Particles as elementary excitations

To give a definition of particle we need time translations T_τ and spatial translations $T_{\mathbf{a}}$ acting on states. These notions allow us to define excitations of translation-invariant stationary state ω.

A state σ is an excitation of ω if for every observable A the expectation value of A in the state $T_{\mathbf{a}}\sigma$ tends to the expectation value of A in the state ω as $\mathbf{a} \to \infty$.

In classical mechanics we can assume that the states are represented by functions of coordinates $\mathbf{x}$, spatial translations shift the argument $\mathbf{x} \to \mathbf{x} + \mathbf{a}$, time evolution is specified by translation-invariant Hamiltonian. Translation -invariant state is represented by a constant function, an excitation is a bump; we do not see this bump anymore when it is shifted to infinity.

In the algebraic approach to quantum theory, every state (linear functional ω on the $*$-algebra of observables $\mathcal{A}$ obeying positivity condition $\omega(x^*x) \geq 0$) can be represented as a vector in some Hilbert space through so-called GNS (Gelfand-Naimark-Segal) construction. (More precisely, there exists a representation $A \to \hat{A}$ of $\mathcal{A}$ by operators in some Hilbert space $\mathcal{H}$ and a cyclic vector θ such that $\omega(A) = \langle \theta, \hat{A}\theta \rangle$.)
If $\omega(1) = 1$ one can say that $\omega(A)$ is the expectation value of A in the state ω

$\mathcal{H}$-pre Hilbert space
$\theta \in \mathcal{H}$ is a cyclic vector if a map $\mathcal{A} \to \mathcal{H}$ defined by the formula $A \to \hat{A}\theta$ is surjective.
The relation
$\langle \hat{A}\theta, \hat{B}\theta \rangle = \omega(A^*B)$
allows us to define pre Hilbert space $\mathcal{H}$ as the quotient of $\mathcal{A}$ with inner product
$\langle A, B \rangle = \omega(A^*B)$
with respect to null vectors

If the state is translation-invariant and stationary then time and spatial translations descend to Hilbert space; infinitesimal translations are interpreted as energy operator $\hat{H}$ and momentum operator $\mathbf{P}$. (Translations can be considered as automorphisms of $\mathcal{A}$; the action of translations on $\mathcal{H}$ is compatible with these automorphisms.)

Elements of $\mathcal{H}$ can be considered as excitations of translation-invariant stationary state ω (at least in the case when ω obeys cluster property).
Simplest form of cluster property
$\lim_{\mathbf{x}\to\infty}\omega(A(\mathbf{x},t)B)=\omega(A)\omega(B)$
Element σ of $\mathcal{H}$ has the form $B\theta$ where $B\in\mathcal{A}$.
Expectation value of A in the state $T_{\mathbf{a}}\sigma$ is proportional to
$\langle T_{\mathbf{a}}B\theta, AT_{\mathbf{a}}B\theta\rangle=\omega(B^*(\mathbf{a},0)AB(\mathbf{a},0))\approx \omega(B^*B)\omega(A)$

Elementary excitation of ω in algebraic approach is defined as a generalized vector function $\Phi(\mathbf{p}) \in \mathcal{H}$ obeying

$$\mathbf{P}\Phi(\mathbf{p}) = \mathbf{p}\Phi(\mathbf{p}), \hat{H}\Phi(\mathbf{p}) = \epsilon(\mathbf{p})\Phi(\mathbf{p}).$$

Here $\mathbf{p} \in \mathbb{R}^d$.
Introducing notation $\Phi(f) = \int d\mathbf{p} f(\mathbf{p})\Phi(\mathbf{p})$ we can consider an elementary excitation as a map of the space of test functions $\mathfrak{h}$ into $\mathcal{H}$.

Let us define the action of spatial and time translations in $\mathfrak{h}$ (in "elementary space") by the formulas $(T_{\mathbf{a}}f)(\mathbf{p}) = e^{i\mathbf{a}\mathbf{p}}f(\mathbf{p})$, $(T_\tau f)(\mathbf{p}) = e^{-i\tau\epsilon(\mathbf{p})}$. More generally, we can define an elementary space as the space of vector functions $f(\mathbf{p})$ where spatial translations are defined by the same formula and time translations commute with spatial translations.

We can identify an elementary excitation with a map of $\mathfrak{h}$ into the set of excitations; this map should commute with translations. The same definition works in the geometric approach. The only difference: in the algebraic approach the map is linear, in the geometric approach it is non-linear. (It is natural to assume that it is quadratic, more precisely, Hermitian).

If we are working in algebraic approach then the state $\sigma(f)$ corresponding to the vector $\Phi(f)$ is a functional

$$(\sigma(f))(A) = \langle \Phi(f), A\Phi(f) \rangle.$$

We assume that in algebraic approach

$$\Phi(f) = \hat{B}(f)\theta$$

where $B(f) \in \mathcal{A}$ *depends linearly on* $f \in \mathfrak{h}$. Then we can obtain the state $\sigma(f)$ acting on ω by the operator $L(f)$ transforming a linear functional $\rho \in \mathcal{A}^\vee$ into the functional $(L(f)\rho)(A) = \rho(B^*(f)AB(f))$.

This remark prompts the following definition of elementary excitation in the geometric approach: *We say that an elementary excitation of translation-invariant stationary state* ω *is specified by a map* σ *of an elementary space* $\mathfrak{h}$ *into the set of excitations of* ω *commuting with space and time translations. We assume that for* $f \in \mathfrak{h}$ *we have* $\sigma(f) = L(f)\omega$*. The operators* $L(f)$ *and* $L(g)$ *should almost commute if the supports of* f *and* g *in coordinate representation are far away.*

Two-particle state

$L(f)L(g)\omega$

where supports of f and g in coordinate representation are far away.

$(T_\tau f)(\mathbf{x}) = \int d\mathbf{p} e^{-i\epsilon(\mathbf{p})\tau + i\mathbf{p}\mathbf{x}} f(\mathbf{p})$

is small outside of the set τU_f where U_f is an open neighborhood of compact set containing all points of the form $\nabla f(\mathbf{p})$ with $f(\mathbf{p}) \neq 0$.

We say that τU_f is an essential support of $T_\tau f$ in coordinate representation.

If the sets U_f and U_g do not overlap we say that functions f and g do not overlap. In this case the essential supports τU_f and τU_g are far away for large $|\tau|$

If the theory came from algebraic approach then $L(f)$ can be obtained from $L(\tilde{f}, f)$ taking $\tilde{f} = f$. (Here $L(\tilde{f}, f)$ denotes an operator acting in $\mathcal{L}$; it should be linear with respect to f and anti-linear with respect to $\tilde{f}$.)

Elementary excitations of the ground state are called particles, in general case they should be called quasiparticles. Particles can be unstable (and quasiparticles are in general unstable). Saying that an elementary excitation is unstable we have in mind that the above relations are satisfied only approximately.

If particles have some discrete quantum numbers (like spin) or there are several kinds of particles we have several elementary excitations in the above terminology. (Alternatively one can define the space of test functions $f(\mathbf{p})$ as the space vector-valued functions on $\mathbb{R}^d$ where spatial translations act as multiplication: $(T_{\mathbf{a}}f)(\mathbf{p}) = e^{i\mathbf{ap}}f(\mathbf{p})$, and time translations commute with spatial translations. Then the same definition that we gave can serve as a definition of elementary excitation in the presence of discrete quantum numbers.)

To have a Lorentz-invariant theory we need an extension of the commutative group of translations to the Poincaré group. In this case we should assume that ω is Poincaré-invariant, then the representation of Poincaré group descends to $\mathcal{H}$. Irreducible subrepresentations of the representation of Poincaré group in the space $\mathcal{H}$ can be regarded as elementary excitations.

One can obtain the "elementary space" $\mathfrak{h}$ quantizing "elementary symplectic manifold" with Darboux coordinates $(\mathbf{q}, \mathbf{p})$ and space and time translations defined by formulas $T_{\mathbf{a}}(\mathbf{q}, \mathbf{p}) = (\mathbf{q} + \mathbf{a}, \mathbf{p})$, $T_\tau(\mathbf{q}, \mathbf{p}) = (\mathbf{q} + \mathbf{v}(\mathbf{p})\tau, \mathbf{p})$ where $\mathbf{v}(\mathbf{p}) = \nabla\epsilon(\mathbf{p})$. In translation-invariant classical theory, a map of "elementary symplectic manifold" into the space of excitations of translation-invariant solution of equations of motion specifies a family of solitons if it commutes with space and time translations.

Solitons and generalized solitons.

Consider a translation-invariant Hamiltonian in an infinite-dimensional phase space M consisting of vector-valued functions $f(\mathbf{x})$, where $\mathbf{x} \in \mathbb{R}^d$ are spatial coordinates. Spatial translations act as shifts of these coordinates, and time translations are governed by a Hamiltonian that is invariant with respect to spatial translations.

The equation of motion can be written as:

$$\frac{\partial f}{\partial t} = Af + B(f),$$

where A is a linear operator and B represents the nonlinear part. Assuming the nonlinear part is at least quadratic, for small f the linear part dominates. We can say that $f \equiv 0$ is a solution, and in its neighborhood, one can neglect the nonlinear part.

Soliton (solitary wave) is defined as a solution of the form $s(\mathbf{x} - \mathbf{v}t)$. We suppose that $s(\mathbf{x})$ tends to zero as $\mathbf{x} \to \infty$. We can visualize the solution $f \equiv 0$ as a horizontal straight line, and then the soliton is a bump moving with constant speed without changing the shape. A generalized soliton is a bump that moves, with a constant average speed, but at the same time it can pulsate, it can change its shape.

In Lorentz-invariant theory, by applying a Lorentz transformation to a soliton we again get a soliton. We obtain a family of solitons- solitons with different velocities. The same reasoning can be used for Galilean invariance and Galilean transformations. In both cases, we have a family of functions $s_{\mathbf{p}}(\mathbf{x}-\mathbf{a})$ that is invariant under temporal and spatial translations (here $\mathbf{p}$ denotes the momentum of soliton). This family can be considered as a symplectic manifold. A family of generalized solitons also can be considered a symplectic manifold that is invariant under temporal and spatial translations; the coordinates on this manifold are the data characterizing a (generalized) soliton.

We assume that the soliton has finite energy. (The fact that the energy is finite means, roughly speaking, that the soliton is more or less concentrated in some finite domain.)

In an old paper we conjectured that for many systems and for almost all initial conditions having finite energy the solution behaves in the following way for times tending to plus or minus infinity. If there are no solitons or generalized solitons in the theory then asymptotically the solution obeys a linear equation. In the general case, we get a few solitons plus something that approximately satisfies a linear equation (a tail).

This is a well-known result for integrable systems in the case $d = 1$; we have conjectured that this is true without the assumption of integrability in any dimension. Later this hypothesis has also been expressed in other papers. Soffer calls it "grand conjecture", Tao calls it "soliton resolution conjecture".

So far there are no results in this direction for $d > 1$ (and even for non-integrable theories in the case $d = 1$) if there exist solitons in the theory.

This conjecture can be justified by the following reasoning. Let us assume that the initial condition is a field concentrated in some domain. In this case, we should expect the spreading of wave packet. That is, if the initial data were concentrated in some domain, then later the solution spreads to a larger domain. The energy is conserved, so this spreading causes the amplitude of the wave to decrease. If the amplitude decreases all the time, then, as we assumed, in the case of small amplitudes the nonlinear part can be neglected, and the solution of the nonlinear equation can be approximated by a solution of a linear equation.

If there is a soliton or a generalized soliton in the theory the height of the bump remains the same, hence the amplitude does not tend to zero. However, we can expect that in the end, we get some solitons or generalized solitons plus a tail that approximately satisfies a linear equation. Of course, our reasoning is not proof, but it is convincing.

To prove the above conjecture one should impose some conditions. In particular, the stability of the translation-invariant state $f \equiv 0$ and of the solitons is necessary, otherwise, the solution can blow up. Nevertheless, it seems that the conjecture is true in many cases.

In these cases, there is a notion of soliton scattering. For solvable models of dimension 1+1 (one space dimension and one time dimension) this is a well-known fact. Two solitons collide, we see something that does not resemble any solitons (" a mess"), and then the same solitons appear again. The situation in the general case is slightly different: after the collision, we get some solitons (not necessarily the same solitons) plus a "tail". The tail asymptotically behaves as a solution of a linear equation.

Let us give some formal definitions. Let us denote the space of possible initial data by $\mathcal{R}$. Our conjecture means that for a dense set of initial data, we can define a mapping $D^+(t) : \mathcal{R} \to \mathcal{R}_{as}$ of initial data at the moment t to asymptotic data at $t \to +\infty$. (The asymptotic data characterize the solitons and the asymptotic behavior of the tail.). We can also consider the asymptotic data at $t \to -\infty$ to get a mapping $D^-(t) : \mathcal{R} \to \mathcal{R}_{as}$.

Now we assume that there is also an inverse mapping, i.e. one can find a solution with given asymptotic behavior. That is, we want to consider inverse operators $S(t, +\infty) = (D^+(t))^{-1}$ and $S(t, -\infty) = (D^-(t))^{-1}$.

In the quantum case, the solution to this problem is well known - it is what is called the Haag-Ruelle scattering theory; a generalization of this theory will be explained in the next lecture.

Now we can define the non-linear scattering matrix:

$$S = S(0, +\infty)^{-1} S(0, -\infty) : \mathcal{R}_{as} \to \mathcal{R}_{as}.$$

Roughly speaking, we fix asymptotic condition at minus infinity, solve the equation and watch the asymptotic behavior at plus infinity.

One should expect that we can get the non-linear scattering matrix from the quantum scattering matrix in the limit $\hbar \to 0$. (More precisely, one should expect that the inclusive scattering matrix has a limit as $\hbar \to 0$ and the non-linear scattering matrix can be expressed in terms of this limit.)

Classical soliton can be considered a model of a quantum particle. In quantum field theory, the notion of a particle is an asymptotic notion: if two particles collide, we get "a mess", which then disintegrates into particles.
Notice, that the analogy with solitons makes it obvious that the existence of identical particles is not surprising.

The following considerations further emphasize the analogy of solitons with quantum particles. Consider a phase space and a Hamiltonian; in other words, we consider a symplectic manifold $\mathcal{M}$ (that can be identified with the space $\mathcal{R}$ of initial data) and an evolution operator. Assume that spatial translations act on $\mathcal{M}$ and time translations commute with spatial translations. Formally this means that on the symplectic manifold $\mathcal{M}$ we have an action of the commutative group $\mathcal{T}$ of spatial and temporal translations. Now let us take a stationary translation-invariant point $m \in \mathcal{M}$ of this symplectic manifold.

In the previous picture, such a point was the solution $f \equiv 0$.
Let us define an excitation of a translation-invariant stationary state as a state with finite energy (we assume that the energy of a translation-invariant state is equal to zero).
We define an elementary symplectic manifold $\mathcal{E}$ as such a symplectic manifold where in Darboux coordinates $\mathbf{p}, \mathbf{x}$ the spatial translations act as shifts $\mathbf{x} \to \mathbf{x} + \mathbf{a}$, while $\mathbf{p}$ does not change. We consider a Hamiltonian $\epsilon(\mathbf{p})$ that depends only on $\mathbf{p}$ (i.e. it is invariant with respect to spatial translations). Then the time translations are transformations $\mathbf{x} \to \mathbf{x} + \mathbf{v}(\mathbf{p})t, \mathbf{p} \to \mathbf{p}$, where $\mathbf{v}(\mathbf{p}) = \nabla \epsilon(\mathbf{p})$.

Suppose now that $\mathcal{M}$ is realized as a space of vector-valued functions $f(\mathbf{x})$ where $\mathbf{x} \in \mathbb{R}^d$ and the spatial translations act as shifts $\mathbf{x} \to \mathbf{x} + \mathbf{a}$. Let us take a symplectic embedding of the elementary symplectic space $\mathcal{E}$ into the set of excitations of translation-invariant state $f(\mathbf{x}) = const$ in $\mathcal{M}$. If this embedding commutes with the space-time translations, then we get a family of solitons.

To verify this we notice that symplectic embedding maps the point $(\mathbf{p}, 0)$ into some function $s_{\mathbf{p}}(\mathbf{x})$ depending on $\mathbf{p}$. Since the embedding $\mathcal{E} \to \mathcal{M}$ commutes with spatial translations, the point $(\mathbf{p}, \mathbf{a})$ maps into a shifted function $s_{\mathbf{p}}(\mathbf{x} + \mathbf{a})$. The condition that the mapping $\mathcal{E} \to \mathcal{M}$ commutes with time shifts means that the function $s_{\mathbf{p}}(\mathbf{x} - \mathbf{v}(\mathbf{p})t)$ satisfies the equation of motion.