Bernoulli Shifts and Induced Automorphisms

LAIF SWANSON

Department of Mathematics, Texas A&M University, College Station, Texas 77843

We show that Bernoulli shifts induce, on a dense class of sets, weakly mixing automorphisms which are not mixing. We also show that the non-Bernoulli K-automorphisms described by Ornstein and Shields induce Bernoulli shifts on a dense class of sets. We are therefore able to see that Bernoulli flows may be written as flows under functions over a wide class of automorphisms.

If $T$ is an automorphism of a measure space $X$ and $A$ is a set of positive measure, the induced automorphism $T_A$ of the normalized probability space $A$ is defined by

$$T_A(x) = T^k x,$$

where $k(x) = \min\{n > 0: T^n x \in A\}$.

(It is not hard to see that $T_A$ is, in fact, an automorphism of $A$ and that $T_A$ is ergodic if $T$ is.)

We will say that a family $\mathcal{A}$ of measurable sets is dense in the measure space $(X, \mu)$ if for each $\epsilon > 0$ and measurable set $B$ there is an $A \in \mathcal{A}$ such that $\mu(A \Delta B) < \epsilon$.

Friedman and Ornstein [4] show that ergodic automorphisms induce mixing automorphisms on a dense class of sets; Friedman shows that Bernoulli shifts induce Bernoulli shifts on a dense class of sets. It is also true, and easy to see using the Rohlin–Kakutani theorem, that the family of sets $A$ for which $T_A$ is not weakly mixing is dense in $(X, \mu)$.

Two ergodic automorphisms $S$ and $T$ are said to be weakly equivalent, or Kakutani equivalent, if they induce isomorphic automorphisms, or equivalently, if they are induced by a common ergodic automorphism. (These concepts are introduced and explored by Kakutani [5]). If $S$ and $T$ are ergodic, the special flows (or flows written under functions) which can be written over $S$ can also be written over $T$ if and only if $S$ and $T$ are weakly equivalent [5, p. 640].

Ornstein has shown that the Bernoulli flow can be written as a flow under a function over a Bernoulli shift. Saleski has shown that all Bernoulli shifts are weakly equivalent. Thus our proofs that Bernoulli shifts induce weakly mixing automorphisms which are not mixing and that the Ornstein–Shields $K$-automorphisms induce Bernoulli shifts give us a wide range of automorphisms over which we can write the Bernoulli flow.

Some unanswered questions related to those answered in this paper are:
Are all automorphisms of positive entropy weakly equivalent? Does every automorphism induce, on a dense class of sets, weakly mixing automorphisms which are not mixing? Does every automorphism of positive entropy induce Bernoulli shifts? On a dense class of sets? If $S$ and $T$ are weakly equivalent and $h(S) > h(T)$, does $T$ induce $S$? Strongest of all, is it true if $h(S) > h(T)$, $T$ induces $S$?

1. Bernoulli Shifts Induce Weak Mixing Automorphisms Which Are Not Mixing on a Dense Class of Sets

In order to show that every Bernoulli shift of finite entropy induces, on a dense class of sets, weakly mixing automorphisms which are not mixing, we will first construct a family $\mathcal{B}$ of automorphisms and show that $\mathcal{B}$ is (up to isomorphism) exactly the family of Bernoulli shifts. We will then show that each $T \in \mathcal{B}$ induces, on sets of arbitrarily large probability, weakly mixing automorphisms which are not mixing.

Each member $T$ of $\mathcal{B}$ will be the common extension of the tower transformations for a sequence of towers $\mathcal{F}(n)$, and $X$ will be $\bigcup_n \mathcal{F}^*(n)$. The columns of $\mathcal{F}(n)$, whose intervals will generators for $T$, will all have the same height $h(n)$. Associated with each $T$ will be integers $k$ and $r$, and a probability vector $(p_1, \ldots, p_k)$, and increasing sequences of integers $f$ and $s$ whose properties are listed below. (The sequences $f$ and $s$ make $T$ a Bernoulli shift; the number $r$ and probability vector $(p_1, \ldots, p_k)$ can be adjusted to change the entropy of $T$.)

The space $X$ will be partitioned by $P$,

$$P = (P_{f_1}, \ldots, P_{f_k}, P_s, P_0)$$

which will generate for $T$, and $P$ will also denote the partition associated with each tower $\mathcal{F}(n)$.

These automorphisms are very much like the Bernoulli shift described by Ornstein [6].

Some Definitions for the Construction

A column $C$ is a finite family of intervals $(A_1, \ldots, A_n)$ of the same length (called the width of $C$), together with a partition $P$ of $\bigcup_i A_i$, each of whose atoms $P_i$ is a union of sets $A_i$, and together with a transformation $T = T(C)$ which carries $A_i$ linearly onto $A_{i+1}$ and is not defined on $A_n$. $C^*$ means $\bigcup_i A_i$ (sometimes called the points of $C$). We call $n$ the height of $C$ or $h(C)$. The base of $C$ is defined to be $A_1$; its roof is $A_n$. By a subcolumn of $C$ we mean a column $(B_1, \ldots, T^r B_1)$ where $B_1$ is a subinterval of $A_1$.

A tower $\mathcal{F}$ is a finite set $(C_1, \ldots, C_n)$ of columns with the property that $C_i^*$ is disjoint from $C_j^*$ if $i \neq j$. The symbol $\mathcal{F}^*$ is used to denote $\bigcup_i (C_i^*)$. (The
columns need not have the same height or width.) The *width of* \( \mathcal{T} \), \( w(\mathcal{T}) \), is defined as \( \sum_{C \in \mathcal{T}} w(C) \). We use \( \text{Base}(\mathcal{T}) \) to mean \( \bigcup_{C \in \mathcal{T}} \text{Base}(C) \). Towers \( \mathcal{T} = (C_1, ..., C_n) \) and \( \mathcal{T}' = (C'_1, ..., C'_n) \) are called *isomorphic* if \( (C_i)w(\mathcal{T}) = w(C'_i)|w(\mathcal{T}') \) and \( h(C_i) = h(C'_i) \) for each \( i \).

If \( \mathcal{T} \) and \( \mathcal{T}' \) are towers of the same width and \( \mathcal{T}^* \) is disjoint from \( \mathcal{T}'^* \), a new tower \( \mathcal{T} \ast \mathcal{T}' \) can be defined by a process called *independent cutting and stacking* so that \( (\mathcal{T} \ast \mathcal{T}')^* = \mathcal{T}^* \cup \mathcal{T}'^* \) and the tower transformation \( T(\mathcal{T} \ast \mathcal{T}') \) extends both \( T(\mathcal{T}) \) and \( T(\mathcal{T}') \). We will describe the process:

Let \( \mathcal{T} = (C_1, ..., C_n) \), and \( \mathcal{T}' = (C'_1, ..., C'_m) \).

Divide each column \( C_i \) of \( \mathcal{T} \) into \( m \) subcolumns \( C_i(j) \), \( 1 \leq j \leq m \), so that

\[
 w(C_i(j)) = \frac{w(C_i)w(C'_i)}{w(\mathcal{T})};
\]
divide each column \( C'_i \) of \( \mathcal{T}' \) into \( n \) subcolumns \( C'_i(i) \), \( 1 \leq i \leq n \), so that

\[
 w(C'_i(i)) = w(C_i(j)),
\]
and let the column \( C(i,j) \) be

\[
 (\text{Base}(C_i(j)), ..., \text{Roof}(C_i(j)), \text{Base}(C'_i(i)), ..., \text{Roof}(C'_i(i))).
\]

The tower \( \mathcal{T} \ast \mathcal{T}' \) is all of the columns \( C(i,j) \), \( 1 \leq i \leq n, 1 \leq j \leq m \).

**The Construction**

We will assume that we have integers \( k \) and \( r \), the probability vector \( (p_1, ..., p_k) \), and increasing sequences \( f \) and \( s \). We will use these inductively to construct a sequence of towers which will define \( T \).

Let \( A_1, ..., A_r \) be intervals of the same length, and let \( \mathcal{T}_1(1) = (C_1) \), where \( C_1 = (A_1, ..., A_t) \). The atom \( T_0 \) of \( P \) is exactly \( \mathcal{T}_1(1) \).

We will define \( \mathcal{T}(n) \), assuming that \( \mathcal{T}(n - 1) \) has already been defined:

First, divide \( \mathcal{T}(n - 1) \) into \( s(n) \) isomorphic towers of equal width, \( \mathcal{T}(n, 1), ..., \mathcal{T}(n, s) \) and let

\[
 \mathcal{T}(n) = \mathcal{T}(n, 1) \ast \cdots \ast \mathcal{T}(n, s).
\]

Then divide each column \( C \) of \( \mathcal{T}(n) \) into \( f(n) - 1 \) subcolumns \( C(j) \) of equal width, \( 1 \leq j < f(n) \), and then divide each of these columns into \( k^j \) subcolumns \( C(j) \) of equal width, \( 1 \leq j < f(n) \), and then divide each of these columns into \( k^j \) subcolumns \( C(j; r_1, ..., r_i) \) \( 1 \leq r_i \leq k \) where

\[
 w(C(j; r_1, ..., r_i)) = w(C(j))(p_{r_1} \cdots p_{r_i}).
\]
Precede \( C(j; r_1, \ldots, r_j) \) by a column \( C'(j; r_1, \ldots, r_j) \) of the same width and of height \( j \) whose \( i \)th interval is in \( f_{r_i} \), and follow it by a column of height \( f(n) - j \) entirely in \( e \). The tower \( \mathcal{T}(n) \) is composed of all these columns.

Note that all the columns in \( \mathcal{T}(n) \) have the same height, which we will call \( h(n) \).

Thus a typical column in \( \mathcal{T}(n) \) looks like:

\[ h(n) = s(n) h(n - 1) + f(n) \]

\[ s(n) \]
\[ n - 1 \]

\[ j \]
\[ i \]

\[ p_i \]

\[ f_i \]
BERNOULLI SHIFTS AND INDUCED AUTOMORPHISMS

Our space $X$ will be $\bigcup_n \mathcal{F}^*(n)$, and our automorphism $T$ will be the common extension of $T(\mathcal{F}(n))$.

Because $s(n)$ is an increasing sequence and

$$w(\mathcal{F}(n)) = \frac{w(\mathcal{F}(n-1))}{s(n)},$$

we see that

$$\lim_{n \to \infty} w(\mathcal{F}(n)) = 0$$

and so $T$ is defined almost everywhere.

We will require the sequences $f$ and $s$ to satisfy

$$\lim_{n \to \infty} \frac{h(n-1)}{f(n)} = 0$$

(which we will need in order that the automorphisms be Bernoulli), and

$$\sum_{n \geq 1} \frac{f(n)}{h(h)} < \infty$$

(which will guarantee that $\bigcup_n \mathcal{F}^*(n)$ will have finite measure and so we can use normalized interval length for measure.)

$T$ is Bernoulli

**Proposition.** Each $T \in \mathcal{B}$ is a Bernoulli shift.

**Proof.** Fix $T$ in $\mathcal{B}$. We will show that the process $(T, P)$ is very weakly Bernoulli.

First note that the independence of the stacking causes $T$ to be ergodic. (This easy fact is left to the reader.)

We must show that for every $\epsilon$ there is a $q$ such that for any $m$ there is a family

$$\mathcal{A} \subset \bigvee_{t=-m}^{0} T^t P$$

with $\mu(\cup \mathcal{A}) > 1 - \epsilon$, such that if $A$ and $B \in \mathcal{A}$,

$$\overline{d}(\{T^i P|A\}_{i=-1}^{q}, \{T^i P|B\}_{i=-1}^{q}) < \epsilon.$$

(See Ornstein [6] for definitions.)

Following Ornstein, we will use "block" to mean that part of $P$-name corresponding to a column of a tower, e.g., an $n$-block for a column of $\mathcal{F}(n)$, and "string" to mean any sequential part of a $P$-name.

Choose $n$ so large that the measure of the $n$-blocks (i.e., $m(\mathcal{F}^*(n))$) is at least $1 - \epsilon/4$. Choose $t$ so large that $(4/5)^t < \epsilon/4$. Choose $q$ so large that for all but a family of $q$-strings of measure less than $\epsilon^3/4$. $q$-strings are at least $(1 - \epsilon/2)$
made up of whole \((n+t)\)-blocks. (That is, use the ergodic theorem to choose a \(q\) so large that for all but a family of \(x\)'s of measure less than \(\varepsilon^q/4\),
\[\text{card}\{i: h(F(n+t)) \leq i \leq q - h(F(n+t)) \text{ and } T^i x \in F^*(n+t)\} \geq (1 - \varepsilon/2)q.\]
We will show that for most atoms \(A\) and \(B\) of \(V_0^\top T^P\),
\[d(\{T^P A\}^q_1, \{T^P B\}^q_1) < \varepsilon.\]
We will exclude those atoms \(A\) for which
\[\mu(\{x \in A: (x_{-q}, ..., x_{-1}) \text{ is at least } (1 - \varepsilon/2) \text{ made up of whole } (n+t)\text{-blocks}\}) \leq (1 - \varepsilon^q/4) \mu(A),\]
which is a family of measure at most \(\varepsilon\) because of our choice of \(q\).
To show that if \(A\) and \(B\) are atoms of \(V_0^\top T^P\),
\[d(\{T^P A\}^q_1, \{T^P B\}^q_1) < \varepsilon,\]
we can of course partition \(A\) and \(B\) into disjoint subsets
\[A = \bigcup A_g, \quad B = \bigcup B_i,\]
and show that for all but a set of \(A_g\)'s of measure less than \(\varepsilon/2\) and all but a set of \(B_i\)'s of measure less than \(\varepsilon/2\),
\[d(\{T^P A_g\}^q_1, \{T^P B_i\}^q_1) < \varepsilon/2.\]
To do that, we establish the following equivalence relation on \(A\) (and a similar one for \(B\)).
For \(x \in A\), define
\[I(x) = \{i: 1 \leq i \leq q - h(F(n+1)) \text{ and } T^{-i} x \in C\} \text{ where } C = \text{Base}(F(n+t))\]
and say that \(x \sim y\) if \(I(x) = I(y)\). (In Ornstein's terminology, we are partitioning according to the position of \((n+t)\)-blocks.)
We will condition only on those equivalence classes \(A_g\) (and \(B_i\)) for which
\[h(F(n+t)) \text{ card}(I(x)) \geq (1 - \varepsilon/2)q\]
for \(x \in A_g(B_i)\), i.e., those made up almost entirely of \((n+t)\)-blocks. But these make up all but a proportion \(\varepsilon/2\) of the measure of \(A\) (or \(B\)).
So now we must show that if \(A_g\) is the set of all points in a certain atom of \(V_0^\top T^P\) with \((n+t)\)-blocks in certain places, and \(B_i\) is the set of all points in another atom of \(V_0^\top\) with \((n+t)\)-blocks in other certain places, then
\[d(\{T^P A_g\}^q_1, \{T^P B_i\}^q_1) < \varepsilon.\]
We could do this by defining an isomorphism $\phi$ from $A_g$ to $B_j$ and showing that

$$\frac{1}{q} \sum_{i=1}^{\delta} |\phi T^n P|B_j - T^n P|B_j| < \epsilon.$$ 

We will instead define a (finite) sequence of "isomorphisms" $\phi_i$ which take successively finer partitions of $A_g$ to successively finer partitions of $B_j$ in a compatible fashion. Any extension $\phi$ of the "isomorphism" $\phi_i$ to a real isomorphism from $A_g$ to $B_j$ will have the desired property

$$\frac{1}{q} \sum_{i=1}^{\delta} |\phi T^n P|B_j - T^n P|B_j| < \epsilon.$$ 

Since $I(\varepsilon)$ is common to all members of $A_g$, we will call it $I(A_g)$; similarly we will use $I(B_j)$ to denote the set of beginning points for $(n + t)$-blocks in $B_j$.

Since

$$h(n + t) \text{ card}(I(A_g)) \geq (1 - \varepsilon/2)q$$

and

$$h(n + t) \text{ card}(I(B_j)) \geq (1 - \varepsilon/2)q$$

we can define a correspondence $\psi_1: I(A_g) \rightarrow I(B_j)$ so that

(i) $\psi_1$ is defined on most members of $I(A_g)$,
(ii) $\psi_1$ is 1-1 and is onto most of $I(B_j)$,
(iii) $|i - \psi_1(i)| \leq \frac{3}{8}h(n + t)$ whenever $\psi_1$ is defined.

(In Ornstein's terminology, we can make a correspondence between most $(n + t)$-blocks so that they overlap at least $\frac{1}{4}$, because most of the $P$-$q$-name is made of $(n + t)$-blocks.)

For each $i$ in the domain of $\psi_1$, partition $A_g$ according to the number of $f$'s at the beginning of that $(n + t)$-block, i.e., according to

$$\min\{\{p > i: T^{-p}x \in \text{Base}(\mathcal{F}(n + t - 1))\})$$

and do the same for $B_j$ using $\psi_1(i)$. Since $s(n + t) h(n + t - 1)$ is almost as big as $h(n + t)$, and $f(n + t)$ is very large compared to $h(n + t - 1)$ (by the properties we listed earlier), we can make a correspondence between these classes which will make at least one-fifth of the $(n + t - 1)$-blocks line up exactly. That is, we can define $\phi_1$ so that for each $x \in A_g$

$$\text{card}\{i: 1 \leq i \leq q - h(\mathcal{F}(n + t - 1)), \quad T^{-i}x \in \text{Base}(\mathcal{F}(n + t - 1)) \quad \text{and} \quad \phi_1 T^{-i}x \in \text{Base}(\mathcal{F}(n + t - 1))\} \geq \frac{1}{5}r,$$

where $r = \text{card}\{i: 1 \leq i \leq q - h(\mathcal{F}(n + t - 1)) \quad \text{and} \quad T^{-i}x \in \text{Base}(\mathcal{F}(n + t - 1))\}$, i.e., the number of whole $(n + t - 1)$ blocks in the $q$-name of $x$. 
Since which \((n + t - 1)\)-block occurs is independent of everything else, we can in fact make these \((n + t - 1)\)-blocks line up and be the same. So, so far, we have made our correspondence good enough that
\[
\text{card}(\{j: 1 \leq j \leq q \text{ and } T^{-i}x = T^{-i}(\phi_1x)\}) \geq \frac{1}{6} \text{card}(I(A_g)) h(n + t - 1) s(n + t),
\]
where "\(\equiv\)" means "in the same atom of \(P\)."

Now we can repeat our argument for \((n + t - 1)\)-blocks:

We restrict our attention to one atom \(A^i_g\) of the last partition (which was according to the position of \((n + t)\)-blocks and the number of \(f\)'s at the beginning of \((n + t)\)-blocks) and its image under \(\phi_1\). Positions of \((n + t - 1)\)-blocks are fixed in \(A^i_g\) and in \(\phi_1(A^i_g)\).

Make a correspondence \(\psi_2\) between \((n + t - 1)\)-blocks in the name of a point in \(A^i_g\) and those in the name of a point in \(\phi_1(A^i_g)\) so that each \((n + t - 1)\)-block in the name of a point in \(A^i_g\) overlaps the corresponding \((n + r - 1)\)-block in \(\phi_1(A^i_g)\) by at least \(\frac{1}{3}\). Then partition the atoms further by the number of \(f\)'s at the beginning of \((n + t - 1)\)-blocks. Now we can define \(\phi_2\) on the atoms of this newest partition so that at least one-fifth of the \((n + t - 2)\) blocks which did not already line up (by virtue of belonging to lined up \((n + t - 1)\)-blocks under \(\phi_1\)) will line up and be the same. Thus, we have now made our correspondence good enough that
\[
\text{card}(\{j: 1 \leq j \leq q \text{ and } T^{-i}x = T^{-i}(\phi_2x)\}) \geq (1 - (\frac{3}{4})^t) \text{card}(I(A_g)) s(n + t - 1) s(n + t) h(n + t - 2).
\]

We can continue this argument, at each stage lining up another one-fifth of the remaining \((n + t - j)\)-blocks, until, after \(t\) steps, we have lined up
\[
\frac{1}{3} \text{ of the } (n + t)\text{-blocks},
\frac{1}{3} \text{ of the } (n + t - 1)\text{-blocks not contained in the blocks above},
\vdots
\frac{1}{3} \text{ of the } n\text{-blocks not contained in any of the blocks above}.
\]

Thus, all together, at least \((1 - (\frac{3}{4})^t)\) of the \(n\)-blocks line up exactly and are the same. Thus the "isomorphism" \(\phi_1: A_g \rightarrow B_j\) has the property that, for each \(x \in A_g\),
\[
\text{card}(\{j: 1 \leq j \leq q \text{ and } T^{-i}x = T^{-i}(\phi_1x)\}) \geq (1 - (\frac{3}{4})^t)(\text{number of whole } n\text{-blocks contained in whole } (n + t)\text{-blocks in } (x_{-q}, \ldots, x_{-l}) \text{ h} (\mathcal{F}(n)) \geq (1 - \epsilon/4)((1 - 3\epsilon/4)q) \geq (1 - \epsilon)q,
\]
where "\(\equiv\)" again means "in the same atom of \(P\)."
That is, if $f$ is the function
$$f(x) = \frac{1}{2} \sum |X_{P_i} - X_{P_j}|, \quad P_i \in P,$$
where $\phi$ is any isomorphism extending $\phi_i$, then
$$\sum_{i=1}^q f(T^{-i}x) \leq \epsilon.$$ 

Now, by integrating $\sum_{i=1}^q f(T^{-i}x)$ over $A_\sigma$, we obtain
$$\sum_{i=1}^q |T^iP|B_j - \phi T^iP|B_j| \leq \epsilon,$$
and so
$$d\left(\{T^iP|A_j\}_{j=1}^q, \{T^iP|B_j\}_{j=1}^q\right) < \epsilon. \quad \text{Q.E.D.}$$

\textbf{B Contains Automorphisms of Each Positive Entropy}

We will show that every positive number is the entropy of some $T \in B$.

\textbf{Lemma.} Let $T$ be any member of $\mathcal{B}$. Let $P_f = \bigcup P_i$, $P'$ be the partition $(P_f, P, P_0)$, and $\mathcal{F}'$ be the factor algebra $\bigvee_{-\infty}^\infty T^iP'$. Denote by $T'$ the factor $T|_{\mathcal{F}'}$, and by $T_B$ the Bernoulli shift on $k$ letters with probabilities $(p_1, \ldots, p_k)$. Then
$$h(T) = h(T') + h(T_B)m(P_f).$$

\textbf{Proof.} $h(T) = h(T, P) = h(T', P') + h(P|\bigvee_{-\infty}^\infty T^iP \vee \mathcal{F}')$. The atoms $P_e$ and $P_0$ of $P$ are measurable with respect to $\mathcal{F}'$, and so
$$h\left(P|\bigvee_{-\infty}^\infty T^iP \vee \mathcal{F}'\right) = m(P_e)h\left(P|_{P_e}|\bigvee_{-\infty}^\infty T^iP \vee \mathcal{F}'\right),$$
but we have constructed $(P_{i_1}, \ldots, P_{i_2})$ so that $P|_{P_j} \perp \bigvee_{-\infty}^\infty T^iP \vee \mathcal{F}'$, and so
$$h\left(P|_{P_j}|\bigvee_{-\infty}^\infty T^iP \vee \mathcal{F}'\right) = h(P|_{P_j}) = h(T_B).$$
Thus
$$h(T) = h(T, P) = h(T', P') + h\left(P|\bigvee_{-\infty}^\infty T^iP \vee \mathcal{F}'\right)$$
$$= h(T') + m(P_f)h(T_B).$$

\textbf{Note.} We could have seen this instead by writing $T$ as the skew of $T'$ with $T_B$ and the identity.
Note. $T'$ depends on $r$ and on the sequences $f$ and $s$, but not on $k$ or on the probability vector $(p_1, ..., p_k)$.

Corollary. Given $r$ and sequences $f$ and $s$, the integer $k$ and the probability vector $(p_1, ..., p_k)$ can be chosen to make $h(T)$ any number larger than $h(T')$.

Lemma. By suitable choices of $r$ and the sequences $f$ and $s$, $h(T')$ can be made as small as we like.

Proof. $h(T') < h(P')$, because $P'$ is a generator for $T'$, and $h(P')$ gets arbitrarily small as $m(P_0)$ approaches one. But for fixed $f$ and $s$, $m(P_0)$ gets arbitrarily close to one as $r$ gets large (which we can do without affecting the requirements on $f$, $s$, and $h(\mathcal{T}(n))$) and so $h(T')$ can be made arbitrarily small by making $r$ large.

Proposition. Given a real number $h$, there is an automorphism $T \in \mathcal{B}$ such that $h(T) = h$.

Proof. Choose any sequences $f$ and $s$ satisfying our growth requirements. Choose $r$ so that $h(T') \leq h$. Choose $k$ and $(p_1, ..., p_k)$ so that

$$h(T_B) = \frac{h - h(T')}{m(P')}.$$ 

Then

$$h(T) = h(T') + m(P_f) h(T_B) = h(T') + h - h(T') = h.$$ 

Induced Automorphisms

We will show that if $T \in \mathcal{B}$, there are sets $A$ of arbitrarily large measure for which $T_A$ is weakly mixing but not mixing.

We assume that a member $T$ of $\mathcal{B}$ is fixed.

The sets we will look at are the sets

$$A(n) = \mathcal{T}^*(n) \cup \left( \bigcup_m \text{Roof}(\mathcal{T}(m)) \right).$$

Since the sets $\mathcal{T}^*(n)$ are of arbitrarily large measure, $\lim_{n \to \infty} \mu(A(n)) = 1$.

Lemma. For each $n$, $\mathcal{T}^*(n)$ is measurable with respect to $\sqrt[n]{T^n P}$.

Proof. This fact, which can be proved by induction on $n$, is left to the reader. Since the measurability of $\mathcal{T}^*(n)$ implies the measurability of $\text{Roof}(\mathcal{T}(n))$, we can define $A(g) = \mathcal{T}^*(g) \cup \left( \bigcup_n \text{Roof}(\mathcal{T}(m)) \right)$.

Lemma. For any $g$, $T_A(g)$ is the common extension of the following tower transformations:
Let $\mathcal{F}_1(1) = \mathcal{F}(g)$.

Given $\mathcal{F}_1(m)$, to form $\mathcal{F}_1(m + 1)$: Divide $\mathcal{F}_1(m)$ into $s(m + g)$ isomorphic towers $\mathcal{F}_1(m, 1), \ldots, \mathcal{F}_1(m, s)$ and let $\mathcal{F}_1(m + 1) = \mathcal{F}_1(m, 1) \times \cdots \times \mathcal{F}_1(m, s)$. Divide each column $C$ of $\mathcal{F}_1(m + 1)$ into $f(m + g) - 1$ subcolumns $C(j)$, $1 \leq j < f(m + g)$, of equal width; divide $C(j)$ into $j^k$ subcolumns of $C(j: r_1, \ldots, r_j)$, $1 \leq r_i \leq k$, of width $w(C(j)) (p_{r_1}, \ldots, p_{r_j})$, add one interval of $P_*$ to the top of each of these columns; let these be the columns of $\mathcal{F}_1(m + 1)$.

Proof. By induction on the towers: $\mathcal{F}_1^*(m - g) = \mathcal{F}_1^*(m) \cap A(g)$ if $m \geq g$. The $P_*$ added at the end of each step is the roof of $\mathcal{F}(m + g)$; the division into subcolumns separates points with different names outside $A(g)$, i.e., according to the $f$ added to $\mathcal{F}(m + g)$ which is not in $A(g)$.

Shown below is a column in the $n$th tower for $T_{A(g)}$.

- one interval of $e$
- lots of $e$ from $\mathcal{F}(n)$ left out
- (n - 1) columns for $T_{A(g)}$
- all the added $f$ from $\mathcal{F}(n)$ left out

**Proposition.** For each $n$, $T_{A(n)}$ is weakly mixing.

Proof. We will use $h(j)$ to denote height ($\mathcal{F}_1(j)$). We will assume that $n$ is fixed and write $A$ for $A(n)$.

If $T_A$ is not weakly mixing, there is an eigenfunction $g$ with eigenvalue $\alpha$, $|\alpha| = 1$. We can assume that $|g| = 1$ a.e. and, since $T$ is ergodic (causing $T_A$ to be) that $\alpha \neq 1$.

Since the intervals in the columns of the towers generate the Lebesgue sets, for each $\varepsilon$ there is an $m$ so that for all but a family of columns of $\mathcal{F}_1(m)$ of total width no more than $\varepsilon w(\mathcal{F}_1(m))$, each interval $I$ of each column has an associated complex number $\lambda(I)$ of absolute value 1 so that

$$\mu(\{x \in I: |g(x) - \lambda(I)| > \varepsilon\}) < \varepsilon \mu(I).$$

Because most of the points in the base of a column of $\mathcal{F}_1(m)$ are images under
In the tower $\mathcal{T}_i(m+1)$, parts of the base of such a $C$ are carried to other parts of the base by $T_A^{s(m+n)}$, $0 \leq i \leq s(m+n)$. Thus for all but possibly $\epsilon s(m+n)$ of the $i$'s:

$$|\alpha^{s(m+n)}h(m) - 1| < 8\epsilon,$$

and so

$$|\alpha^{s(m+n)}h(m+1) - 1| < 8\epsilon.$$

But $h(m+1) = s(m+n)h(m) + 1$, and so we get

$$|\alpha - 1| = |\alpha^{s(m+n)}h(m) - \alpha^{s(m+n)}h(m)| \\ 
\leq |\alpha^{s(m+n)}h(m+1) - 1| + |\alpha^{s(m+n)}h(m) - 1| \\ 
\leq 16\epsilon.$$

But $\alpha \neq 1$ and so we have reached a contradiction and $T_A$ is weakly mixing.

**Proposition.** $T_A$ is not mixing.

**Proof.** Suppose $T_A$ is mixing. Let $B$ be $\text{Base}({\mathcal{T}_i(n)})$, for any $n$. Since $\mu(B) < \frac{1}{2}$, there is a $p$ so that if $q > p$, $\mu(T_A^qB \cap B) < \frac{1}{2}\mu(B)$. Choose $j$ so that $h(j) > p$ and $j > n$.

$$\mu(T_A^jB \cap B) > \frac{s(j) - 1}{s(j)}\mu(B) \geq \frac{1}{2}\mu(B).$$

But this is a contradiction, and so $T_A$ is not mixing.

**The Main Theorem**

By putting together all of these propositions, we get:

**Theorem.** If $T$ is a Bernoulli shift, there are sets $A$ of arbitrarily large probability for which $T_A$ is weakly mixing but not mixing.
Corollary. Bernoulli shifts induce weakly mixing automorphisms which are not mixing on a dense class of sets.

Proof. Bernoulli shifts induce Bernoulli shifts on a dense class of sets [3], but those Bernoulli shifts in turn induce weakly mixing automorphisms which are not mixing on sets of arbitrarily close measure.

Corollary. Bernoulli flows may be written as flows under functions over automorphisms which are weakly mixing but not mixing.

Proof. Bernoulli flows may be written as special flows over Bernoulli shifts, and we have just seen that Bernoulli shifts are weakly equivalent to automorphisms which are weakly mixing but not mixing. (See Ornstein [6] and Kakutani [5].)

Corollary. Any ergodic automorphism induces, on a dense class of sets, automorphisms which are weakly mixing but not Bernoulli.

Proof. Any ergodic automorphism induces, on a dense class of sets, mixing automorphisms. But if such an induced automorphism is Bernoulli, it induces weakly mixing non-Bernoulli automorphisms on sets of arbitrarily large measure.

Note. The same argument that shows that $T$ is a Bernoulli shift will show that $T_{\mu}$ is also a Bernoulli shift.

2. ORNSTEIN-SHIELDS $K$-AUTOMORPHISMS INDUCE BERNOULLI SHIFTS ON A DENSE CLASS OF SETS

We will show that the $K$-automorphisms (which are not Bernoulli) described by Ornstein and Shields induce Bernoulli shifts on a dense family of sets.

Description of the Automorphisms

Here is a brief description of the Ornstein-Shields construction [7].

We have two increasing sequences of integers, $f(n)$ and $s(n)$, and a sequence $g(n)$ of zeros and ones. We will, for convenience, assume that $f(1) > 2$. Two sequences of gadgets $G(n)$ and $G^*(n)$ are constructed; $G^*(n) \subset G^*(n + 1)$, $\lim_{n \to \infty} w(G(n)) = 0$, and, because of the choice of the sequences of $f$ and $s$, $\lim_{n \to \infty} w(G(n)) h(G(n)) < \infty$. The gadget transformations are consistent. $\bigcup_n G^*(n)$ is a probability space with normalized interval length for measure, and the transformation $T$, the common extension of all the gadget transformations, is shown to be a $K$-automorphism which is not Bernoulli. Each gadget partition $P(n)$ is labeled (consistently) $P = (P_s, P_f, P_0)$.

We first form $G(1)$. It consists of $s(1)$ intervals, all in $P_s$. 
To construct $G(1)$, divide each column $C$ of $G(1)$ into $f(1) - 1$ subcolumns $C(i), 1 \leq i < f(1)$; precede $C(i)$ by $i$ intervals in $P_f$ and follow it with $f(1) - i$ intervals in $P_e$. These are the columns of $G(1)$.

Assuming that $G(n)$ has been constructed, we will construct $G(n + 1)$, and then $G(n + 1)$.

Divide the base of $G(n)$ into $2^{n+1}$ disjoint sets of equal measure so that each is the base of a gadget isomorphic to $G(n)$. Call these gadgets $G_1, ..., G_{2^{n+1}}$.

For each $i, 1 \leq i \leq 1 + 2^{n+1}$, construct a gadget $H_i$, consisting only of sets in $P_s$, by taking $i s(n + 1)$ disjoint sets $S_{i,j}$, each of length $w(G(n))/2^{n+1}$, which do not meet the set $G^s(n)$, and letting $H_i = (S_{i,1}, ..., S_{i,s(n+1)})$.

$G(n + 1)$ is determined by $g(n)$:

If $g(n) = 0$, then $G(n + 1)$ is defined to be

$$H_1 * G_1 * \cdots * G_2 * H_{1+2^{n+1}}$$

and

if $g(n) = 1$, then $G(n + 1)$ is

$$H_{1+2^{n+1}} * G_1 * \cdots * G_2 * H_1.$$ 

Now to construct $G(n + 1)$, first divide each column $C$ of $G(n + 1)$ into $f(n + 1) - 1$ subcolumns $C(j), 1 \leq j < f(n + 1)$, of equal width. Precede $C(j)$ by $j$ intervals in $P_f$ and follow it by $f(n + 1) - j$ intervals in $P_e$. These are the columns of $G(n + 1)$.

Ornstein and Shields show that the measure automorphism $T$ which is the common extension of all the gadget transformations is a $K$-automorphism which is not Bernoulli. They also show that, for fixed sequences $f$ and $s$, if $T$ and $T'$ are two such automorphisms, depending on two different sequences $g(n)$ and $g'(n)$ of zeros and ones, they are isomorphic if and only if $g(n) = g'(n)$ for all but finitely many $n$.

An Induced Automorphism

Let $A$ be all of $G(1)$ except the first interval of the first column. We will show that $T_A$ is a Bernoulli shift.

First notice that $T_A$ is simply $T$ on all of $A$ except $\text{Root}(G(1))$. To see what $T_A$ is on all of $A$, we will write $A$ as a subset of each gadget.

Observe the following sequences of towers, described inductively:

Let $J(1) = A$.

Assume that $J(n)$ has been formed.

Divide $J(n)$ into $2^{n+1}$ disjoint isomorphic towers, $J_1, ..., J_{2^{n+1}}$.

Let $J(n + 1) = J_1 * J_2 * \cdots * J_{2^{n+1}}.$
Divide each column of \( J(n) \) into \( f(n + 1) - 1 \) columns. Call the tower composed of these columns \( J(n + 1) \).

Note that \( f(n) \) is just that part of \( G(n) \) which intersects \( A \). Thus \( T_A \) is exactly the common extension of the tower transformations \( T(f(n)) \).

It is known that a transformation formed by the independent cutting and stacking of a tower which has two columns whose heights differ by one is a Bernoulli shift \([3, 9]\) so we have obtained the fact that \( T_A \) is a Bernoulli shift.

**Proposition.** All Ornstein-Shields K-automorphisms and all Bernoulli shifts of finite entropy are in the same weak equivalence class.

**Proof.** It is known that all Bernoulli shifts are weakly equivalent (Saleski [8] also Friedman [3]). But we have just seen that every Ornstein-Shields K-automorphism is weakly equivalent to a Bernoulli shift.

**Corollary.** The Bernoulli flows of finite entropy may be written as special flows over any Ornstein-Shields K-automorphism.

**Proof.** The Bernoulli flows may be written as special flows over Bernoulli shifts [6]. Then, since Bernoulli shifts and Ornstein-Shields K-automorphisms are weakly equivalent, the Bernoulli flows may also be written as special flows over any Ornstein-Shields K-automorphism [5].

**Proposition.** Ornstein-Shields K-automorphisms induce Bernoulli shifts on a dense family of sets.

**Proof.** Let \( T \) be any Ornstein-Shields K-automorphism. Choose \( n \). If \( A = A(n) \) is all of the \( n \)th gadget \( G(n) \) except the first interval of the first column, then, since \( T_A \) is formed by independent cutting and stacking, \( T_A \) is a Bernoulli shift. But \( \lim_{n \to \infty} \mu(A(n)) = 1 \), and so \( T \) induces Bernoulli shifts on sets of arbitrarily large measure. It is known that Bernoulli shifts induce Bernoulli shifts on a dense family of sets [3], and so \( T \) induces Bernoulli shifts on a dense family of sets.

**Ornstein-Shields K-Automorphisms and Entropy**

Ornstein and Shields show that for a fixed pair of sequences \( f \) and \( s \),

\[
\mu(T(f, s, g)) = \mu(T(f, s, g'))
\]

for any two sequences \( g \) and \( g' \) of zeros and ones. This can be seen by looking at induced automorphisms as well.

It is known that \( \mu(A) \mu(T_A) = \mu(T) \) for any ergodic automorphism \( T \) and set \( A \) of positive measure [1]. If \( A \) is \( G(1) \), neither \( T_A \) nor \( \mu(A) \) depends on the sequence \( g \), and so \( \mu(A) \mu(T_A) \) is constant, and thus for any sequences \( g \) and \( g' \),

\[
\mu(T(f, s, g)) = \mu(T(f, s, g')).
\]

Also, since the measure of the first gadget is smaller in cases where the sequences \( f \) and \( s \) grow faster, the entropy of \( T(f, s, g) \) is smaller for sequences \( f \) and \( s \) which grow faster.
ACKNOWLEDGMENT

This work was part of a Ph.D. thesis written at the University of California, Berkeley, under the supervision of Jacob Feldman. He was all the things an advisor should be and I thank him. Thanks are also due Nathaniel Friedman, who suggested the first problem, and Doug Lind, with whom I discussed many of these ideas.

Note added in proof. Since this paper was submitted, the work of Jack Feldman (Israel Journal 24 (1976), 16–38) and other work following that have answered most of the questions in the introduction.

REFERENCES