On the Decay Rate of the Charged Pion.

J. Bernstein (*)

Faculté des Sciences - Orsay

S. Fubini

Istituto di Fisica dell'Università - Padova
CERN - Geneva

M. Gell-Mann (**) 

Collège de France and Ecole Normale Supérieure - Paris (***)

W. Thirring

Institute for Theoretical Physics, University of Vienna - Vienna
CERN - Geneva

(ricevuto il 7 Giugno 1960)

Summary. — The Goldberger-Treiman relation between the pion decay amplitude and the axial vector coupling constant in β-decay is shown to be plausible if the divergence of the axial vector current is a highly non-singular operator (i.e., one that emphasizes low frequencies). Such a situation is probably realized, for example, in a renormalizable theory for which this operator is proportional to the conventional pion field. Some other possible consequences of the low-frequency character of the operator are mentioned.

1. — Introduction.

Goldberger and Treiman (*) have stated a formula for the rate of charged pion decay that is in excellent agreement with experiment. Their derivation of the formula was based, however, on a number of assumptions, some of

(*) National Science Foundation Post-Doctoral Fellow.
(**) National Science Foundation Senior Post-Doctoral Fellow.
(***) Permanent address: California Institute of Technology, Pasadena, Cal., U.S.A.
which are not easy to justify. Various authors (2, 3) have therefore tried to find other chains of argument that would lead, at least approximately, to the same formula. Some of these attempts have involved abandoning the conventional form of the strong pion-nucleon interaction or the conventional form of the weak axial vector current or both.

Recently (4) the possibility has been considered that the strong interactions and the weak current are such that the divergence of the axial vector current \( P_\alpha \) (for baryons and mesons with \( \Delta S = 0 \)) is proportional to a component of the pion field. Under that condition the G-T formula can be made particularly plausible. A number of models of strong and weak interactions possessing the property in question were presented, but none of them was free of difficulties.

It is interesting, therefore, to see if the kind of argument that was used in A to obtain the G-T formula can be made more general. It is evident at once that there is something artificial about the requirement that \( \delta P_\alpha \) be proportional to \( \pi^- \), since it is somewhat arbitrary which pseudoscalar field with \( S = 0 \) and \( I = 1 \) is called the pion field. With this in mind, let us try to isolate, in a way that does not involve a choice of which field is to be called \( \pi^- \), the assumptions in A that are relevant to the deduction of the G-T formula.

2. – The essence of the G-T relation.

Whatever arbitrariness there may be in the definition of the renormalized pion field operator \( \pi^- \), its matrix elements between one-pion states and the vacuum are uniquely defined. The corresponding matrix element of the divergence of the axial vector current can always be written in the form

\[
\langle 0 | \delta P_\alpha | \pi^- \rangle = im_\pi 2^{-\frac{1}{2}} \langle 0 | \pi^- | \pi^- \rangle,
\]

where the constant \( a_1 \) corresponds to what was called \( a_\sqrt{Z} \) in A. The pion decay rate is then

\[
\Gamma_{\pi} = \frac{G^2}{16\pi m_\pi^2} \frac{m_\pi^2}{m_\pi^2} \left( 1 - \frac{m_\pi^2}{m_\pi^2} \right) a_1^2,
\]

where we have chosen the phases so that \( a_1 \) is real.

(5) M. GELL-MANN and M. LÉVY: Nuovo Cimento, 16, 705 (1960), to be referred to as A; and J. BERNSTEIN, M. GELL-MANN and L. MICHEL: Nuovo Cimento, 16, 560 (1960), to be referred to as B. We shall employ the notation of these articles and shall quote equations from them as (A-1), (B-2), etc.
The operator $\lambda^-$ defined by the equation

\begin{equation}
\lambda^- = -i\sqrt{2} a^- a^+ \gamma_\alpha P_\lambda
\end{equation}

evidently behaves like the renormalized pion field $\pi^-_\alpha$ between one-pion states and the vacuum and may always be thought of as a kind of candidate for the title of pion field. It was suggested in $\Lambda$ that the condition $\lambda^- = \pi^-_\alpha$ would make the G-T relation plausible. Whether or not this is so, the argument is that if certain properties that might reasonably be attributed to the operator $\pi^-_\alpha$ are possessed by the "candidate" $\lambda^-$, then the G-T relation could be understood in a simple way. We must try to specify, as nearly as possible, what are the desirable properties of $\lambda^-$. We are especially interested in the matrix-element of $\lambda^-$ between onenucleon states. Putting, as in (B-18),

\begin{equation}
\langle p | P_\alpha | n \rangle = -\frac{G}{\sqrt{2}} \bar{u}_i \gamma_\alpha \gamma_\beta u_i + ik_\beta(k^2) \bar{u}_i \gamma_\beta u_i ,
\end{equation}

and defining $K_\lambda$ by the relation

\begin{equation}
\langle p | \lambda^- | n \rangle = i\sqrt{2} \bar{u}_i \gamma_\beta u_i K_\lambda(k^2) ,
\end{equation}

we have

\begin{equation}
2m \left( -\frac{G}{\sqrt{2}} \right) \frac{z(k^2)}{k^2} = \frac{a_1}{K_\lambda(k^2)} .
\end{equation}

Since $z(0)$ is unity by definition, the formula relating the axial vector coupling constant to the pion decay amplitude is just

\begin{equation}
-\frac{G}{\sqrt{2}} = \frac{a_1}{2m} K_\lambda(0) .
\end{equation}

Now the function $K_\lambda$ is the same one that occurs in the creation of a nucleon-antinucleon pair from the vacuum by the operator $\lambda^-$, i.e., in the matrix element $\langle p, \bar{n} | \lambda^- | 0 \rangle$. As remarked in $\Lambda$, the function always has a pole at $k^2 = -m_\pi^2$ with residue $-g_1$, where $g_1$ is the renormalized pion-nucleon coupling constant, because $\lambda^-$ behaves as the renormalized pion field for the creation of real pions from the vacuum. This situation is perfectly general and was pointed out long ago (a) by Goldberger and Treiman and others.

in their work on the induced pseudoscalar term $\beta(k^2)$. The pole in question, coming from the intermediate one-pion state in the sense of dispersion theory, occurs only in the $\beta$ term of eq. (6). We have, then, in any theory,

$$K_2(k^2) \approx \frac{-g_1}{k^2 + m_\pi^2} \quad \text{for} \quad k^2 \approx -m_\pi^2,$$

or, what is the same thing,

$$K_2(k^2) = \frac{-g_1}{k^2 + m_\pi^2} \varphi_2(k^2) \quad \text{where} \quad \varphi_2(-m_\pi^2) = 1.$$

In this language, the G-T relation is simply the statement that $\varphi_2(k^2)$ remains close to unity between $-m_\pi^2$ and 0 so that $\varphi_2(0) \approx 1$. If $\varphi_2$ is slowly varying in this way, then we have

$$K_2(0) \approx \frac{-g_1}{m_\pi^2},$$

and thus

$$G_2 \sim \frac{-g_1}{2m_\pi k^2},$$

which is the G-T relation.

In order to explain the pion lifetime, all that must be established in any theory of the strong and weak couplings is that the one-pion pole in $K_2$ continues to dominate at $k^2 = 0$.

Similarly, in order to derive the prediction of Goldberger andTreiman for the induced pseudoscalar term in p capture by protons, all that must be established is that the one-pion pole in $\beta(k^2)$

$$\beta(k^2) \approx \frac{a_1 g_1}{m_\pi^2} \left( \frac{1}{k^2 + m_\pi^2} \right) \quad \text{for} \quad k^2 \approx -m_\pi^2,$$

continues to dominate at $k^2 = -m_\pi^2$.

Besides the pole at $k^2 = -m_\pi^2$, the functions $\beta$ and $K_2$ have branch lines running from $-9m_\pi^2$ to $-\infty$, which represent the contributions of the intermediate states other than the one-pion state, in the sense of dispersion theory. Either function, if it vanishes at infinity, may be written simply as the sum of an integral over the branch line and a term representing the pole. In other words, we have a dispersion relation without subtractions. For instance, if
in question, in ion theory, by,

\[ K_2 \rightarrow 0 \text{ as } |k^2| \rightarrow \infty, \text{ then we have} \]

\begin{equation}
K_2(k^2) = \frac{-g_1}{k^2 + m_i^2} + \int_{9m_i^2}^{\infty} dM^2 \frac{\sigma(M^2)}{k^2 + M^2}.
\end{equation}

If, however, \( K_2 \) fails to approach zero at infinity, the we must make a subtraction in the dispersion relation (13) and the value of \( K_2 \) at some particular value of \( k^2 \), such as \( k^2 = 0 \), become an arbitrary constant.

The weight function \( \sigma \) (which is not necessarily positive) is the sum of products of two factors, one representing the amplitude for the creation of a state of mass \( M \) from the vacuum by the operator \( \lambda^- \) and the other representing the transition amplitude from such a state to a final state of proton and antineutron.

One way to understand the success of the G-T relation, then, is to suppose that \( \lambda \) is an operator which, acting on the vacuum, emphasizes low frequencies, in such a way that:

\( a) \) The matrix elements for creation of states of high mass tend to zero as the mass goes to infinity; in particular \( K_2 \), which represents \( \langle p, \bar{\pi} | \lambda^- | 0 \rangle \), tends to zero (\(^1\)), giving the unsubtracted dispersion relation (13). (We note that if \( K_2 \rightarrow 0 \) and if \( \alpha \) is not to increase like \( k^2 \) at infinity, then \( \beta \rightarrow 0 \) and obeys an unsubtracted dispersion relation).

\( b) \) The matrix elements for creation of higher states than the pion are not large, so that, for normal values of the transition amplitudes mentioned above, the one pion term \( -g_1/m_i^2 \) dominates the integral

\[ \int_{9m_i^2}^{\infty} \frac{\sigma(M^2) dM^2}{M^2}. \]

This point of view is not without experimental implications, although the experiments involved are difficult to perform at the present time. For example, we should expect the one-pion pole to predominate near \( k^2 = 0 \) in the case of the matrix elements \( \langle \Sigma^+ | \lambda^+ | 0 \rangle \) and \( \langle A, \Sigma^- | \lambda^- | 0 \rangle \), giving an approximate relation like the G-T formula linking the renormalized pion coupling constant for \( \Lambda \leftrightarrow \Sigma \) transitions to the effective axial vector coupling constants for \( \Sigma^+ \rightarrow \Lambda + e^+ + \nu \) and \( \Sigma^- \rightarrow \Lambda + e^- + \bar{\nu} \) (or, more accurately, the average of these two.)

\(^1\) This crucial point has recently been made independently by Y. Nambu: Phys. Rev. Lett., 4, 380 (1960).
If we apply the same idea to the matrix element \( \langle N', P; 0 \mid Z^- \mid 0 \rangle \), we obtain an approximate prediction for small \( k^2 \) of the value of the axial vector amplitude for processes like \( \bar{\Psi} + N \rightarrow \pi^+ + N' + \pi \), in terms of the pion-nucleon scattering amplitude.

In summary, then, what we suppose to be at the root of the G-T relation is that \( Z^- \) is a highly non-singular operator that emphasizes low frequencies. In the limit of high frequencies, then, the axial vector current is conserved.

The formal equality \( Z^- = \pi^- \) (where \( \pi^- \) is a conventionally defined pion field) is useful in that, in a renormalizable theory like the \( \sigma \) model of article A it makes plausible the non-singular character of \( Z^- \).

3. – The propagator.

For a conventional pion field \( \pi^- \), it is customary to consider the renormalized pion propagator \( 0 \mid P(\pi_-(x), \pi_+(y)) \mid 0 \) with Fourier transform (8)

\[
(14) \quad \frac{d_x(k^2)}{k^2 + m_\pi^2 - i\epsilon} = \frac{1}{k^2 + m_\pi^2 - i\epsilon} + \int_{\frac{m_\pi^2}{k^2}}^{\infty} \frac{g_2(M^2) dM^2}{k^2 + M^2 - i\epsilon}
\]

and to suppose that the integral converges, i.e., that the pion propagator in momentum space is finite.

For the operator \( Z^- \), we may also construct a propagator \( 0 \mid P(Z_-(x), Z_+(y)) \mid 0 \) and write its Fourier transform formally as

\[
(15) \quad \frac{d_x(k^2)}{k^2 + m_\pi^2 - i\epsilon} = \frac{1}{k^2 + m_\pi^2 - i\epsilon} + \int_{\frac{m_\pi^2}{k^2}}^{\infty} \frac{g_2(M^2) dM^2}{k^2 + M^2 - i\epsilon}
\]

If \( Z \) is a sufficiently non-singular operator so that it gives a finite propagator in momentum space, then it is true \( a fortiori \) that \( K_z \rightarrow 0 \) at infinity. We may see this by writing \( g_2(M^2) = g_2^{\text{str}} + g_2^{\text{var}} \), where \( g_2^{\text{var}} \) is the contribution to the weight function \( g_2 \) from states containing one nucleon and one antinucleon and is given by

\[
(16) \quad g_2^{\text{var}} = (8\pi^2)^{-1} [M^2(M^2 - 4m^2)]^2 (M^2 - 4m^2) |K_z(-M^2)|^2.
\]

Then we have
\[ m_\pi^2 d_2(0) = m_\pi^2 + (8\pi^2)^{-1} \int \frac{dM^2}{4M^2} |K_\lambda(M^2)|^2 \left( 1 - \frac{4m_\pi^4}{M^2} \right)^\frac{3}{2}, \]
which, of course, forces \( K_\lambda \) to go to zero at infinity. Even if \( d_2(0) - 1 = m_\pi^2 \frac{(d_2(M^2))}{M^2} \) diverges, but not so badly as to make the next moment \( \int \frac{dM^2}{4M^2} |K_\lambda(M^2)|^2 \) converge as well, that is sufficient to make \( K_\lambda \to 0 \).

It is interesting that we can formally calculate \( d_2(0) \), using the definition (3) of \( \lambda^\pm \) as proportional to the divergence of the axial vector current. We find
\[ 0 \langle P(\lambda^+(x), \lambda^-(y)) | 0 \rangle = -i \sqrt{2} a_1^{-1} \left\{ \frac{\gamma^\pm}{\gamma^+} \right\} \langle P(\lambda^+(x), \lambda^-(y)) | 0 \rangle + i \partial(x_0 - y_0) \langle [P(\lambda^+, \lambda^-), \lambda^+(y)] | 0 \rangle. \]

We may put
\[ [P_\lambda(x, x_0), \lambda^+(y, x_0)] = \delta(x - y)(-i 2 \gamma^\pm) \lambda^+(x, x_0). \]

Going over to the Fourier transform \( (k^2 + m_\pi^2)^{-1} d_2(k^2) \), we remark that the divergence term in (18) contributes nothing at \( k = 0 \), so that we have
\[ m_\pi^2 d_2(0) = \langle 0 | X | 0 \rangle a_1^{-1}. \]

In the \( \sigma \) model considered in \( \Lambda \), \( X \) is simply \( 2m_\pi^2 \sigma = 2\sigma Z_\pi^{-1} \), so that we have
\[ m_\pi^2 d_2(0) = m_\pi^2 d_3(0) = 2a_1^{-1} \langle 0 | \sigma Z_\pi^{-1} | 0 \rangle. \]

According to the \( \sigma \) model, the quantities \( d_3(0), a_3, \) and \( \langle 0 | \sigma Z_\pi^{-1} | 0 \rangle \) are all finite, at least in the series expansion of perturbation theory.

In the gradient coupling model, \( X \) is just the number \( \mu_0^2/2a_1^{-1} \). Since in that model \( a_1 = -Z_\pi^2 \mu_0^2 f_a^{-1} \), eq. (20) becomes
\[ m_\pi^2 d_2(0) = m_\pi^2 d_3(0) = \mu_0^2 Z^{-1}, \]
which we know to be true for the gradient coupling model. (See eq. (A-35)).

In the usual pseudoscalar pion theory with pseudoscalar coupling to nucleons and with \( P_\pi \) taken to be simply \( \overline{\gamma}_\lambda \gamma_\lambda \), we obtain formally
\[ X = -4a_1^{-1} (m_\pi \overline{\gamma}_\lambda \gamma_\lambda - ig_0 \overline{\gamma}_\lambda \gamma_\lambda \gamma_5 \pi \cdot \pi N) = 4a_1^{-1} \overline{\gamma}_\lambda \gamma_\lambda \gamma_5 \pi \cdot \pi N, \]
and it looks as if \( X \) and \( d_2(0) \) are highly divergent. Of course that does not exclude the possibility of \( K_\lambda \) going to zero any way.
4. – Correspondance with the arguments of G-T.

In a theory in which $\lambda^-$ is not the conventional $\pi^-$ field and in which we cannot establish that $\lambda^-$ gives a sensible pion propagator, we may try to get some information about $K_2$ by looking at a mixed propagator, which is neither

$$\langle 0 | P(\pi_-(x), \pi^+_+(y)) | 0 \rangle$$

nor

$$\langle 0 | P(\lambda^-(x), \lambda^+_+(y)) | 0 \rangle,$$

but

$$\frac{1}{2} \left( \langle 0 | P(\lambda^-(x), \pi^+_+(y)) | 0 \rangle + \langle 0 | P(\pi^-(x), \lambda^+_+(y)) | 0 \rangle \right).$$

The Fourier transform of the first propagator gives, of course

$$(21a) \frac{d_\pi(k^2)}{k^2 + m_\pi^2 - i\epsilon} = \frac{1}{k^2 + m_\pi^2 - i\epsilon} + \int_0^\infty \frac{\sqrt{(M^2 - 4m^2)M^2} \, dM \, |K_\pi(-M^2)|^2}{8\pi^2} \frac{1}{k^2 + M^2 - i\epsilon} + \cdots,$$

where the remaining terms, like the one-pion and one-pair terms, have positive weight functions. Likewise the second propagator gives, in Fourier transform

$$(21b) \frac{d_\lambda(k^2)}{k^2 + m_\lambda^2 - i\epsilon} = \frac{1}{k^2 + m_\lambda^2 - i\epsilon} + \int_0^\infty \frac{\sqrt{(M^2 - 4m^2)M^2} \, dM \, |K_\lambda(-M^2)|^2}{8\pi^2} \frac{1}{k^2 + M^2 - i\epsilon} + \cdots,$$

with positive weight functions everywhere. The mixed propagator gives

$$(21c) \frac{d_{\text{mixed}}(k^2)}{k^2 + m_\pi^2 - i\epsilon} = \frac{1}{k^2 + m_\pi^2 - i\epsilon} + \int_0^\infty \frac{\sqrt{(M^2 - 4m^2)M^2} \, dM \, \text{Re} \, K_\pi(-M^2)K_\lambda^*(-M^2)}{8\pi^2} \frac{1}{k^2 + M^2 - i\epsilon},$$

where the weight functions need not be positive.

Now just as we showed that $m_\pi^{-2} d_\pi(0) = a_1^{-1} \langle 0 | X | 0 \rangle$, we can define

$$[P_\lambda(x, x_0), \pi^+_+(y, y_0)] = \delta(x - y) \left( -\frac{i}{\sqrt{2}} \right) Y,$$

and obtain

$$(22) \quad m_\pi^{-2} d_{\text{mixed}}(0) = \langle 0 | Y | 0 \rangle a_1^{-1}.$$
In the simple pseudoscalar theory with $P_\pi = \bar{P}_\pi \gamma_5 \gamma_\mu n$, the operator $Y$ vanishes and so we obtain

$$0 = \frac{1}{m_\pi^2} \int_{m_\pi^2}^{\infty} \frac{\sqrt{(M^2 - M^2 - 4m^2) \ dM^2}}{8\pi^2 M^2} \ Re \ K_\lambda(-M^2)K_\pi^*(M^2 - M^2) + \ldots,$$

which is the starting point of the method used by Goldberger and Treiman to derive their relation. Using eq. (9) we may rewrite the equation in the form

$$\frac{1}{m_\pi^2} \int_{m_\pi^2}^{\infty} \frac{\sqrt{(M^2 - M^2 - 4m^2) \ dM^2}}{8\pi^2 M^2} \ Re \ q_\lambda (-M^2)q_\pi^*(M^2 - M^2) \ (M^2 - m_\pi^2)^2 + \ldots.$$

Unfortunately, the absence of any requirement of positiveness means that we cannot use inequalities. Goldberger and Treiman therefore are obliged to throw away the remaining terms, such as contributions from $3\pi$ states, and an unknown error is thereby introduced.

Even if we do throw away the other terms, we do not have clear sailing since we do not know the functional forms of $q_\lambda$ or $q_\pi^*$; Goldberger and Treiman get around these difficulties by two more assumptions:

1) They estimate $q_\pi$ by a dispersion theory calculation using one-pion and one-pair states only.

2) They assume a simple trial form for $q_\lambda$, namely $1 + (k^2 m_\pi^2 y/m_\pi^2)$ where $y$ is a number. (Note that $K_\lambda$ then approaches a constant at infinity, rather than zero, which would be our conjecture. This is not a very serious difference between the points of view, because the trial form need hold only up to a few GeV in the work of G-T, not necessarily all the way to infinity.)

Finally they use eq. (32) to calculate $y$ and find that it is small; it follows that $q_\lambda(0)$ is close to 1, which is of course, their important result.

5. - Conclusions.

The success of the G-T formula for the rate of pion decay indicates that the function $K_\lambda$ is dominated near $k^2 = 0$ by its one pion pole. Probably the operator $\lambda^\pi$ is sufficiently non-singular so that $K_\lambda$ vanishes as $k^2 \rightarrow \infty$.

If so, it is possible that $\lambda$ is even more non-singular and that $\lambda$, treated as the renormalized pion field, gives a pion propagator with a convergent spectral representation. It is a description of such a situation that was attempted in
A and B, particularly in the σ model. If the situation obtains, then the vacuum expectation value of the commutator of \( P_i \) and \( \lambda \) must be finite.

That \( \lambda \) be a suitable pion field is, however, only a sufficient condition for the vanishing of \( \mathcal{K}_2 \) at \( \infty \); it is not necessary, however, and we may if we like, concentrate on the implications of the dominance of the one-pion pole as the explanation of the G-T formula.

First, there is no uncertainty in the relative sign of the axial vector term \(-\partial_z z(k^2)\) and the induced pseudoscalar term \(\theta_p(k^2)\). (In the problem of muon capture, the possibility of varying the relative sign has been discussed \(^\text{(*)}\); crude comparisons of theory and experiment favor the sign given above, which is also the original sign of Goldberger and Treiman.)

Second, the matrix element of \( \lambda \) between \( \Sigma \) and \( A \) may also be dominated by the one-pion pole at small \( k^2 \). Thus we can relate the strength of the axial vector interaction in the decay \( \Sigma \rightarrow \Lambda + \text{leptons} \) to the value of the renormalized coupling constant for the pion to \( \Sigma \) and \( \Lambda \).

Third, the matrix element of \( \lambda \) between a nucleon state and a state with one nucleon and one pion may be given mainly by the one-pion pole. That would permit the calculation of the axial vector amplitude for \( n-\nu \rightarrow \pi^0 + e^- + \nu \), etc., at small momentum transfers.

A discussion of these points and of the role of the one-pion pole in nuclear \( 2 \beta \)-transitions will be given elsewhere \(^{(*)} \).


\(^{(*)} \) \textit{Note added in proof:} Results analogous to ours have been obtained independently by Chou Kuang-Chao - Dubna Report D514 (1969).

**RIASSUNTO**

Si dimostra la plausibilità della relazione di Goldberger e Treiman fra vita media del pione e costante d'accoppiamento assiale nel caso in cui la divergenza della corrente pseudovettoriale è altamente non-singolare (cioè che contiene in prevalenza basse frequenze). Tale situazione ha probabilmente luogo per esempio in una teoria renormalizzabile in cui quest'operatore è proporzionale al campo mesonico. Si discutono alcune possibili conseguenze del carattere non singolare dell'operatore.