 would be in some sense the best way to permute the cells after cutting them up. Note that if  $\sigma$  is an optimal permutation then any conjugate of  $\sigma$  by the  $N$ -cycle  $(1, 2, \dots, N)$  would also be optimal and so one cannot expect the optimal permutation to be unique. We now address these questions for  $N = 3, 4$  and  $5$ .

#### 4.1. Optimal speedup of permute-diffuse for $N = 3$

From the numerical calculations discussed in detail above we see that interleaving with  $N = 3$  gives an asymptotic speedup of mixing by using a mixture of permutation and diffusion over that obtained just using diffusion. Note that there are only six elements in  $S_3$ ; these are either transpositions, the identity or 3-cycles. Hence for the case  $N = 3$  we can answer Q1—at least numerically; any permutation that is a simple transposition (exchange of any two elements) will lead to the same speedup which is optimal. Hence, if the transposition causes any speedup, Q2 must also be true; the optimal permutation is independent of the value of  $\rho$ .

#### 4.2. Optimal speedup of permute-diffuse for $N = 4$

The situation for  $N = 4$  is more interesting; Table 2 lists cycle types and the corresponding values of  $\lambda_{pd}^1$  for  $\rho = \lambda_d^1 = 0.5$  using the finite different approximation with

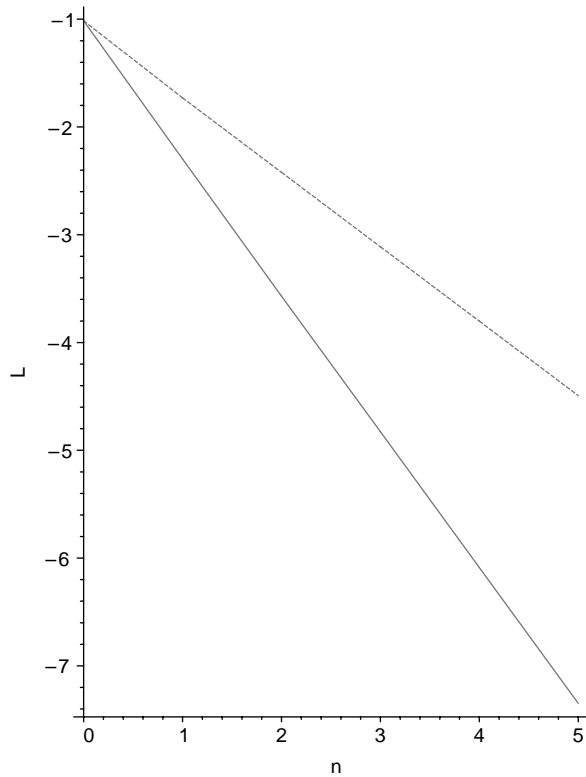


Fig. 4. Progress to mixing as measured by  $L = \ln(\|p(x, n) - 1\|_2)$  versus  $n$  for  $\rho = 0.5$ , comparing pure diffusion (dashed line) and permute-diffuse with  $N = 3$  interleaving (solid line). Identical initial distributions were considered and the finite difference method was used with grid  $(N_s, N_t) = (30, 10)$ ; note the steeper slope of decay for the permute-diffuse operation, implying a faster time to 95% mixing.

40 space and 10 time intervals per step. Observe that

1. Permutations with the same cycle type (e.g. the 4-cycles here) may have different leading eigenvalues.
2. The smallest leading eigenvalue is for the three cycle (123) with a ‘speed-up ratio’ of 0.497 over pure diffusion.

We have investigated several other values of  $\rho$ , and these give rise to leading eigenvalues preserving the corresponding order in Table 2.

#### 4.3. Optimal speedup of permute-diffuse for $N = 5$

For  $N=5$  intervals we have a choice of  $5! = 120$  different permutations. Nevertheless, numerical experiments indicate that there are optimal permutations, namely those with a 4-cycle of the form (2453). The absolute values of leading eigenvalues are listed in

Table 2

Absolute values of the leading eigenvalue for permute-diffuse mixing as a function of permutation for permutation of four equal length subintervals

Permutation	$\lambda_{pd}^1$
(123),(243),(134)	0.2494
(12)(34), (14)(23)	0.3223
(13),(24)	0.3223
(1243),(1324)	0.3967
(12),(14),(23),(34)	0.4239
(1234),(13)(24),()	0.5

These values were obtained using the finite-difference approximation with  $\rho = 0.5$  and  $(N_s, N_t) = (40, 10)$ . The optimal asymptotic rate of mixing is found for any three cycle of the subintervals.

Table 3

Absolute values of the leading eigenvalue for permute-diffuse mixing as a function of permutation for permutation of five equal length subintervals

Permutation	$\lambda_{pd}^1$
(1342)	0.1266
(134)	0.1663
(12354)	0.2101
(12453)	0.2824
(12)(354)	0.3018
(13)(254)	0.3283
(123)	0.3461
(1234)	0.3472
(14)(23)	0.3837
(1324)	0.3886
(12)(34)	0.3976
(13)(24)	0.4167
(13)	0.4196
(12)	0.4718
()	0.5017

These values were obtained using finite-difference with  $\rho = 0.5$  and  $(N_s, N_t) = (25, 10)$ . The optimal speedup is found for certain four-cycles of the subintervals. Only representative permutations are listed; all permutations fall into one of these classes. Note that the fastest speedup is achieved for  $N = 5$  interleaving.

Table 3. We have tried several values of  $\rho$  and find that the optimal permutations appear to be independent of  $\rho$ .

However, the ordering is not preserved at different values of  $\rho$  even after accounting for numerical inaccuracy of the computations here. For example, for  $\rho = 0.5$  we have that (12354) mixes asymptotically faster than (12)(354); their leading eigenvalues have modulus 0.210 and 0.301, respectively. By contrast, for  $\rho = 0.95$  these are reversed in order and their eigenvalues have modulus 0.827 and 0.812, respectively. This leads us

to suspect that the optimal choice of permutation to realize the best asymptotic rate of mixing depends on  $\rho$ .

## 5. Discussion

We have shown that by using a mixture of discontinuous (permutation) and smooth mixing processes one can improve the asymptotic mixing rates in one dimensional domains. This work complements much of that done on deterministic mixing in that it is merely one dimensional, but uses the discontinuity in an important way. In principle one can make progress on similar problems in two- or three-dimensional domains in the same way.

More direct derivations from practical (non-turbulent) mixing regimes would be very helpful – in fact necessary to make the project applicable. One could in principle model mixing in turbulent flow by a mixing protocol including discontinuous maps (approximating detached boundary layers within the turbulence) that vary chaotically with time.

A generalization of the permutation of cells considered here are the *interval exchange maps* [12, p. 470]; these allow us to model permutations of cells of differing sizes and can give rise to maps that are not periodic. Note however that  $\|p(x, t) - 1\|$  will still not change with  $t$  and so they trivially cannot mix in the sense of any such norm decaying. In fact, the mixing properties of interval exchange maps are very subtle and relatively poorly understood and depend on parameters in a sensitive way. For this reason we have not considered these more general maps in this paper.

### 5.1. Implications for applications to mixing

Although the model considered here is too idealized to apply directly to mixing of a substance within a highly viscous material, it does give a quantitative study of a number of features that we expect to see in such situations:

1. Combining permutation and diffusive mixing can give asymptotic mixing rates considerably better than either acting alone.
2. Even permutations that do not mix on their own can be used to effectively speed up mixing.
3. The effectiveness is sensitive on the form of permutation used; this is paralleled in the work of Clifford et al. [7] where the arrangement of lamellar structures affects reaction yield.
4. As a general principle, this work implies that it is possible to find an optimal way of cutting up and shuffling parts of the domain as a way of improving the asymptotic mixing rate.



*Further questions.* We finish by stating some other questions that present themselves as extensions of this work:

1. We have focused on improving the asymptotic mixing rate or equivalently reducing the magnitude of the leading eigenvalues. How much in general does the constant  $C$  vary and how good a guide to the time to achieve 95% mixing is the leading eigenvalue?
2. Suppose instead of taking the same permutation every time, we choose one from a number of possible shuffles before applying the diffusion constant; this could be viewed as a random dynamical system. Is there any way to improve the asymptotic mixing rate over what one could obtain using an optimal permutation repeatedly? We believe not. By comparison with Refs. [6,8] we expect simple periodic protocols to work best.
3. Rather than fixing on a mixing protocol one could consider the following control problem: given a distribution can one find an algorithm to pick the best permutation to mix it? This would mean we need to split up regions where the largest gradients exist.

## Acknowledgements

We thank J.M. Smith for encouragement and some very helpful conversations concerning this paper, and also A.D. Gilbert and A.R. Champneys for their helpful comments.

## References

- [1] H. Aref, Stirring by chaotic advection, *J. Fluid Mech.* 143 (1984) 1–21.
- [2] J.M. Ottino, Mixing, chaotic advection and turbulence, *Ann. Rev. Fluid Mech.* 22 (1990) 207–253.
- [3] J.M. Ottino, Mixing and chemical reactions—a tutorial, *Chem. Eng. Sci.* 49 (1994) 4005–4027.
- [4] H.E.A. van den Akker, J.J. Derksen (Eds.), *Proceedings of the 10th European Conference on Mixing*, Delft, Netherlands, July 2–5, 2000, Elsevier, Amsterdam, 2000.
- [5] E. Saatdjian, N. Midoux, M.I.G. Chassaing, J.C. Leprevost, J.C. Andre, Chaotic mixing and heat transfer between confocal ellipses: experimental and numerical results, *Phys. Fluids* 8 (1996) 677–691.
- [6] D. D'Alessandro, M. Dahleh, I. Mezic, Control of mixing in fluid flow: a maximum entropy approach, *IEEE Trans. Automat. Control* 44 (1999) 1852–1863.
- [7] M.J. Clifford, S.M. Cox, E.P.L. Roberts, Reaction and diffusion in a lamellar structure: the effect of the lamellar arrangement on yield, *Physica A* 262 (2001) 294–306.
- [8] F.H. Ling, The effect of mixing protocol on mixing in discontinuous cavity flows, *Phys. Lett. A* 177 (1993) 331–337.
- [9] Z. Toroczkai, G. Károlyi, Á. Péntek, T. Tél, Autocatalytic reactions in systems with hyperbolic mixing: exact results for the active baker map, *J. Phys. A* 34 (2001) 5215–5235.
- [10] A. Lasota, M. Mackey, *Chaos, Fractals and Noise*, Applied Math Science, Vol. 97, Springer, Berlin, 1994.
- [11] W. Ledermann, *Introduction to Group Theory*, Longman, New York, 1973.
- [12] A. Katok, B. Hasselblatt, *Introduction to Modern Theory of Dynamical Systems*, Cambridge University Press, Cambridge, 1995.