VECTORS OF MINIMAL NORM

ANGELA SPALSBURY

(Communicated by Theodore W. Gamelin)

Abstract. We characterize the minimal vectors (or extremal vectors) for the operator of multiplication by \((z - 1)\) on \(H^2\). We then give an application to infinite systems of equations.

1. Main result

In this paper we study the extremal vectors of a concrete operator. Extremal vectors were recently introduced by Ansari and Enflo [1], and were used to show the existence of invariant subspaces for certain classes of operators.

Our setting is a Hilbert space \(H\), and we consider a bounded, linear operator \(T\) on \(H\) that has dense range. Fix a vector \(x_0 \in H\), fix \(\epsilon > 0\), with \(\epsilon < ||x_0||\), and fix \(n \geq 1\). Since we are working in a Hilbert space, we can find a unique vector \(y_{n,x_0}^\epsilon\) of minimal norm such that

\[||T^n y_{n,x_0}^\epsilon - x_0|| \leq \epsilon,\]

i.e.,

\[||y_{n,x_0}^\epsilon|| = \inf\{||y|| : ||T^n y - x_0|| \leq \epsilon\}.\]

We call the vector \(y_{n,x_0}^\epsilon\) (or when the context is clear, \(y_n\)) the minimal vector for \(T^n\).

Ansari and Enflo used these minimal vectors to construct invariant subspaces for a large class of operators containing both the compact operators and the normal operators (but larger than their union). They showed that if \(T\) is compact then there is a subsequence of \((T^n y_n)\) that converges weakly to a vector that is noncyclic, and if \(T\) is normal then \((T^n y_n)\) converges in norm to a vector that is noncyclic. (In fact in both cases the vector is noncyclic for all operators commuting with \(T\).) Thus the orbit of this vector under the operator \(T\) is a (hyper) invariant subspace for \(T\).

We study an operator that is neither compact nor normal. Let \(T\) be the operator of “multiplication by \((z - 1)\)” on \(H^2\), i.e.,

\((Tf)(z) = (z - 1)f(z)\)

for any \(f \in H^2\). Note that \(T\) has dense range but is not invertible. We begin by characterizing the minimal vectors of \(T^n\).

Received by the editors February 11, 1997.

1991 Mathematics Subject Classification. Primary 47A15, 47B38.

Key words and phrases. Invariant subspaces, extremal vectors.

©1998 American Mathematical Society

2737
Theorem 1. Let \( T : H^2 \rightarrow H^2 \) be the operator of multiplication by \((z - 1)\). Fix \( \epsilon > 0 \), \( n \geq 1 \), and set \( x_0 = 1 \in H^2 \). Let \( y_n \) be the unique function in \( H^2 \) of minimal norm such that

\[
||T^n y_n(z) - 1|| = ||(z - 1)^n y_n(z) - 1|| \leq \epsilon.
\]

Then for each \( i = 1, \ldots, n \), there exist numbers \( a_i, |a_i| > 1 \), and there exists a constant \( A \) such that

\[
y_n(z) = A \frac{1}{(z - a_1) \cdots (z - a_n)}.
\]

In proving this theorem we need a geometric result from Ansari and Enflo [1, Theorem 1].

Theorem A (Ansari, Enflo). Let \( T \in L(H) \), with dense range. Fix \( x_0 \in H \), fix \( \epsilon > 0 \), with \( \epsilon < ||x_0|| \), and fix \( n \geq 1 \). Let \( y_n = y_{n,x_0} \) be the minimal vector for \( T^n \). Then there exists a \( \delta < 0 \) such that

\[
y_n = (\delta(T^*)^n T^n - I)^{-1} \delta(T^*)^n(x_0).
\]

Moreover, the following orthogonality relation holds:

if \( z \in H \) such that \( z \perp y_n \), then \( T^n z \perp T^n y_n - x_0 \).

We now give a proof of Theorem 1.

Proof. By Theorem A there exists a \( \delta < 0 \) such that

\[
y_n = (\delta(T^*)^n T^n - I)^{-1} \delta(T^*)^n(1).
\]

We need to introduce some notation. Let \( P_+ \) denote the operator of taking the “analytical part” of a power series, i.e.,

\[
P_+ \left[ \sum_{n=-\infty}^{\infty} a_n z^n \right] = \sum_{n=0}^{\infty} a_n z^n.
\]

Since \( T \) is multiplication by \((z - 1)\), \( T^* \) is

\[
P_+ \left[ \text{multiplication by } \left( \frac{1}{z} - 1 \right) \right],
\]

i.e., for any \( g \in H^2 \), and any \( z \in \mathbb{C} \),

\[
(T^* g)(z) = P_+ \left[ \left( \frac{1}{z} - 1 \right) g(z) \right].
\]

Without loss of generality, we may assume that

\[
(T^*)^n(1) = 1.
\]

Therefore by (1.1)

\[
\frac{1}{\delta} y_n = (\delta(T^*)^n T^n - I)^{-1}(1),
\]

so

\[
(\delta(T^*)^n T^n - I) \left( \frac{1}{\delta} y_n \right) = 1.
\]

In other words, we have that for any \( z \in \mathbb{C} \),

\[
(1.2) \quad (T^*)^n T^n y_n(z) - \frac{1}{\delta} y_n(z) = 1.
\]
Using our notation it is easy to check that for any \( g \in H^2 \), and \( z \in \mathbb{C} \),
\[
((T^*)^n T^n g)(z) = \mathcal{P}_+ \left[ \left( \left( \frac{1}{z} - 1 \right)(z-1) \right)^n g(z) \right]
\]
\[
= \mathcal{P}_+ \left[ \left( 2 - \frac{1}{z} - z \right)^n g(z) \right].
\]

Now fix \( z \in \mathbb{C} \). Then by (1.2)
\[
\mathcal{P}_+ \left[ \left( 2 - \frac{1}{z} - z \right)^n y_n(z) \right] - \frac{1}{\delta} y_n(z) = 1.
\]

Hence there exists a polynomial \( Q_n \) of degree less than or equal to \( n \) such that
\[
\left( 2 - \frac{1}{z} - z \right)^n y_n(z) + Q_n \left( \frac{1}{z} \right) - \frac{1}{\delta} y_n(z) = 1.
\]

Therefore
\[
y_n(z) \left( \left( 2 - \frac{1}{z} - z \right)^n - \frac{1}{\delta} \right) = 1 - Q_n \left( \frac{1}{z} \right),
\]
or
\[
\frac{1}{\delta} y_n(z) \left( \delta \left( 2 - \frac{1}{z} - z \right)^n - 1 \right) = 1 - Q_n \left( \frac{1}{z} \right).
\]

We have then that
\[
\frac{1}{\delta} y_n(z) = \frac{1 - Q_n \left( \frac{1}{z} \right)}{\delta \left( 2 - \frac{1}{z} - z \right)^n - 1}.
\]

Write \( Q_n(z) = b_0 + b_1 z + \cdots + b_n z^n \). Then
\[
y_n(z) = \frac{\delta \left( 1 - b_0 - \frac{b_1}{z} - \cdots - \frac{b_n}{z^n} \right)}{\delta(2 - \frac{1}{z} - z)^n - 1}
\]
\[
= \frac{\delta \left( z^n(1 - b_0) - b_1 z^{n-1} - \cdots - b_n \right)}{z^n(\delta \left( 2 - \frac{1}{z} - z \right)^n - 1)}.
\]

Set \( p_n(z) = \delta(z^n(1 - b_0) - b_1 z^{n-1} - \cdots - b_n) \), and \( s_n(z) = \delta(2 - \frac{1}{z} - z)^n - 1 \). Then
\[
y_n(z) = \frac{p_n(z)}{z^n s_n(z)}.
\]

Now \( y_n \in H^2(U) \), so \( y_n \) can have no poles inside the unit disk (or even on the unit circle). Clearly \( z_0 \) is a pole of \( y_n(z) \) only if
\[
z_0 = 0 \quad \text{or} \quad s_n(z_0) = 0.
\]

It is easy to see that \( s_n(z) \) has exactly \( 2n \) zeroes. Also for any \( z \)
\[
s_n \left( \frac{1}{z} \right) = s_n(z).
\]

Thus if \( z_0 \) is a zero of \( s_n \) then so is \( \frac{1}{z_0} \). Therefore the same number of zeroes of \( s_n(z) \) lie inside the disk as lie outside the unit disk.

Since \( s_n(z) \) has at most \( n \) zeroes outside the unit disk, it follows that \( y_n(z) \) can have at most \( n \) poles.
On the other hand, \( y_n \) has at least \( n \) poles. To see this just note that (1.3) tells us that

\[
\lim_{z \to \infty} y_n(z) = \lim_{z \to \infty} \frac{\text{constant}}{z^n}.
\]

This also shows that \( y_n \) must be a rational function.

Therefore \( y_n(z) \) is a rational function with exactly \( n \) poles. Hence there exist numbers \( a_i, |a_i| > 1 \) (for each \( i = 1, \ldots, n \)), and there exists a constant \( A \) such that

\[
y_n(z) = \frac{A}{(z-a_1)(z-a_2)\cdots(z-a_n)}. \quad \square
\]

We would like to thank A. Volberg who introduced the idea of using \( \mathcal{P}_+ \) and who showed that \( y_n \) has at least \( n \) poles.

Now that we have characterized the minimal vectors \( y_n \) for the operator \( T \) of multiplication by \( (z-1) \) on \( H^2 \) we consider the question of convergence of \( (T^n y_n) \).

From the proof of Theorem 1 we see that \( y_n \) is of the form

\[
y_n(z) = \frac{A}{(z-a_1)(z-a_2)\cdots(z-a_n)},
\]

where \( a_1, a_2, \ldots, a_n \) are the \( n \) zeros of the function

\[
s_n(z) = \delta \left( 2 - \frac{1}{z} - z \right)^n - 1
\]

found outside the unit disk, i.e., \( a_1, a_2, \ldots, a_n \) are the \( n \) solutions to the quadratic equation

\[
z^2 + \left( \frac{1}{\delta} \right)^n - 2 z + 1 = 0
\]

found outside the unit disk. We are currently working to show that \( (T^n y_n) \) is norm convergent. By arguments similar to those used in [1] the limit would be a noncyclic vector for \( T \). By Beurling’s theory of the invariant subspaces for the shift operator we would then have that \( (T^n y_n) \) converges to a nonsingular inner function.

2. An application to infinite systems of equations

In this section we look at the relationship between the Taylor coefficients of the minimal vectors of our operator \( T \) (multiplication by \( (z-1) \) on \( H^2 \)). It is interesting to note that the Taylor coefficients satisfy certain recurrence relations, or equivalently they are solutions to certain infinite systems of equations. In Theorem 4 we characterize the solutions to such systems of equations using Theorem 1.

**Proposition 2.** Let \( T : H^2 \to H^2 \) be the operator of multiplication by \( (z-1) \). Fix \( \epsilon > 0 \), and let \( y (= y_1) \) be the unique function in \( H^2 \) of minimal norm such that

\[
||Ty - 1|| = ||(z-1)y(z) - 1|| \leq \epsilon.
\]

If we express \( y \) as a power series

\[
\sum_{n=0}^{\infty} a_n z^n,
\]

License or copyright restrictions may apply to redistribution; see https://www.ams.org/journal-terms-of-use
then the Taylor coefficients must satisfy the following recursive relation:

\[ a_0 = \delta (2a_0 + 1 - a_1), \]
\[ a_1 = \delta (-a_0 + 2a_1 - a_2), \]
\[ a_2 = \delta (-a_1 + 2a_2 - a_3), \]
\[ \vdots \]
\[ a_k = (-a_{k-1} + 2a_k - a_{k+1}), \]

or equivalently,

\[ a_1 = \left( \frac{2\delta - 1}{\delta} \right) a_0 + 1, \]
\[ a_2 = \left( \frac{2\delta - 1}{\delta} \right) a_1 - a_0, \]
\[ \vdots \]
\[ a_k = \left( \frac{2\delta - 1}{\delta} \right) a_{k-1} - a_k, \]

for all \( k \geq 2 \). (Here \( \delta \) is as in Theorem A.)

**Proof.** If we write

\[ y(z) = \sum_{i=0}^{\infty} a_i z^i, \]

then we see that

\[ (z - 1)y(z) = z \sum_{i=0}^{\infty} a_i z^i - \sum_{i=0}^{\infty} a_i z^i \]
\[ = \sum_{i=0}^{\infty} a_i z^{i+1} - \sum_{i=0}^{\infty} a_i z^i \]
\[ = \sum_{j=1}^{\infty} a_{j-1} z^j - \sum_{i=0}^{\infty} a_i z^i \]
\[ = -a_0 + \sum_{j=1}^{\infty} (a_{j-1} - a_j) z^j, \]

and therefore

\[ (z - 1)y(z) - 1 = (-a_0 - 1) + \sum_{j=1}^{\infty} (a_{j-1} - a_j) z^j. \]

Now suppose that \( g \in H^2 \) with \( g \perp y \). By Theorem A we know that

\[ Tg \perp Ty - 1. \]

Therefore if we write

\[ g(z) = \sum_{i=0}^{\infty} b_i z^i, \]
then

\[(z-1)g(z) \perp (z-1)y(z) - 1,\]

and so

\[-b_0(-a_0 - 1) + \sum_{j=1}^{\infty} (b_{j-1} - b_j)(a_{j-1} - a_j) = 0.\]

Therefore

\[0 = -b_0(-a_0 - 1) + \sum_{j=1}^{\infty} b_{j-1} \overline{a_{j-1}} + \sum_{j=1}^{\infty} b_j \overline{a_j} - \sum_{j=1}^{\infty} b_j a_{j-1} - \sum_{j=1}^{\infty} b_{j-1} a_j\]

\[= -b_0(-a_0 - 1) + \sum_{n=0}^{\infty} b_n \overline{a_n} + \sum_{n=1}^{\infty} b_n a_n - \sum_{n=1}^{\infty} b_n a_{n-1} - \sum_{n=0}^{\infty} b_n a_{n+1}\]

\[= b_0(a_0 + 1) + b_0 \overline{a_0} + \sum_{n=1}^{\infty} b_n \overline{a_n} + \sum_{n=1}^{\infty} b_n a_n - \sum_{n=1}^{\infty} b_n a_{n-1} - \sum_{n=0}^{\infty} b_n a_{n+1} - b_0 \overline{a_1}\]

\[= b_0(2a_0 + 1 - a_1) + \sum_{n=1}^{\infty} b_n(-a_{n-1} + 2a_n - a_{n+1}).\]

Set

\[h(z) = (2a_0 + 1 - a_1) + \sum_{n=1}^{\infty} (-a_{n-1} + 2a_n - a_{n+1})z^n.\]

Then \(h \in H^2\), since

\[||h(z)||^2 = |2a_0 + 1 - a_1|^2 + \sum_{n=1}^{\infty} |a_{n-1} + 2a_n - a_{n+1}|^2\]

\[\leq |2a_0 + 1 - a_1|^2 + \sum_{n=1}^{\infty} (|a_{n-1}|^2 + 2|a_n|^2 + |a_{n+1}|^2)\]

\[\leq |2a_0 + 1 - a_1|^2 + 4 \sum_{n=1}^{\infty} |a_{n-1}|^2\]

\[= |2a_0 + 1 - a_1|^2 + 4||f||^2 < \infty.\]

Now \(h\) has the property that

if \(g \in H^2\) with \(g \perp y\), then \(g \perp h\).

Therefore there exists a number \(\alpha\) such that \(y = \alpha \cdot h\). In fact, more is true:

\[\alpha = \delta,\]

where \(\delta\) is found as in Theorem A. (This is apparent on examining the proof of Theorem A.)

Thus

\[y(z) = \sum_{n=0}^{\infty} a_n z^n = \delta(2a_0 + 1 - a_1) + \delta \sum_{n=1}^{\infty} (-a_{n-1} + 2a_n - a_{n+1})z^n.\]
Equating coefficients, we have
\[ a_0 = \delta(2a_0 + 1 - a_1), \]
\[ a_1 = \delta(-a_0 + 2a_1 - a_2), \]
\[ a_2 = \delta(-a_1 + 2a_2 - a_3), \]
\[ \vdots \]
\[ a_n = \delta(-a_{n-1} + 2a_n - a_{n+1}), \]
or
\[ a_1 = \left(\frac{2\delta - 1}{\delta}\right) a_0 + 1, \]
\[ a_2 = \left(\frac{2\delta - 1}{\delta}\right) a_1 - a_0, \]
\[ \vdots \]
\[ a_n = \left(\frac{2\delta - 1}{\delta}\right) a_{n-1} - a_{n-1}. \]

We now characterize the solutions to this infinite system of equations.

**Theorem 3.** Let \((a_0, a_1, \ldots) \in \ell^2\). If \((a_0, a_1, \ldots)\) are solutions to the infinite system of equations,
\[ a_0 = \delta(2a_0 + 1 - a_1), \]
\[ a_1 = \delta(-a_0 + 2a_1 - a_2), \]
\[ a_2 = \delta(-a_1 + 2a_2 - a_3), \]
\[ \vdots \]
\[ a_n = \delta(-a_{n-1} + 2a_n - a_{n+1}), \]
(\(\delta\) any nonzero real number), then \((a_0, a_1, \ldots)\) are the Taylor coefficients of some \(H^2\) function \(g\) of the form
\[ g(z) = \frac{B}{(z - b)}, \]
where \(b\) is the solution to the quadratic equation
\[ z^2 + \left(\frac{1}{\delta}\right)^n - 2 = 0 \]
found outside the unit disk. The converse is also true provided \(\delta\) is as in Theorem A.

**Proof.** Suppose that \((a_0, a_1, \ldots)\) are solutions to the infinite system of equations. Define a function \(g \in H^2\) by
\[ g(z) = \sum_{k=0}^{\infty} a_k z^k. \]
Then it is easy to verify that (1.2) holds for any $\delta$, i.e., for any $\delta$,

$$(T^*)^n T^n g(z) - \frac{1}{\delta} g(z) = 1,$$

for all $z \in \{z : |z| < 1\}$. (Here $T$ is the operator of multiplication by $(z-1)$ on $H^2$.) Then, exactly as in the proof of Theorem 1, we can again conclude that $g$ is of the form

$$g(z) = \frac{B}{(z - b)}.$$

The converse follows directly from Theorem 1 and Proposition 2. \hfill \Box

We conclude by stating an analogous result for $T^n$.

**Theorem 4.** Let $(a_0, a_1, a_2, \ldots) \in l_2$. Fix $n \geq 1$. If $(a_0, a_1, a_2, \ldots)$ are solutions to the infinite system of equations,

\begin{align*}
a_0 &= \delta \left( \binom{2n}{n} a_0 - \binom{2n}{n+1} a_1 + \cdots + (-1)^n \binom{2n}{2n} a_n + (-1)^{n+1} \right), \\
a_1 &= \delta \left( - \binom{2n}{n-1} a_0 + \binom{2n}{n} a_1 + \cdots + (-1)^{n+1} \binom{2n}{2n-1} a_n \\
&\quad\quad\quad+ (-1)^n \binom{2n}{2n} a_{n+1} \right), \\
&\vdots \\
a_{n-1} &= \delta \left( (-1)^{n+1} \binom{2n}{1} a_0 + (-1)^n \binom{2n}{2} a_1 + (-1)^{n+1} \binom{2n}{3} a_2 \\
&\quad\quad\quad+ \cdots + (-1)^n \binom{2n}{2n} a_{2n-1} \right), \\
a_j &= \delta \left( (-1)^n \binom{2n}{0} a_{j-n} + (-1)^{n+1} \binom{2n}{1} a_{j-(n+1)} \\
&\quad\quad\quad+ \cdots + (-1)^n \binom{2n}{2n} a_{j+n} \right), \text{ for all } j \geq n
\end{align*}

($\delta$ any arbitrary nonzero real number), then $(a_0, a_1, a_2, \ldots)$ are the Taylor coefficients of an $H^2$ function $g$ of the form

$$g(z) = \frac{B}{(z - b_1) \cdots (z - b_n)},$$

where $b_1, b_2, \ldots, b_n$ are the $n$ solutions to the quadratic equation

$$z^2 + \left( \frac{1}{\delta} \right)^n - 2 + 1 = 0$$

found outside the unit disk. The converse is also true provided $\delta$ is as in Theorem A.

**Remark 5.** Using Proposition 2, it is possible to show that the minimal vector $y_1$ for $T$ is of the form

$$y_1(z) = \frac{A}{(z - a)}.$$
However, it is not clear how to generalize this technique to show that the minimal vector $y_n$ for $T^n$ ($n > 1$) is of the form

$$y_n(z) = \frac{A}{(z - a_1)(z - a_2) \cdots (z - a_n)}.$$ 

References


Department of Mathematics, University of the Witwatersrand, PO Box 3, Wits 2050, South Africa

Current address: Department of Mathematics and Computer Science, Kent State University, Kent, Ohio 44242

E-mail address: angie@mcs.kent.edu