

# Cumulative effect of random linear transformations\*

S.G. Dani

August 2, 2005

## Abstract

Let  $V$  be a finite-dimensional real vector space. It is well known that if we start with a vector  $v$  and repeatedly apply a linear transformation  $T$  to it, then the resulting trajectory is either bounded, or diverges. This talk will be concerned with the analogous question as to what happens if we apply not a fixed linear transformation  $T$ , but linear transformations chosen randomly with respect to a probability law. It turns out that a similar phenomenon as above holds in this general case also, under appropriate conditions. The work on the theme involves also understanding how products of random matrices behave and, in technical terms, the concentration functions of convolution powers of probability measures on linear groups, which will be described.

Let  $V$  be a finite-dimensional real vector space (that is,  $V = \mathbb{R}^d$  for some  $d$ ), and let  $\text{GL}(V)$  be the group of nonsingular linear transformations of  $V$ . Let  $v_0 \in V$  and  $T \in \text{GL}(V)$  and consider the trajectory  $\{v_0, Tv_0, T^2v_0, \dots\}$ . Using the Jordan decomposition of  $T$  it can be seen that the trajectory either goes off to infinity (i.e.  $\|T^i v_0\| \rightarrow \infty$  as  $i \rightarrow \infty$ ) or is bounded (contained in a compact subset); in other words, if there is any subsequence in the trajectory going to infinity, then the whole trajectory goes to infinity; this is a special feature of the dynamics of this class of systems, in the general class of dynamical systems. In a general dynamical system the behaviour along different subsequences can be drastically different.

---

\*Abdi Endowment Lecture, at the 20 th annual conference of the Ramanujan Mathematical Society, Calicut, 30 July 2005.

Here we want to consider an analogue of this for *random linear transformations*. This means that we may not apply at each successive stage a fixed linear transformation  $T$  as above, but apply possibly different transformations, selected with certain preassigned probabilities. For example let  $T_1, \dots, T_k \in \text{GL}(V)$  and  $p_1, \dots, p_k \in [0, 1]$  such that  $p_1 + \dots + p_k = 1$ , and consider the “random transformation” which sends a given vector  $v$  in  $V$  to  $T_j v$  with probability  $p_j$ ,  $j = 1, \dots, k$ . We then repeat the process. The probability that we move from a point  $v \in V$  to a point  $v' \in V$  in two steps is given by  $\sum p_{j_1} p_{j_2}$ , the summation being taken over all pairs  $(j_1, j_2)$  such that  $T_{j_1} T_{j_2} v = v'$ ; more generally, the probability of going from a point  $v$  to  $v'$  in  $n$  steps is given by  $\sum p_{j_1} p_{j_2} \dots p_{j_n}$ , the summation being taken over all  $n$ -tuples  $(j_1, j_2, \dots, j_n)$  such that  $T_{j_1} T_{j_2} \dots T_{j_n} v = v'$ .

We shall now express the probabilities as above in a concise form as follows: Let  $\mu$  be the probability measure on  $\text{GL}(V)$  defined by  $\mu(E) = \sum_{T_j \in E} p_j$ , for all Borel subsets of  $E$  (notice that  $\text{GL}(V)$  is a locally compact topological space and hence has a Borel structure arising from it). On the space of probability measures on  $\text{GL}(V)$  we have the *convolution product* defined, given two probability measures  $\mu_1$  and  $\mu_2$  on  $\text{GL}(V)$ , by

$$\mu_1 * \mu_2(E) = \int \chi_E(xy) d\mu_1(x) d\mu_2(y),$$

for any Borel set  $E$  of  $\text{GL}(V)$ ,  $\chi_E$  being the characteristic function of  $E$ ; (notice that in this definition it does not matter whether  $\mu_1, \mu_2$  are discrete measures as above, or more general measures, provided they are finite, ensuring that the integral is finite). In particular we can form the  $n$ th power of any probability measure  $\mu$  (not necessarily discrete), which we shall write as  $\mu^n$  and call the  $n$ th *convolution power* of  $\mu$ . Also, given a probability measure  $\mu$  on  $\text{GL}(V)$  and a probability measure  $\nu$  on  $V$  we can define an action of  $\mu$  on  $\nu$  by

$$\mu * \nu(E) = \int \chi_E(gv) d\mu(g) d\nu(v),$$

for all Borel sets  $E$  of  $V$ , with  $\chi_E$  the characteristic function of  $E$  on  $V$ .

It is then straightforward to verify that the probability of moving from  $v \in V$  to  $v' \in V$ , for the measure  $\mu$  as in the preceding discussion that we computed numerically, can now be expressed as  $\mu^n * \delta_v(\{v'\})$ , where  $\delta_v$  is the point measure based at the point  $v$ .

In the place of the measure  $\mu$  as above which was supported on a discrete set, we may also consider a more general probability measure. In this case the rule for

transition is to be understood as follows: suppose we start from an initial point  $v \in V$ . Then the probability of landing in a (Borel) subset  $E$  of  $V$  is given by  $\mu(\{g \in \text{GL}(V) \mid gv \in E\})$ ; in the case when  $E$  is the singleton subset  $v'$  this corresponds to the value assigned above. Observe that the probability of moving from  $v \in V$  to a specific point  $v' \in V$  could be 0 in general, while the probability of landing in certain nonempty open subsets is positive. The cumulative effect of such a process, starting from an initial point  $v$ , is that the probability of being in a Borel subset  $E$  after (exactly)  $n$  steps is  $\mu^n(\{g \in \text{GL}(V) \mid gv \in E\})$ , where  $\mu^n$  is the  $n$ th convolution power defined above (for a general measure). It is also the same as  $\mu^n * \delta_v(E)$ .

We may also let the initial point  $v$  to be “random”, chosen according to a probability law, say  $\nu$ . In other words, the point  $v$  in the above discussion need not be a specific one, but a “random point” whose probability of being in a Borel subset  $E$  is given by  $\nu(E)$ ,  $\nu$  being a probability measure on  $V$ , the “initial distribution”. In this case the cumulative effect of application of the random transformations corresponding to a probability law  $\mu$  on  $\text{GL}(V)$  is that the probability of being in a subset  $E$  after  $n$  steps is given by  $\mu^n * \nu(E)$ , where  $\mu^n * \nu$  is the probability measure on  $V$  defined as above. Thus  $\mu^n * \nu$  is the corresponding probability distribution.

We are thus led to studying, in this framework of random transformations, the sequence of probability measures  $\{\nu, \mu * \nu, \dots, \mu^i * \nu, \dots\}$  as the analogue of the trajectory  $\{v_0, Tv_0, \dots, T^i v_0, \dots\}$  in the deterministic setup. In the deterministic setup we noted that the trajectory tends to infinity unless it is bounded. What is the analogue to look for in the more general random framework? There are a few different natural possibilities in extending the question, especially the aspect of going to infinity. One analogue (that we shall be interested in) for the sequence of probability measures  $\mu^i * \nu$ ,  $i = 1, 2, \dots$ , of the trajectory  $T^i v_0$  going to infinity is that  $\mu^i * \nu(K) \rightarrow 0$  for all compact subsets  $K$  of  $V$ ; the latter condition means, intuitively, that the associated probability distributions “dissipate” to infinity. It is also possible to consider, instead, the behaviour of the “sample paths”, such as  $T_{j_i} T_{j_{i-1}} \cdots T_{j_1} v$ ,  $i = 1, 2, \dots, i \dots$ , in the case of the introductory example above, and ask whether almost all of them (in a sense that we shall not go into here) go to infinity. This is a stronger question. There is a lot of literature on the latter question; the pioneering work of H. Furstenberg in the 1960’s on the issue has been followed up by several authors; see [6] and references therein. This stream however involves stronger hypothesis (and deeper techniques) in both analytic and algebraic terms (condition of finiteness of moments, strong irreducibility etc.). We shall not

concern ourselves with it here, beyond pointing out some interconnections.

Given a probability measure  $\mu$  we denote by  $S(\mu)$  and  $G(\mu)$ , respectively the smallest closed subsemigroup and the smallest closed subgroup of  $\text{GL}(V)$ , whose complement in  $\text{GL}(V)$  has zero measure; thus  $S(\mu)$  is the smallest closed subsemigroup containing the support of  $\mu$ , and  $G(\mu)$  is the smallest closed subgroup containing the support of  $\mu$  (we recall that the “support” of  $\mu$  is the smallest closed subset whose complement has zero measure). As all the random transformations involved essentially come from the support of  $\mu$ , the analogue of the trajectory  $\{T^i v \mid i = 1, 2, \dots\}$  being bounded, is that  $\{gv \mid g \in S(\mu)\}$  be bounded.

In the light of the observations above, in analogy of the dichotomy that the trajectory  $\{v_0, Tv_0, \dots T^i v_0, \dots\}$  either goes off to infinity or is bounded we may ask, given the probability measures  $\mu$  on  $\text{GL}(V)$  and  $\nu$  on  $V$ , whether it is true that either  $\mu^i * \nu(K) \rightarrow 0$ , as  $i \rightarrow \infty$ , for every compact subset  $K$ , or  $\nu(W) > 0$  for the subspace  $W$  consisting of all  $w$  in  $V$  such that  $\{gw \mid g \in S(\mu)\}$  is bounded.

The answer turns out to be in the negative in general, as we shall see below. However, it was shown in [4] that such a statement holds under certain “mild” additional conditions. We now describe the results from [4]. The situation is somewhat simpler when the measure  $\mu$  has its support on the subgroup  $\text{SL}(V)$  consisting of  $g$  in  $\text{GL}(V)$  such that  $\det g = 1$ , and it will be convenient to describe certain results in this special case first.

**Theorem 1** *Let  $\mu$  be a probability measure on  $\text{SL}(V)$ . Suppose that  $G(\mu)$  is non-compact. Let  $\nu$  be a probability measure on  $V$  such that  $\nu(W) = 0$  for every proper subspace  $W$  which is invariant under a subgroup of finite index in  $G(\mu)$ . Then  $\mu^i * \nu(K) \rightarrow 0$  for all compact subsets  $K$  of  $V$ .*

In particular if there is no proper nonzero subspace which is invariant under the action of  $G(\mu)$  (in this case the  $G(\mu)$ -action is said to be strongly irreducible), then this answers our question in the affirmative. This special case has an analogue in the stronger version of the question as in the work of Furstenberg mentioned above, under the moment condition  $\int \log \|g\| d\mu(g) < \infty$ ; the original result of Furstenberg involves a stronger moment condition, with  $\|g\|$  in the place of  $\log \|g\|$ , but Guivarc’h and Raugi uphold the statement under the weaker moment condition. I may mention, without going into the technical details that the results of Furstenberg (and the later papers along the stream) in fact show, under the respective conditions, that not only almost all sample paths go off to infinity, but in fact their norms (in  $\mathbb{R}^d$ ) grow exponentially as  $i \rightarrow \infty$ .

We note that for the special case when  $\nu$  is a point measure  $\delta_v$ ,  $v \in V$ , the theorem says that if  $v$  is not contained in a subspace  $W$  which is invariant under a subgroup of finite index in  $G(\mu)$  then  $\mu^i * \delta_v(K) \rightarrow 0$  for all compact subsets  $K$  of  $V$ . A subspace  $W$  of  $V$  is invariant under a subgroup of finite index in  $G(\mu)$  if and only if there exist finitely many subspaces  $W_1 = W, W_2, \dots, W_k$  such that  $\cup_{j=1}^k W_j$  is  $G(\mu)$  invariant. Given a  $v \in V$  there exists a unique smallest  $G(\mu)$ -invariant subset  $E$ , containing  $v$ , which is a union of finitely many subspaces, and the theorem tells us that if  $E = V$  then  $\mu^i * \delta_v(K) \rightarrow 0$  for all compact subsets  $K$  of  $V$ . Now consider the complementary case, that  $E$  is a proper subset, so a finite union of proper subspaces. Suppose that in fact  $E$  is just one proper subspace, say  $W$ , (this is possible depending on  $G(\mu)$  and  $v$ ). Then one may consider proceeding by induction on dimension. There is however one hitch: when we restrict to the subspace  $W$ , which by the assumption is  $G(\mu)$  invariant, and view  $\mu$  as a measure on  $GL(W)$ , the determinants of the transformations involved, namely of the restrictions to  $W$ , may not be 1, so the lower dimensional case of the theorem would not suffice in proceeding by induction. This difficulty can be gotten over if we put a further condition on  $\mu$  that we now discuss.

We recall that  $T \in GL(V)$  is said to be *unipotent* if  $(T - I)^r = 0$  for some  $r$  (and hence for  $r = d$ , the dimension of  $V$ ), or equivalently if 1 is the only eigenvalue of  $T$ . We now introduce the following.

**Definition** A subgroup  $H$  of  $GL(V)$  is said to be of *type  $\mathcal{S}$*  if there exists a closed subgroup  $\tilde{H}$  of  $GL(V)$  containing  $H$ , such that the following holds:

- i) if  $E$  is a subset of  $V$  which is  $H$ -invariant and can be expressed as a finite union of subspaces of  $V$ , then  $E$  is also  $\tilde{H}$ -invariant, and
- ii) if  $S$  is the subset of  $\tilde{H}$  consisting of all the elements  $x$  such that  $x$  is either unipotent or contained in a compact connected subgroup of  $\tilde{H}$ , then the subgroup generated by  $S$  is dense in  $\tilde{H}$ .

Though the condition may look rather artificial it turns out that there are natural situations in which it holds. For a unipotent linear transformation  $T$ , if a subspace  $W$  is  $T^m$ -invariant for some  $m \geq 1$  then it is also  $T$ -invariant. Therefore if a finite union of subspaces is invariant under a unipotent transformation  $T$  then each of the maximal subspaces in the union is invariant under  $T$ . Similarly, if  $C$  is a compact connected subgroup of  $GL(V)$  then whenever a finite union of subspaces is  $C$ -invariant then the maximal subspaces in the union are  $C$ -invariant. If  $H$  is a subgroup which is generated by the unipotent elements and compact connected

subgroups contained in it, then it is of type  $\mathcal{S}$ . However, a subgroup  $H$  can be of type  $\mathcal{S}$  even without this being satisfied, as we allow a larger subgroup  $\tilde{H}$  to be considered.

In terms of the theory of algebraic groups, there is neat sufficient condition for a subgroup to be of type  $\mathcal{S}$ , which we shall briefly recall, without going into the technical detail. If  $H$  is a subgroup of  $\mathrm{GL}(V)$  such that the Zariski closure of  $H$  in  $\mathrm{GL}(V)$  is a Zariski connected algebraic group in which the maximal split tori are contained in semisimple subgroups, then  $H$  is of type  $\mathcal{S}$ .

The type  $\mathcal{S}$  condition enables us to proceed by induction as proposed above, and thus we are able to deduce from Theorem 1 the following, which gives a complete answer to our question, for a class of measures.

**Corollary** *Let  $\mu$  be a probability measure on  $\mathrm{SL}(V)$ . Suppose that  $G(\mu)$  is of type  $\mathcal{S}$ . Let  $B$  be the subspace of  $V$  consisting of all the elements  $v$  such that  $\{gv \mid g \in S(\mu)\}$ , is bounded in  $V$ . Then for a probability measure  $\nu$  on  $V$ ,  $\mu^i * \nu(K) \rightarrow 0$  for all compact subsets  $K$  of  $V$  if and only if  $\nu(B) = 0$ .*

Though in the above discussion we restricted to measures supported on  $\mathrm{SL}(V)$  for simplicity, the study in [4] extends to all measures on  $\mathrm{GL}(V)$ . Moreover, we consider not only sequences of measures of the form  $\mu^i * \nu$  as above, but more generally those of the form  $\mu_i * \nu$ , where  $\mu_i$  is any sequence of measures on  $\mathrm{GL}(V)$  commuting with a (fixed) probability measure  $\mu$  on  $\mathrm{GL}(V)$ , and address the question whether  $\mu_i * \nu(K) \rightarrow 0$  for all compact subsets  $K$  of  $V$ . In this respect we assume also that  $\mu_i(C) \rightarrow 0$  for all compact subsets of  $\mathrm{End}(V)$ , the vector space of all endomorphisms of  $V$ . The latter can be seen to be a necessary condition for the desired conclusion, that  $\mu_i * \nu(K) \rightarrow 0$  for all compact subsets  $K$  of  $V$ , to hold. The following general result is proved in [4].

**Theorem 2** *Let  $\{\mu_i\}$  be a sequence of probability measures on  $\mathrm{GL}(V)$  such that  $\mu_i(C) \rightarrow 0$  for all compact subsets  $C$  of  $\mathrm{End}(V)$ . Suppose that there exists a measure  $\mu$  on  $\mathrm{GL}(V)$  such that  $\mu_i * \mu = \mu * \mu_i$  for all  $i$ . Let  $\nu$  be any probability measure on  $V$  such that  $\nu(W) = 0$  for any proper subspace  $W$  which is invariant under the action of a subgroup of finite index in  $G(\mu)$ . Then  $\mu_i * \nu(K) \rightarrow 0$ , as  $i \rightarrow \infty$ , for any compact subset  $K$  of  $V$ .*

Theorem 1 is deduced from Theorem 2 by choosing  $\mu_i = \mu^i$ , for all  $i$ , where  $\mu$  is the given measure as in the hypothesis. To apply the theorem with that choice we need to know that  $\mu^i(C) \rightarrow 0$  for all compact subsets  $C$  of  $\mathrm{End}(V)$ . Since

$\mathrm{SL}(V)$  is a closed subset of  $\mathrm{End}(V)$  and all  $\mu^i$  are supported on  $\mathrm{SL}(V)$  it suffices to know that  $\mu^i(C) \rightarrow 0$  for all compact subsets  $C$  of  $\mathrm{SL}(V)$ . But this is true for any probability measure on  $\mathrm{SL}(V)$  for which  $G(\mu)$  is noncompact. It may be worthwhile to recall here the following result from which this readily follows; it may be noted that Theorem 3 is for all probability measures on  $\mathrm{GL}(V)$ , though the conclusion as required above can be deduced only for measures supported on  $\mathrm{SL}(V)$ ; this is because  $\mathrm{SL}(V)$  is a closed subset of  $\mathrm{End}(V)$  while  $\mathrm{GL}(V)$  is not.

**Theorem 3** (see [1], Theorem 3.2) *Let  $\mu$  be a probability measure on  $\mathrm{GL}(V)$  such that  $G(\mu)$  is noncompact. Let  $K$  be any compact subset of  $\mathrm{GL}(V)$ . Then at least one of the following holds.*

i)  $\sup_{g \in \mathrm{GL}(V)} \mu^i(Kg) \rightarrow 0$  as  $i \rightarrow \infty$ ;

ii)  $G(\mu)$  has an open normal subgroup  $H$  such that  $G(\mu)/H$  is infinite and  $\mu(gH) = 1$  for some  $g \in G(\mu)$ , and, consequently,  $\mu^i(K) = 0$  for all large  $i$ .

Given a probability measure  $\mu$  on a locally compact group  $G$  the map  $K \mapsto \sup_{g \in G} \mu(Kg)$  is called the *concentration function* of  $\mu$ ; notice that if a measure is highly “concentrated” in the intuitive sense then the values of the function would be close (or equal) to 1, while if the measure is thinly spread out on the group then the values would be close to 0. There has been a detailed study of this concept of classical origin, in the context of measures on locally compact groups. The reader is referred to [1] and [2] for further details and references on the topic.

We have seen above that if we could get the conclusion as in Theorem 1 under the weaker hypothesis that  $\nu(W) = 0$  for all proper  $G(\mu)$ -invariant subspaces, in the place of  $\nu(W) = 0$  for all proper subspaces invariant under a subgroup of finite index, it would facilitate obtaining a complete version as in the Corollary to Theorem 1. One may in this context ask if the conclusion could be true under the weaker hypothesis. That however is not the case, as may be seen from the following example.

**Example** Let  $V = \mathbb{R}^2$  and  $\{e_1, e_2\}$  be the standard basis of  $V$ . Let  $\lambda > 1$  be fixed. Let  $g_1, g_2 \in \mathrm{GL}(V)$  be the elements defined by  $g_1(x_1e_1 + x_2e_2) = \lambda x_1e_1 + \lambda^{-1}x_2e_2$ , and  $g_2(x_1e_1 + x_2e_2) = \lambda x_2e_1 + \lambda^{-1}x_1e_2$ , for all  $x_1, x_2 \in \mathbb{R}$ . Let  $\mu$  be the probability measure on  $\mathrm{GL}(V)$  defined by  $\mu(\{g_1\}) = \mu(\{g_2\}) = \frac{1}{2}$ , and let  $\nu = \delta_{e_1}$ . Clearly there is no proper nonzero subspace invariant under both  $g_1$  and  $g_2$ , both of which belong to  $G(\mu)$ . However it turns out that for all compact subset  $K$  of  $V$ ,  $\mu^i * \nu(K) \rightarrow \frac{1}{2}$ ; this can be proved by comparing the process with the classical simple random walk

on the line. Thus we see that the conclusion as in Theorem 1 does not hold under the weaker hypothesis that  $\nu(W) = 0$  for every  $G(\mu)$ -invariant proper subspace. Observe that the condition as in that theorem is not satisfied in this case, since the union of the one-dimensional subspaces spanned by  $e_1$  and  $e_2$  is indeed invariant under  $G(\mu)$ , and hence each of the subspaces is invariant under a subgroup of index 2 in  $G(\mu)$ .

We conclude this description of the results with the following remark. Unlike for measures supported on  $\text{SL}(V)$ , in general conditions on  $G(\mu)$  do not determine whether  $\mu^i(C) \rightarrow 0$  for all compact subsets  $C$  of  $\text{End}(V)$ ; namely for different measures  $\mu_1$  and  $\mu_2$  with  $G(\mu_1) = G(\mu_2)$ , the behaviour could be different. Some of the literature in the area provides conditions under which this does hold (see [6]), and these may be used in applying Theorem 2 for measures  $\mu$  not supported on  $\text{SL}(V)$ . We shall however not go into the details of this here.

### Sketch of proof of Theorem 2

Let  $\{\mu_i\}$ ,  $\mu$  and  $\nu$  be as in the hypothesis. Consider the sequence of measures  $\{\mu_i * \nu\}$ . We proceed along the following steps. (The description below is intended only to give a flavour of what is involved; the reader is referred to [4] for detailed arguments.)

i) We realise the sequence  $\{\mu_i * \nu\}$  as a sequence of probability measures on the one-point compactification  $X = V \cup \{\infty\}$ , the point  $\infty$  being assigned mass 0 for each of the measures. The space of probability measures on  $X$  is a compact second countable space, under the weak\* topology (with the space of measures on  $X$  viewed as the dual of the space of continuous functions). We note that the assertion as in the theorem would follow if  $\{\mu_i\}$  converges to the point measure  $\delta_\infty$ , as a sequence of measures on  $X$ . Suppose this is not the case. Then there exists a subsequence of  $\{\mu_i * \nu\}$  converging to a measure, say  $\lambda$ , on  $X$  such that  $\lambda(V) > 0$ . Since we are only interested in showing that the assumption as above leads to a contradiction, by passing to a subsequence we may assume that  $\mu_i * \nu$  converges to the measure  $\lambda$ , such that  $\lambda(V) > 0$ .

Now, using the assumption that  $\mu_i(C) \rightarrow 0$  for all compact subsets  $C$  of  $\text{End}(V)$  we conclude that there exist a compact subset  $K$  of  $V$ ,  $a > 0$ , and a sequence  $\{g_i\}$  in  $\text{GL}(V)$ , such that  $\{g_i\}$  is divergent as a sequence in  $\text{End}(V)$ , and  $g_i\nu(K) \geq a$  for all  $i$ .

ii) We next use some elementary features of behaviour of measures on vector spaces under application of sequences of linear transformations (see [2]) and con-

clude that there exists a proper subspace  $W$  of  $V$  such that  $\nu(W) > 0$ . (As yet we do not have anything about the proper subspace  $W$  being invariant under any transformation.)

iii) Since  $\mu_i * \mu = \mu * \mu_i$ , for all  $i$ , for any  $j$  we have  $\mu_i * \mu^j * \nu = \mu^j * \mu_i * \nu \rightarrow \mu^j * \lambda$ . Hence we get that for every  $j$  there exists a proper subspace  $W_j$  such that  $\mu^j * \nu(W_j) > 0$ . The key step in the proof is to show that  $W_j$  satisfying the above condition can be chosen such that they form an eventually periodic sequence, and  $(\text{supp } \mu)W_j = W_{j+1}$  for all large  $j$ .

iv) Let  $\tilde{\mu}$  be the measure defined by  $\tilde{\mu}(E) = \mu(E^{-1})$  for all Borel subsets  $E$  of  $\text{GL}(V)$ , and  $N(\mu)$  be the closed subgroup generated by  $\text{supp}(\tilde{\mu} * \mu) \cup \text{supp}(\mu * \tilde{\mu})$  (supp stands for “support of”). The conclusion in (iii) then implies that there exists a finite union of proper subspaces, with positive  $\nu$  measure, and invariant under  $N(\mu)$ . If  $N(\mu) = G(\mu)$  (which holds for instance if  $\mu$  is symmetric or, more generally if  $\text{supp } \mu$  contains the identity element) then this completes the proof.

v) In general  $N(\mu)$  is a normal subgroup such that  $G(\mu)/N(\mu)$  is cyclic, and  $\text{supp } \mu$  is contained in a single coset of  $N(\mu)$ . The proof in the general case is via an averaging argument, based on the following

**Proposition** *Let  $T : V \rightarrow V$  be a linear transformation. Then for every  $\delta > 0$  there exists  $k$  such that the following holds: if  $W$  is a proper subspace of  $V$  and there exists a proper subspace  $W'$  of  $V$  such that*

$$|\{1 \leq j \leq k \mid T^j W \subset W'\}| > \delta k,$$

(where  $|\cdot|$  denotes the cardinality of the set), then there exists  $m$  such that  $W$  is contained in a proper  $T^m$ -invariant subspace of  $V$ .

There ought to be a simpler proof for this, than we know at present. The proof in [4], for a general  $T \in \text{GL}(V)$ , is based on a deep theorem of Szemerédi in the following form (Gowers version [5]).

**Theorem** *For any  $\delta > 0$  and natural number  $m$  there exists a natural number  $k$  such that every subset of  $\{1, 2, \dots, k\}$  of cardinality exceeding  $\delta k$  contains an arithmetic progression of length  $m$ .*

The concluding part of the proof of Theorem 2 can be further simplified if the following stronger statement, which we expect to be true, actually holds.

**Conjecture:** *Let  $T \in \text{GL}(V)$  and  $v \in V$ . Then there exists  $n$  such that for every*

subspace  $W$  for which  $\{i \mid T^i v \in W\}$  has more than  $n$  elements there exist natural numbers  $k$  and  $m$  such that  $T^{k+im}v$  is contained in  $W$  for all  $i$ .

It is not difficult to see that this statement holds when  $T$  is a unipotent linear transformation. It is also possible to prove it, using elementary calculus, for a somewhat larger class of  $T$ , those for which all eigenvalues are real; see [3] for details. The general case however remains open.

## References

- [1] S.G. Dani, Scattering of products of random matrices, in: *Analysis, Geometry and Probability* (ed: R. Bhatia), pp. 33–53, Hindustan Book Agency, New Delhi, 1996.
- [2] S.G. Dani, Asymptotic behaviour of measures under automorphisms, in *Probability Measures on Groups: Recent Directions and Trends* (Proceedings of CIMPA-TIFR School, Mumbai 2002), Narosa Publishing House, New Delhi, (international distribution by the American Mathematical Society), to appear.
- [3] S.G. Dani, Dynamical properties of linear and projective transformations and their applications, *Indian J. Pure Appl. Math.* 35 (2004), 1365-1394.
- [4] S.G. Dani and Riddhi Shah, Asymptotic behaviour under iterated random linear transformations, *Math. Res. Lett.* 11 (2004), 467-480.
- [5] W.T. Gowers, A new proof of Szemerédi's theorem, *Geom. Funct. Anal.* 11 (2001), 465–588.
- [6] Y. Guivarc'h, Limit theorems for random walks and products of random matrices, in *Probability Measures on Groups: Recent Directions and Trends* (Proceedings of CIMPA-TIFR School, Mumbai 2002), Narosa Publishing House, New Delhi, (international distribution by the American Mathematical Society), to appear.

School of Mathematics  
Tata Institute of Fundamental Research  
Homi Bhabha Road, Colaba  
Mumbai 400 005  
India