Theorem (Textbook 2.2)
Let $x_1$ and $x_2$ be solutions of
\[
\begin{bmatrix}
\Phi & \Psi
\end{bmatrix} x = b.
\]
Then,
\[
\| x_1 \|_0 + \| x_2 \|_0 \geq \frac{2}{\mu(A)}.
\]

Proof: Let $e = x_1 - x_2$

Since $Ax_1 = b$ and $Ax_2 = b$, we have
\[
Ae = A(x_1 - x_2) = 0.
\]
So, $e \in \text{Null}(A)$.
Partition $e$ into $e = \begin{bmatrix} e_\Phi \\ e_\Psi \end{bmatrix}$.

Then,
\[
\Phi e_\Phi + \Psi e_\Psi = 0
\]
\[
\Rightarrow \quad \Phi e_\Phi = -\Psi e_\Psi
\]
Let $\Phi e_\Phi = y = -\Psi e_\Psi$.

Then, theorem 2.1 implies that
\[
\| e_\Phi \|_0 + \| e_\Psi \|_0 \geq \frac{2}{\mu(A)}
\]
\[
= \| e \|_0 , \text{ because } e = \begin{bmatrix} e_\Phi \\ e_\Psi \end{bmatrix}.
\]
Since $e = x_1 - x_2$,

\[ \|e\|_0 \leq \|x_1\|_0 + \|x_2\|_0 \] (by triangle inequality)

\[ \Rightarrow \|x_1\|_0 + \|x_2\|_0 \geq \frac{2}{\mu(A)}. \]

**Theorem (Textbook 2.3)**

If a candidate solution for $(\Phi, \Phi)x = b$ has less than $\frac{1}{\mu(A)}$ non-zeros, then it is necessarily the sparsest solution.

**Proof:** If $\|x\|_0 \leq \frac{1}{\mu(A)}$, and if $y$ is another solution. Then, by theorem 2.2,

\[ \|x\|_0 + \|y\|_0 \geq \frac{2}{\mu(A)} \]

\[ \Rightarrow \frac{1}{\mu(A)} + \|y\|_0 \geq \frac{2}{\mu(A)} \]

\[ \Rightarrow \|y\|_0 \geq \frac{1}{\mu(A)} \]

\[ \geq \|x\|_0. \]

So $y$ has at least $\|x\|_0$ non-zeros.
Spark

The spark of a matrix $A$ is the smallest number of columns that are linearly dependent.

Spark VS Rank

Rank is the largest number of columns that are linearly independent.

Computing Spark($A$) is NP hard.

Property: if $x \in \text{Null}(A)$, then $\|x\|_0 \geq \text{spark}(A)$.

**Proof:** if $x \in \text{Null}(A)$, then $A x = 0$.

\[
\begin{pmatrix}
a_1 & a_2 & \ldots & a_m
\end{pmatrix}
\begin{pmatrix}
x_1 \\
\vdots \\
x_m
\end{pmatrix} = 0
\]

$\|x\|_0$ = number of linearly dependent columns using this particular $x \in \text{Null}(A)$.

Spark($A$) = smallest among all possible $x$.

So $\|x\|_0 \geq \text{spark}(A)$. 
Theorem (Textbook 2.4)
If a system $Ax = b$ has a solution such that $\|x\|_0 < \text{spark}(A)/2$, then $x$ is necessarily the most sparse solution.

Proof
Let $y$ be another solution. Then $Ay = b$, and $Ax = b$.

So $A(x-y) = 0$.

$\Rightarrow$ $A^t(x-y) \in \text{Null}(A)$.

By property of $\text{spark}(A)$, we have

$\|x-y\|_0 \leq \text{spark}(A)$.

So by triangle inequality we have

$\|x\|_0 + \|y\|_0 \geq \|x-y\|_0 \geq \text{spark}(A)$.

Therefore, if $\|x\|_0 < \text{spark}(A)/2$, then

$\|y\|_0 > \text{spark}(A)/2$.

So $y$ cannot be sparser than $x$.

Importance of Theorem 2.4

$p_0$ is NP-hard. Theorem 2.4 provides a global optimal criteria.