Solitons

Anzor khelashvili

Institute of High Energy Physics, Iv. Javakhishvili Tbilisi State University, University Str. 9, 0109,

Tbilisi, Georgia and St. Andrea the First-called Georgian University of Patriarchy of Georgia,

Chavchavadze Ave. 53a, 0162, Tbilisi, Georgia

Preface

Solitons are the solutions of classical non-linear equations. They are analogous to extended classical particles. They became popular in early 1970-s.The theory of solitons is attractive and exciting, it brings together many branches of mathematics and theoretical physics.

Now there are a great number of books and review articles devoted to soliton problems. In preparation of these lectures the Author made use only part of them, the list of which is given below. They are:

1. P.G. Drazin and R.S. Johnson, "*Solitons: an introduction",* Cambridge univ. Press.1988.
2. A.T. Filippov, *"The Versatile solitons",* the Quant library, Moscow, 1986.
3. A.Kasman, *"Glimpses of Soliton Theory", AMS*, 2010.
4. T.Dauxois and M. Peyrard, "*Physics of Solitons",* Cambridge univ. Press, *2006.*
5. Lui Lam (Edit.), "*Introduction to nonlinear physics"*, Springer, 1996.
6. S. Nettel, "*Wave Physics –Oscillations-Solitons-Chaos",* Springer, 1995, 2003.
7. T.Miwa, M.Jimbo, E. Date."*Solitons",* Cambridge Tracts in mathematics, 135, 2000.
8. A.C.Newell, "*Solitons in Mathematics and Physics",* Univ. of Arizona, 1985.
9. С.П.Новиков (Edit.) . "*Теория Солитонов, Метод обратной задачи"* ,Наука.1980.
10. S.Leibovich and R.Seebass (Edit.),*"Nonlinear waves",* 1974.
11. Г.М.Заславский, Р.Зю Сагдеев, *Введение в нелинейную физику* ,Наука, 1988.
12. R.Rajaraman, *"Solitons and Instantons"*, Bangalore, 1982.
13. V. Rubakov, *"Classical Gauge Fields",*
14. В.Рубаков,"*Классические калибровочные поля",* Москва, 1999.
15. N. Manton, P. Sutcliffe, *"Topological Solitons",* Cambridge monographs, 2007.
16. Y.Yang, "*Solitons in Field theory and Nonlinear Analysis",* Springer, 2000.
17. E.J. Weinberg,*"Classical solutions in quantum field theory",* Cambridge monographs, 2012.

Review Articles:

18. Д.И.Дъяконов, В.Ю. Петров. "Барион как солитон", Лияф.Ленинпрад. 1984.

19. S.Gasiorowicz, D.A.Geffen, *"Effective Lagrangians and field theory with chiral Symmetry",* Rev. Mod. Physics, V.41, #3, 531. 1969.
20. I. Zahed, G.E. Brown, *"The Skyrme model" (*Phys. Reports,) 1986.
21. D.Holzwarth, R. Schwesinger, "*Baryons in the Skyrme model*", Siegen Univ. 1985.

ACKNOWLEDGEMENTS:

The author was very lucky to be acquainted with materials about solitons in various subjects of physics from the beautiful books and review articles, listed above. They were so impressionable that in many places of these lectures the author included various fragments without any modifications. Therefore the author impress them his great gratitude.

The author thanks also his colleagues for reading a preliminary versions of lectures.

Content

Part I

Part I

Lecture 1

1. Introduction (What is a Soliton)

Solitons are robust, localized traveling waves of permanent form. They are found everywhere. They exist in the sky as density waves in spiral galaxies, and giant Red Spot in the atmosphere of Jupiter; they exist in the ocean as waves bombarding oil wells; they exist in much smaller natural and laboratory systems such as plasmas, molecular systems, laser pulses propagating in solids, superfluid He^3 , fluid flow phase transitions, liquid crystals, polymers, and fluid flows, as well as elementary particles. They may even have something to do with high temperature T_C superconductors. So, what is soliton?

Solitons are special nonsingular solutions of some nonlinear partial differential equations (PDEs):

(i) they are spatially localized;
(ii) a single soliton is a travelling wave (i.e. it is a wave of permanent form);
(iii) they are stable;
(iv) When a single soliton collides with another one, both of them retain their identities after collision – like the elastic collision of two particles.

However in many systems the fourth property cannot hold. It turns out that this nice but stringent elastic collision property is intimately related to a specific property of a system, which is called integrability. We therefore differentiate two kinds of systems; namely, integrable and nonintegrable systems.

Note that many mathematicians insist that the name "soliton" should be reserved for those wave solutions that possess simultaneously all four properties listed above. This is not true for most physicists. All pragmatic physicists have to deal with the real world. There are simply too many real physical systems in nature that are nonintegrable. For these systems the concept of soliton even without (iv)th property is found to be so useful and fruitful that one cannot afford not to use it. The word "soliton" is used so loosely these days that sometimes not even properties (ii) and (iii) are retained.

Hydrodynamic solitons are dynamical structures. They move with a constant speed and shape, but they cannot exist at rest. On both sides of the soliton the state of the medium is the same. They are called *nontopological solitons in* contrast to another class of solitons, which interpolates between two different states of a medium, and can exist at rest.

Solitons are solutions of the classical field equations, which, in their own right, without quantization, are similar to particles. They are lumps of fields (energy) of finite size. More precisely, the fields decrease rapidly from the center of a lump. The existence and stability of solitons is due, in the first place, to the nonlinearity of the field equations. In quantum theory solitons correspond to extended particles, which, roughly speaking, are composed of the elementary particles in each specific model.

Among various types of solitons, the class of *topological solitons* is of particular interest. In particle physics the use of the soliton concept is rather limited, although it is sometimes very fruitful. At the same time solitons are to be found very frequently in condensed matter physics.

Brief history of solitons

Historically, soliton was first observed by John Scott Russell in 1834 on the Edinburg-Glasgow channel. He called it the "Great wave of translation". J. Russell reported his observations to the British Association in his 1944 "Report on waves" in the following words:

"I believe I shall best introduce the phenomenon by describing the circumstances of my own first acquaintance with it. I was observing the motion of a boat which was rapidly drawn along a narrow channel by a pair of horses, when the boat suddenly stopped – not so the mass of water in the channel which it had put in motion, it accumulated round the prow of the vessel in a state of violent agitation, the suddenly leaving it behind, rolled forward with great velocity, assuming the form of a large solitary elevation, a pounded, smooth and well-defined heap of water, which continued its course along the channel apparently without change of form or diminution of speed. I followed it on horseback, and overtook it still rolling on at a rate of some eight or nine miles an hour, preserving its original figure some thirty feet long and a foot to a foot and half in height. Its height gradually diminished, and after a chase of one or two miles I lost it in the windings of the channel.

In other words, J. Russell saw a hump of water created by a boat on the canal and followed it for several miles. Certainly, other people had seen such waves before since the circumstances that created it were not particularly unusual. But, it may be that nobody before gave it such careful thought.

The point is that the wave he saw did not do what you might expect. From our experiences with waves in a bathtub or at the beach, you might expect a moving hump of water to either:

- Get wider and shallower and quickly disappear into tiny ripples as we see with a wave that you might generate with your hand in a swimming pool
- Or "break" like the waves at the beach, with the peak becoming pointy, racing ahead of the rest of the wave until it has nothing left to support it and comes crashing down.

It was therefore of great interest to Russell that the wave he was watching did neither these things, but basically kept its shape and speed as it travelled down the canal unchanged for miles.

J. Russell called it "the Wave of Translation" and later the "Great solitary wave". Russell was so excited by this "singular and beautiful" phenomenon that he tried to explain many things in the inverse with it (which turned out to be wrong), but more importantly, being a good engineer, Russell went on and did experiments, recreating these "great solitary waves" in his laboratory dropping a weight at one of the end of water channel. (Fig.1). By "solitary wave" he was clearly referring to the fact that this wave has only a single hump, unlike the more familiar repeating sine wave pattern that one might first imagine upon hearing the word "wave". As for "wave of translation", it may be that he was referring to the question of whether the individual molecules of water were moving along with the hump or merely moving up and down, but, that is not how the term is generally used in soliton theory today not how we will use it in practice. To us, "translation" refers to the fact that *the profile of the wave –the shape it* *has when viewed from the side – stays the same as time passes, as if it was a cardboard cutout that was merely being pulled along rather than something whose shape could vary moment to moment.*

To study his solitary waves, Russell built a 30 foot long wave tank in his back garden. He found that he could reliably produce them in his tank and study them experimentally (Fig.1). Among the most interesting things he discovered was that there was a mathematical relationship between the height of the wave (a), the depth of the water when at rest (h) and the speed at which the wave travels (c) (Fig.2). He believed that this phenomenon would be of great importance and so reported on it to the British Association for Advancement of Science.

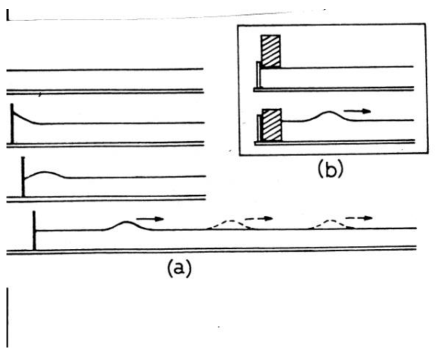


Fig.1. **Two ways of generating solitons in a tank of shallow water:**
(a) and (b) in experiments of J. Russell (1844)

Russell was able to deduce empirically that *the volume of water in the wave is equal to the volume of water displaced and further that the speed c of the solitary wave is obtained from*

$$c^2 = g\left(h + a\right)$$

where a is the amplitude of the wave, h - the undisturbed depth of water and g - the acceleration of gravity (Fig.2).

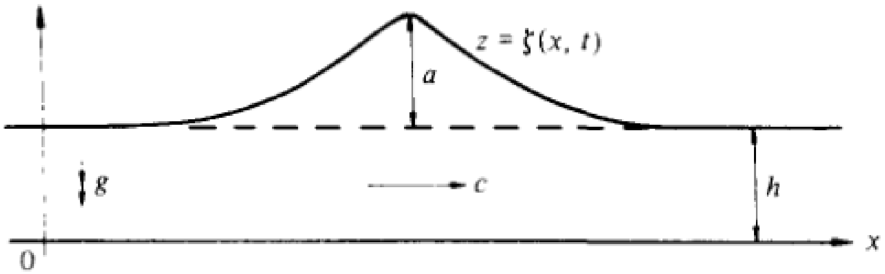


Fig.2. The parameters and variables in the description of the solitary wave

The solitary wave is therefore a *gravity wave.* We note immediately an important consequence of this equation: higher waves travel faster.

To put Russell's formula on a firmer footing both Boussinesq (1871) and Lord Rayleigh (1876) assumed that a solitary wave has a length much greater than the depth of the water. They deduce from the equations of motion for an inviscid incompressible fluid Russell's formula for c. In fact they also showed that the wave profile $z = \varsigma(x,t)$ is given by

$$\varsigma(x,t) = a\,\mathrm{sech}^2\left[\beta(x-ct)\right],$$

where

$$\beta^{-1} = 4h^2(h+a)/3a \ , \quad \text{for any} \quad a > 0$$

Although the $\sec h^2$ profile is strictly only correct, if $a/h \ll 1$.

These authors did not, however, write down a simple equation for $\varsigma(x,t)$, which admit abovementioned profile as a solution. This final step was completed by Korteweg and de Vries in 1895. We'll give this equation below and show that the Russell's solitary wave is a solution of the Korteweg and de Vries (KdV) equation. Moreover, we'll see that the "width" of the wave is proportional to $\sqrt{a}$. In other words, taller waves travel faster and are narrower.

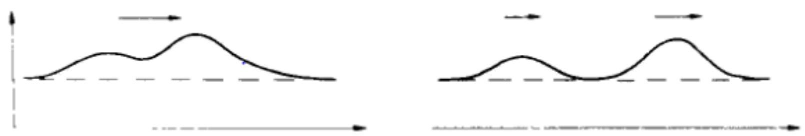


Fig.3. Lower wave lags behind

Let us take attention on a property (iv) of two solitons collision. When two solitons collide after collision the two solitons separate from each other with original shapes and velocities as before collision, but with a phase shift (Fig. 4)

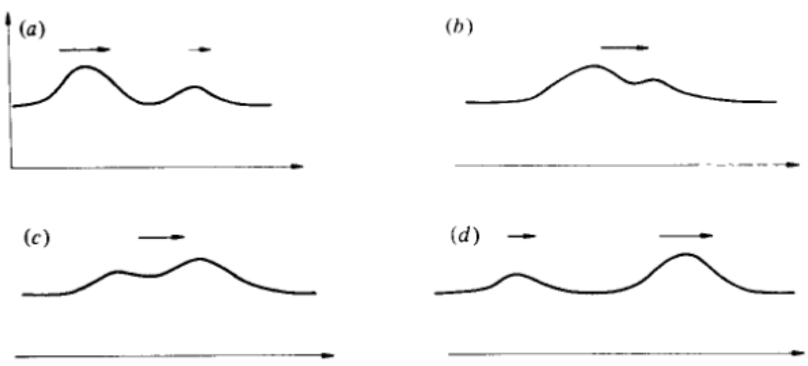


Fig.4. Interaction of two solitary waves

One interpretation of this result is that an arbitrary initial profile (which, in other words, is not an exact solitary wave) will evolve into two (or more) waves which then move apart and progressively approach individual solitary waves as $t \to \infty$ (our solitary wave is defined on $(-\infty, +\infty)$). This alone is rather surprising, but another remarkable property can also be observed. If we start with an initial profile like that given in Fig. 4a, but with the taller wave somewhat to the left of the shorter, then the development is as depicted in Fig.4b.

In this case the taller wave catches up, interacts with and then passes the shorter one. The taller one therefore appears to overtake the shorter one and continue on its way intact and undistorted. This, of course, is what we would expect *if the two waves were to satisfy the linear superposition principle*. But they certainly do not. This suggests that we have a special type of nonlinear process at work here (In fact, the only indication that a linear interaction has not occurred is that the two

waves are phase-shifted, i.e. they are not in the positions after interaction which would be anticipated if each were to move at a constant speed throughout the collision).

Surprising in such collision experiments is that after very long time, the initial profile –or something very close to it - reappears, a phenomenon requiring topology of the torus for its explanation. At the heart of those observations is the discovery that these (nonlinear) waves can interact strongly and then continue thereafter almost as if there had been no interaction at all. These persistence of the wave led Zabusky and Kruskal to coin the name "soliton" (after photon, proton, electron etc.)

We emphasize the particle-like character of these waves which seem to retain their identities in a collision. Owing to this particle-like (iv) property, it is expected that solitons may have a broad application in particle physics, as particles with finite sizes. In this concern the interest to solitons in particle physics was considerable grown after 60-ies of previous century. However, as studies of principal properties of solitons took their origin from the travelling wave observation on a water, we are not able go by description of solitons in ordinary water surfaces.

Therefore, first of all, we consider an example of Korteweg and de Vries model and demonstrate how the non-linearity and dispersion compensate each other's and provide stability of solitary wave. It is not excluded that something like happens in non-linear field theory models of particles while the direct and transparent physical analogy is difficult to obtain (the fields are represented as a Fourier series of various frequencies and the wave packets are also expanded according to frequencies (dispersion), which must be balanced by non-linearity. We'll see that in some models more profound compensation can be achieved thanks to topological reasons).

2. The Korteweg and de Vries (KdV) Equation

(a) ***– general characteristic of wave, its dispersion and dissipation***

Waves play a significant role in nature. There are mechanical waves, seismic waves in air, water waves. There are electromagnetic waves, and underlying all matter, quantum mechanical waves. These deserve wave phenomena are understood on the basis of a few unifying mathematical conceptions. In all these areas it is common practice to develop the concept of wave propagation from the simplest albeit idealized –model for one-dimensional motion

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0 , \tag{1}$$

where $u(x,t)$ is the amplitude of the wave and c is a positive constant. This equation has a simple and well-known general solution, expressed in terms of characteristic variables $(x \pm ct)$ as

$$u(x,t) = f(x-ct) + g(x+ct) , \tag{2}$$

where f and g are arbitrary functions (*<u>comment:</u>* t is a time coordinate and x - spatial coordinate, although here they are interchangeable since they differ only by "scaling" factor c).

The functions f and g (not necessarily differentiable) can be determined from, for example, initial data $u(x,0)$, $\frac{\partial u}{\partial t}(x,0)$. The solution (2), usually referred *as d'Alembert solution*, describes two distinct waves, one moving to the left and one to the right direction correspondingly (both at the speed c . The waves do not interact with themselves and not with each other. This is a consequence of *linearity* of Eq. (1) and hence these solutions may be added (or superposed).

Furthermore, the waves described by (2) do not change their shape as they propagate. This is easily verified if we consider one of the components – say f and choose a new coordinate which is moving with this wave, $\xi = x - ct$. Then $f = f(\xi)$ and it does not change as x and t change, for fixed ξ . In other words shape given by $f(x)$ at $t=0$ is exactly the same at later times but shifted to the right by an amount ct .

Let us restrict ourselves to waves which propagate in only one direction. This is allowable choice in solution (2) merely set $g=0$, for example. f and g will move apart and no longer overlap, since they never interact, one can now follow one of them and ignore the other. To be more specific, we may restrict the discussion to the solution of

$$u_t + cu_x = 0 \tag{3}$$

The general solution of (3) is

$$u(x,t) = f(x-ct)$$

We may set $c=1$. Then if $u_t+u_x=0$, we obtain $u(x,t)=f(x-t)$.

We may also retain the connection with Eq. (1). The d'Alembert operator can be factorized, and either factor may be zero

$$\left(\frac{\partial}{\partial t}\mp c\frac{\partial}{\partial x}\right)\left(\frac{\partial}{\partial t}\pm c\frac{\partial}{\partial x}\right)u=\left(\frac{\partial^2}{\partial t^2}-c^2\frac{\partial^2}{\partial x^2}\right)u=0$$

In general, when wave equations are derived from some underlying physical principles, certain simplified assumptions are made: in extreme cases we might derive Eqs. (1) or (3). However, if these assumptions are less extreme, we might obtain equations which retain more of the physical detail, for example, wave dispersion or dissipation or nonlinearity.

Consider first the equation

$$u_t+u_x+u_{xxx}=0, \tag{4}$$

which is simplest dispersive wave equation. To see this let us examine the form of harmonic wave solution:

$$u(x,t)=e^{i(kx-\omega t)} \tag{5}$$

Now (5) is a solution of (4) if

$$\omega=k-k^3 \tag{6}$$

It is the dispersion relation which determines $\omega=\omega(k)$ for a given k. Here k is the wave number, taken to be real. The solution (5) is certainly oscillatory at $t=0$, and so is the frequency. From (5) we see for the phase

$$kx-\omega t=k\left\{x-\left(1-k^2\right)t\right\}$$

and the solution (5) with condition (6) describes a wave propagating at the velocity

$$c=\frac{\omega}{k}=1-k^2,$$

which is a function of k. Waves with different wave numbers propagate at different velocities. It is characteristic of a dispersive wave, thus a single wave profile (5) which can be (suppose) by the sum of just two components, each like (5), will change its shape as time evolves by virtue of the different velocities of the two components.

To extend this idea we need only add as many components as we desire, or, for greater generality, integrate over all k to yield

$$u(x,t)=\int_{-\infty}^{\infty} A(k)\exp i\left(kx-\omega(k)t\right)dk \tag{8}$$

$A(k)$ is essentially the Fourier transform of $u(x,t)$.The overall effect is to produce a wave profile which changes its shape as it moves, in fact, since the different components travel at different velocities. The profile will necessarily *spread out or disperse.*

Velocity from Eq. (6) is usually termed the phase velocity. Another velocity is the *group velocity* defined by

$$\upsilon_g=\frac{d\omega}{dk}=1-3k^2$$

It determines the velocity of a wave packet

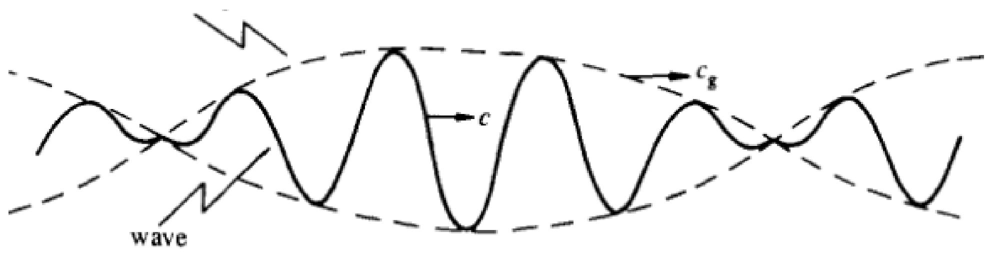


Fig. 5. The sketch of a wave packet, showing the wave and its envelope. The wave moves at the phase velocity, C and the envelope at the group velocity, c_g

The group velocity is the velocity of propagation of energy.

Our assumption that $\omega(k)$ is real for all k remains true only if add to Eq. (4) odd derivative of u. If we choose to use even derivatives, taking for example

$$u_t+u_x-u_{xx}=0 \tag{9}$$

then the picture is quite different. From Est. (5) and (9) we obtain

$$\omega=k-ik^2$$

and therefore

$$u(x,t)=\exp\left\{-k^2t+ik(x-t)\right\} \tag{10}$$

is a solution of Eq. (9).

This describes a wave which propagates at a speed of unity for all k, but which also decays exponentially for any real $k\neq 0$ as $t\to+\infty$. The decay exhibited in Eq. (10) is usually called *dissipation*.

Clearly we could have Eqs. like (4) or (9) which incorporate linear combinations of even and odd derivatives. In this case the harmonic wave solution may be both *dissipative and dispersive.*

(b) *Nonlinearity*

Finally, let us briefly look at one rather more involved aspect of wave motion, namely that of *nonlinearity* in particular, for a nonlinear equation without dispersion and dissipation such as

$$u_t + uu_x = 0, \tag{11}$$

one has the formal solution $u = f\left(x - ct\right)$ with "velocity" $u = c$. So, different points of a pulse then travel with different velocities proportional to their heights (Fig.6), resulting in a squeeze of the pulse width as it travels.

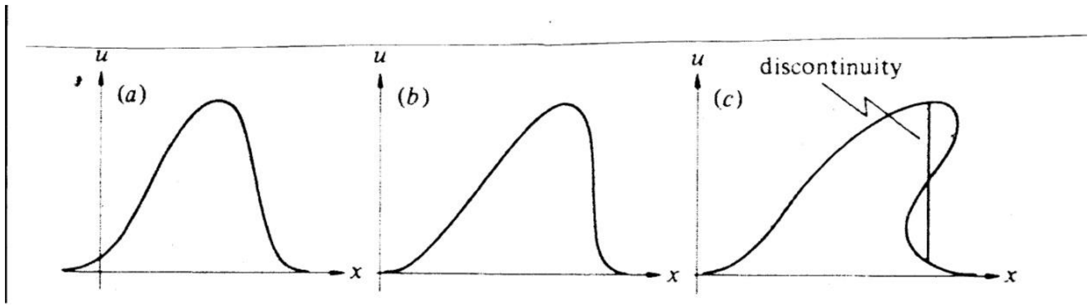


Fig.6 **Evolution of a nonlinear wave as time increases: at:** $(a)\ t = t_1;\ (b)\ t = t_2 > t_1\ \ (c)\ t = t_3 > t_2$

The solution obtained by construction exhibits the non-uniqueness as a wave which has a "broken" (Fig.6) – the solution must necessarily change its shape as it propagates. This difficulty is usually overcome by the insertion of a jump (or discontinuity) which models a shock. Strictly, a discontinuous solution is not a proper solution of Eq. (11).

Another complication arises with nonlinear equations: We have a superposition principle for linear equations. Namely, any linear combination of two solutions u_1 and u_2 is also a solution. However this is not true, in general, for nonlinear equations. The solutions of nonlinear equations cannot superpose to form new solutions, although a related principle is available for certain nonlinear partial differential equations.

For some nonlinear equations with dispersion it is then possible that this squeeze of the profile width due to nonlinearity is balance exactly by the expansion of the width due to dispersion leading to a travelling wave of permanent shape, i.e. soliton.

Lecture 2

Beginning with the extremely accurate but unwieldy Navier-Stokes equations, Korteweg and de Vries made some simplifying assumptions including a sufficiently narrow body of water so that the wave can be described with only one spatial variable and constant, shallow depth as one would find in a canal. Putting all of this together, they settled on the equation

$$u_t = \frac{3}{2} u u_x + \frac{1}{4} u_{xxx}, \tag{12}$$

Due to their initials, this famous equation is now known as the "KdV equation". We will see that in general KdV equation incorporates both nonlinearity and dispersion.

It was by making use of results from the area of "pure mathematics" that they were able to derive a large family of solutions to this equation which translate and maintain their shape. Among these solutions were the functions

$$u_{sol(k)}(x,t) = \frac{8k^2}{\left(e^{kx+k^3t} + e^{-kx-k^3t}\right)^2} , \tag{13}$$

Which satisfy the KdV equation for *any* value of the constant k. This formula gives a translating solitary wave, like Russell's, that travels at speed k^2 and has a height $2k^2$.

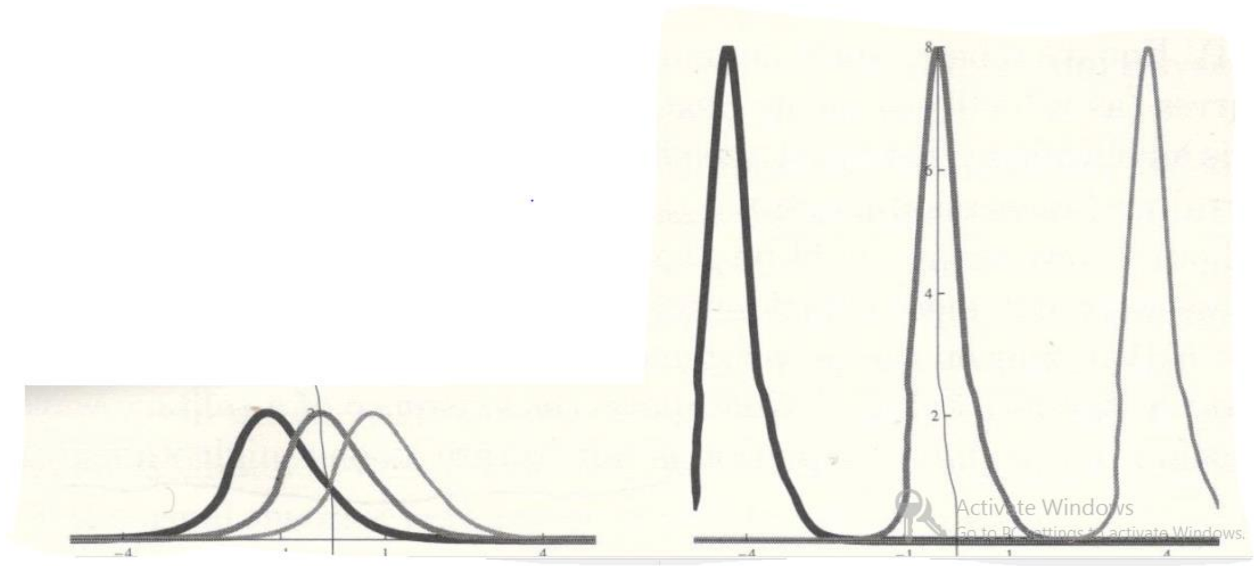


Fig.7. **Two solitary wave solutions of the form (13) to the KdV Equation (12). The figure on the left shows the solution with k=1 and the right is k=2. In each case, the figure illustrates the solution at times t=-1, t=0 and t=1. Note that the speed with which the wave translates is k^2 and that the height is twice the speed.**

Here the solutions $u_{sol(1)}(x,t)$ and $u_{Sol(2)}(x,t)$ are compared side-by-side. Note that in each case the height of the wave is twice its speed.

The u_{xxx} term as we saw previously, resulted there in separation of the different frequency components of a "single-humped" initial profile, resulting in a dispersion. More dramatically, the uu_x term induced a non-linear distortion, which soon destroyed any "single-humped" initial profile. However, somehow, the combination of these two terms seems to avoid both of these problems.

Specifically, the fact that the solutions could be written explicitly was a consequence of the *coincidence* that the KdV equation bears some similarity to an equation related to elliptic curves – and, one might say that it is a coincidence here that effects of distortion and dispersion are perfectly balanced so that they cancel out. However, it would be a long time before anyone realized that these were not mere coincidence. Something interesting also happens when one views solutions that just appear to combine two different solitary waves. For these solutions there are two humps each moving to the left with speed equal to half their height. As will be seen, it is not the case that this is simply a sum of two of the solitary wave solutions found by Korteweg and de Vries. If the taller of two humps is on the left, then they simply move apart. The amazing thing, however, it to consider the situation in which a taller hump is to the right of a shorter one. Since it is moving to the left at a greater speed it will eventually catch up.

(c) *Kruskal and Zabusky numerical experiment*

Intuition about nonlinear differential equations would have suggested to any expert at the time that even though the KdV equation has this remarkable property of having solitary wave solutions, when two solitary waves come together, the result would be a mess. One would expect that whatever coincidence allows them to exist in isolation would be destroyed by the overlap and that the future dynamics of the solution would not resemble solitary waves at all. However, the numerical experiments of Kruskal and Zabusky showed the humped shapes *surviving* the "collision" and seemingly separating again into two separate solitary waves translating left at speeds equal to the half their heights! Moreover the same phenomenon could be seen to occur when three or more separate peaks were combined to form an initial profile: the peaks would move at appropriate speeds, briefly "collide: and separate again.

More specifically, we now refer to the solitary wave solutions as 1-soliton solutions of the KdV equation. In general, an n-soliton solution of the KdV equation has n separate peaks. **Fig. 8** illustrates a 2-soliton solution of the KdV equation, in which a taller soliton traveling at speed 4 catches up to a shorter one with speed 1. Briefly, at time $t = 1.0$ we cannot see two separate peaks, but later again they separate so that we can clearly see a soliton of height 2 and another of height 8. However, you should not mistakenly think that this is the same as two 1-solitons viewed together.

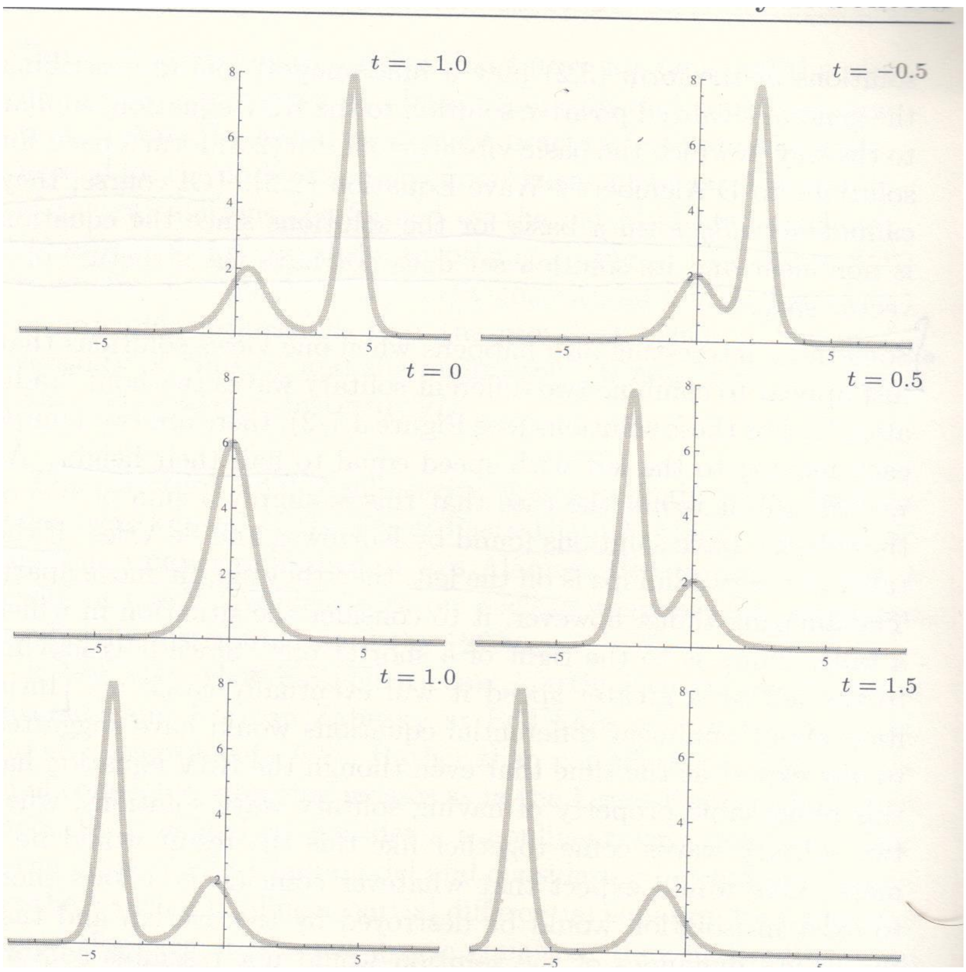


Fig.8 ***A solution to the KdV equation as it would have appeared to Kruskal and Zabulsky in their numerical experiments. Note that two humps, each looking like a solitary wave, come together and then separate.***

It is easily seen in the following figure, which illustrates the combination of two one soliton solutions.

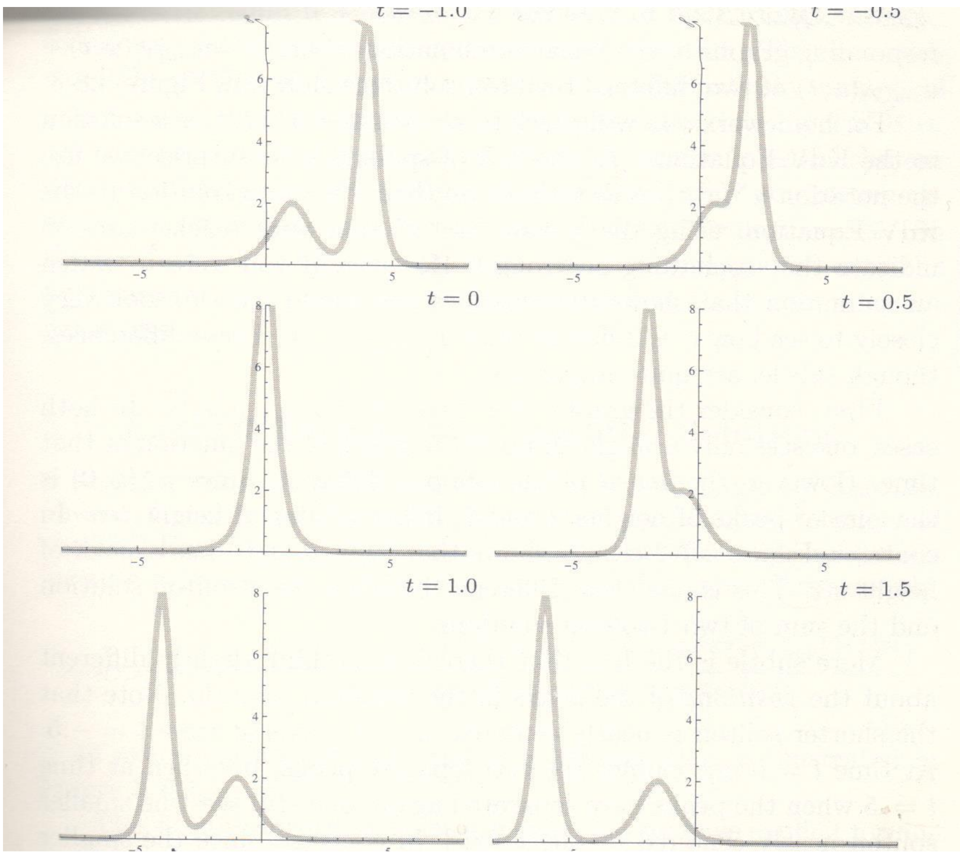


Fig.9 **This is not a solution to the KdV equation! This is a sum of the one soliton solutions $u_{sol(1)}(x,t)$ and $u_{sol(2)}(x,t)$. Compare to previous figure, which is a KdV solution, to see the subtle differences despite that each shows a hump moving to the left at speeds 1 and 4, respectively, at most times and a single hump centered on the x-axis at time $t=0$.**

For now it is enough to think of it as an indication that there is some sort of nonlinear *interaction* going on in the 2-soliton solution. If we think of the solitons as particles, then they have not simply passed through each other without any effect, but have actually "collided" and in some sense the KdV equation incorporates their "bounce".

Lecture 3

(d) ***Explicit solution of the KdV equation***

A nonlinear equation will normally determine a restricted class of profiles which often play an important role in the solution of the initial value problem.

It is clear that, by making suitable assumptions in a given physical problem, we might obtain an equation which is both nonlinear and contains dispersive or dissipative terms (or both). So, for example, we might derive

$$u_t + (1+u)u_x + u_{xxx} = 0, \qquad or \quad u_t + (1+u)u_x - u_{xx} = 0$$

The first of these is the simplest equation embodying nonlinearity and dispersion; this, or one of its variants, is known as the KdV equation, of which we shall say much more lately. The second equation, with nonlinearity and dissipation, is the *Burgers equation.*

Our main concern will be with the method of solution of the KdV equation. However, before upon a more detailed discussion, we mention the various alternative forms of this equation. We can transform this one under

$$1+u \to \alpha u, \qquad t \to \beta t, \qquad x \to \gamma x,$$

where α, β, γ are non-zero real constants, to yield

$$u_t + \frac{\alpha\beta}{\gamma} u u_x + \frac{\beta}{\gamma^3} u_{xxx} = 0$$

This is a general form of the KdV equation, and a convenient choice, which are often used, is

$$u_t - 6uu_x + u_{xxx} = 0$$

This form is *invariant* under the continuous group of transformations

$$X = kx, \qquad T = k^3 t, \qquad U = k^{-2} u$$

After these transformations the above KdV equation becomes

$$U_T - 6UU_X + U_{XXX} = 0$$

These transformations with $k \neq 0$ form an *infinite group,* where k is a parameter of this continuous group.

(e) *Solitary waves*

Now turn to the KdV equation in the standard form

$$u_t - 6uu_x + u_{xxx} = 0 \tag{13}$$

The travelling wave solutions of this equation are

$$u(x,t) = f(\xi) \qquad where \quad \xi = x - ct \quad and \quad c = const$$

Thus Eq. (13) becomes

$$-cf' - 6ff' + f''' = 0$$

It may be integrated once to yield

$$-cf - 3f^2 + f'' = A,$$

where A is an arbitrary constant. If we now use f' as an integrating factor we may integrate once more to give

$$\frac{1}{2}(f')^2 = f^3 + \frac{1}{2}cf^2 + Af + B, \qquad (14)$$

where B is a second arbitrary constant. At this stage let us impose the boundary conditions

$$f, f', f'' \to 0, \qquad as \;\; \xi \to \pm\infty,$$

which describe the solitary wave. Thus A and B are both zero, and it remains

$$(f')^2 = f^2(2f + c) \qquad (15)$$

Now we can see immediately that a real solution exists only if $(f')^2 \geq 0, \quad i.e.\ (2f + c) \geq 0$.

Eq. (15) can be integrated as follows: first write

$$\int \frac{df}{f(2f+c)^{1/2}} = \pm \int d\xi$$

Then use the substitution

$$f = -\frac{1}{2}c\,\text{sech}^2\,\theta, \qquad (c \geq 0)$$

we have obtained

$$f(x - ct) = -\frac{1}{2}c\,\text{sech}^2\left\{\frac{\sqrt{c}}{2}(x - ct + x_0)\right\} \qquad (16)$$

where x_0 is an arbitrary constant of integration. Note that the choice $\pm$ is redundant since the solution is even function and also the constant x_0 (a phase shift) plays a minor role. It merely denotes the position of the peak at $t = 0$.

Note that the speed with which the wave translates is c and the height of the wave is twice its speed?

The solitary wave solution (16) forms a one parameter family (ignoring x_0) and in fact the solution exists for all $c \geq 0$ no matter how large or small the wave may be.

In fact that $f \leq 0$ reflects our choice of KdV equation (13) with negative nonlinearity, we may recover the classical wave of elevation by transforming: $u \to -u$ $u \to -u$.

The solution is negative (see, Fig.10), because it is determined by the sign in front of the nonlinear term.

The solution (16) shows that solitary waves propagate to the right with a velocity c , which is proportional to the square root from the amplitude.

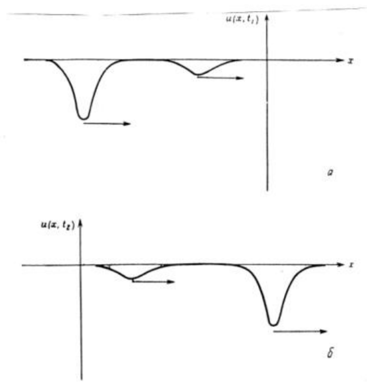


Fig. 10 Soliton interaction $(t_2 > t_1)$

One of the interesting properties of this solution is a "linear" behavior of solitary waves. While superposition of solutions of the nonlinear equations do not lead to the new solutions, but the calculations of Zabuski and Kruskal have shown that two solitary waves with different amplitudes after nonlinear interaction remain immutable. Exactly this property, analogy with particles, gave rise the name "soliton". In particular, let us consider two separated from each other solitons, moreover a large one is on the left from a little one. Because taller soliton has larger speed, it catches up with smaller soliton and after their nonlinear interaction according to the KdV equation, they remain immutable. Nonlinearity of such solutions consists in that the solitons become shifted in compare to their locations they were without interaction.

In the context of the KdV equation, and other similar equations, it is usual to refer to the single-soliton solution as the *solitary wave,* but when more than one of them appears in a solution they are called *solitons.* Another way of expressing this is to say that the soliton becomes a solitary wave when it is infinitely separated from any other soliton.

Also, we must mention the fact that for equations other than the KdV equation the solitary-wave solution may not be a sech^2 function; for example, we shall meet a such function and also $\arctan(e^{2x})$. Furthermore, some nonlinear systems have solitary waves but ***not*** solitons, whereas others (like KdV equation) have solitary waves which are solitons.

.

Exercise: **more explicit calculation of the integral**

We made a substitution $f = -\frac{1}{2}c \cdot \mathrm{sech}^2\,\theta$

Then $df = f' d\theta$ and

$\left(\sec h\theta\right)' = \sec h\theta \cdot \tanh\theta d\theta$

Therefore $df = c \cdot \mathrm{sech}^2\,\theta \cdot \tanh\theta d\theta$

On the other hand $2f + c = -c\,\mathrm{sech}^2\,\theta + c = c\left(1 - \sec h^2\theta\right) = c\tanh^2\theta$

Thus $f\left(2f + c\right)^{1/2} = -\frac{c\sqrt{c}}{2}\mathrm{sech}^2\,\theta \cdot \tanh\theta$

Finally

$$\pm\left(\xi + \xi_0\right) = \int \frac{c \cdot \mathrm{sech}^2\,\theta \cdot \tanh\theta \cdot d\theta}{-\frac{c\sqrt{c}}{2}\mathrm{sech}^2\,\theta \cdot \tanh\theta} = -\frac{2}{\sqrt{2}}\int d\theta = -\frac{2}{\sqrt{2}}\left(\theta + \theta_0\right)$$

$\theta + \theta_0 = \xi + \xi_0$ or $\theta = -\frac{\sqrt{2}}{2}\left(\xi + \xi_0 - \theta_0\right) \equiv -\frac{\sqrt{2}}{2}\left(\xi + \tilde{\xi}_0\right)$

Therefore, the solution is $f = -\frac{1}{2}c\,\mathrm{sech}^2\left[\frac{\sqrt{2}}{2}\left(x - ct + x_0\right)\right]$

(e) ***General waves of permanent form***

The qualitative nature of the solution $f\left(\xi\right)$ of Eq. (14) for arbitrary values of constants c, A *and* B can be determined by corresponding analysis. The quantitative behavior, however, requires the use of elliptic functions or numerical computations. For practical applications, we are interested only in real bounded solutions $f\left(\xi\right)$ of

$$\frac{1}{2}\left(f'\right)^2 = f^3 + \frac{1}{2}cf^2 + Af + B \equiv F\left(f\right)$$

Thus we require $\left(f'\right)^2 \geq 0$ and the form of $F\left(f\right)$ shows that $f\left(\xi\right)$ vary monotonically until f' vanishes. In other worlds the zeros of $F\left(f\right)$ are important. Now we can integrate this equation formally as

$$\xi = \xi_i \pm \int_{f_i}^{f} \frac{df}{\sqrt{2F(f)}},$$

where f_i is i^{th} zero of $F(f)$, i.e. solution of equation $F(f)=0$ and correspondingly, ξ_i is the point, where this zero appears. In the book of Drazin and Johnson detailed analysis of all possibilities is performed.

(g) ***Generality of the KdV equation***

There are numerous examples in physics which can be approximately described by the KdV equation. Apart from J.Russell's observation, there are many other physical examples where the KdV equation arises, such as: non-linear electrical lines – linear approximation shows that the electrical chain behaves as a weakly dispersive medium for long wavelength signals, blood pressure waves appear also as KdV solitons, within suitable approximation, internal waves in oceanography and so on. These examples are however sufficient to show the situations that lead to the KdV equation:

- It applies to systems, which at the first level of approximation, are described by a *hyperbolic linear equation* such as the wave equation $u_{tt} - c_0^2 u_{xx} = 0$
- Moreover a *weak nonlinearity* as $\varepsilon f(u)$, with $f(u) = Au^2 + Bu^3$ must exist.
- Finally the system must show a *weak dispersion* with a dispersion relation for small wave vectors q of the form $\omega(q) = c_0 q\left(1 - \lambda_0^2 q^2\right)$, which can arise from terms like u_{xxxx} or u_{xxtt} in the equation of motion.

In order to stay within the weakly dispersive range, let us consider signals $u(x,t)$ with a slow variation. This implies that their Fourier spectrum $F(q)$ only includes components at a small wave vector q (such that $\delta = q\lambda_0 \ll 1$, *i.e.* $|q|$ is below some value $q_{\max}$). Therefore they can be written as

$$u(x,t) = \int_{-\infty}^{+\infty} F(q) e^{i(qx-\omega t)} dq \simeq F(0) \int_{-q_{\max}}^{q_{\max}} e^{i\left(qx - c_0 qt + c_0 \lambda_0^2 q^3 t\right)} dq$$

Introducing dimensionless variables $X = x/\lambda_0 \quad and \quad T = c_0 t/\lambda_0$, we get

$$u(x,t) \simeq F(0) \int_{-q_{\max}}^{q_{\max}} e^{i\delta(X-T)} e^{i\delta^3 T} dq$$

In order to derive the KdV equation, we change to a frame moving at speed c_0 by defining $\xi = X - T \quad and \quad \tau = T$. We can see that this last equation leads to a time variation of order δ^3 if the space variation is of order δ. This is what leads to the time variation of order $\varepsilon^{3/2}$ once we assume that dispersion and nonlinearity balance each other ($\delta = \varepsilon^{1/2}$ in the cases that we investigated earlier).

Thus it appears that a weak nonlinearity and a rather general form of dispersion relation are enough to predict that, in some range of excitation, a given physical system may show a behavior approximately described by the KdV equation.

Exercises:

1. For what values of the constants c_1 and c_2 is the function $u(x,t)=\dfrac{c_1}{(x+c_2)^2}$ a solution of KdV equation?
2. For what value(s) of the constant c is the function $u(x,t)=\dfrac{cx}{t}$ a solution to the KdV equation? Describe the dynamics "The graph of this function at any fixed time looks like … and as time passes …"
3. Let $a\neq 0$ be a constant and $u(x,t)$ a solution of the KdV equation.
   (a) For what number n will $\hat{u}(x,t)=a^2u(ax,a^nt)$ be a solution of the KdV equation for every choice of a and every KdV solution $u(x,t)$?
4. Verify that if $u(x,t)$ is any solution to the KdV equation and α is any constant, then $w(x,t)=u(x+\gamma t,t)+\alpha$ is also a KdV solution if you choose γ correctly. Derive a formula for γ as a function of α so that this will be true.

Lecture 4

3. Inverse Scattering Method

Given nonlinear partial differential equation (PDE), there is no general way of knowing whether it has soliton solution or not, or how the soliton solutions can be found. One of the powerful tool is the Inverse Scattering method. Bellow we give briefly main features of this method and apply to the KdV equation.

We want to study the time evolution of a spatially localized initial condition $u(x,t=0)$ which evolves according to the KdV equation

$$u_t-6uu_x+u_{xxx}=0 \qquad (17)$$

The basis of the method is to define an associated linear problem as follows:
We consider the Schrodinger equation

$$\left[\frac{d^2}{dx^2}+u(x,t)\right]\psi(x)=\lambda\psi(x), \qquad (18)$$

which define an eigenvalue problem. The potential is chosen to be the solution of the associated KdV equation (17) that we are looking for. Therefore it depends on one parameter, t , the time variable of the KdV equation. As we assume that u is localized solution $\left[\lim_{|x|\to\infty} u(x,t)=0\right]$, the Schrodinger equation generally has a spectrum:

- *discrete* eigenvalues, λ_m, associated with spatially localized ψ solutions,
- *a continuous spectrum,* $\lambda_k = k^2$, $(\mathrm{k} \geq 0)$ associated with eigenfunctions, which behave like a plane wave $\psi \simeq e^{\pm ikx}$ at infinity.

The potential u can be characterized by the discrete spectrum of the Schrodinger equation (18) and by its scattering properties, i.e. the transmission T and reflection R of an incident wave e^{-ikx}. The three quantities λ_m, T *and* R depends on the potential and therefore they are functions of its parameter, t .

It is known the following important ***theorem:***

If $u(x,t)$ evolves according to the KdV equation, then

* the discrete eigenvalues do not depend on the parameter t .

* the coefficients T *and* R can be easily calculated at any time as a function of their initial values calculated at $t=0$.

The consequence:

If we know the initial condition $u(x,t=0)$, we can compute λ_m, T, R for $t=0$ by solving the *linear* problem. Then, knowing these quantities at any time $t>0$ using the results of this theorem, we know the scattering properties of the potential, which we wish to determine, because it is the solution of the KdV equation. But there is a linear method to solve the inverse scattering problem, which can be used to build up a potential from the scattering properties. So that $u(x,t)$ can be computed *through a sequence of linear steps.*

(a) *Inversion of the scattering data*

The inversion of the scattering data determines the potential from the behavior of the solutions at large distances. For the bound states this behavior can be derived by giving the quantities

$$C_m = \lim_{x\to+\infty} \psi_m(x) e^{\kappa_m x} , \qquad (19)$$

where ψ_m is the normalized eigenfunction. The condition $\lim_{|x|\to\infty} u(x)=0$ guarantees that this limit exists. The eigenfunctions corresponding to *continuum states* $\lambda_k = k^2$ with positive eigenvalues, cannot be normalized, because at infinity they have an oscillatory behavior, proportional to $e^{\pm ikx}$. When we study the scattering by the potential of a wave, coming from $+\infty$, the solution is of the form

$$\psi_k(x) = \begin{cases} e^{-ikx} + R(k)e^{ikx} & \text{for} \quad x \to +\infty \\ T(k)e^{-ikx} & \text{for} \quad x \to -\infty \end{cases}$$

with a condition which expresses energy conservation $|R|^2 + |T|^2 = 1$

Knowing the behavior of ψ when x tends to $+\infty$, i.e. Knowing $\lambda_m\left(or\ \kappa_m\right),\ C_m\ \ and\ \ R(k)$ is sufficient to determine the scattering potential $u(x)$. The mathematical problem of the inversion of the scattering data has been extensively studied because it is of great practical importance.

The final results in the one-dimensional case which is of interest here are completed in the equations of Gelfand-Levitan- Marchenko:

$$u(x) = -2\frac{dK(x,x)}{dx}, \tag{20}$$

where K is a solution of the integral equation

$$K(x,y) + B(x+y) + \int_x^{\infty} dz B(z+y)K(x,z) = 0, \tag{21}$$

with

$$B(\xi) = \sum_{m=1}^{N} C_m^2(t)e^{-i\kappa_m t} + \frac{1}{2\pi}\int_{-\infty}^{+\infty} R(k,t)e^{-ikx}dk \tag{22}$$

Eq. (21) is *linear* for the unknown function $K(x,y)$ here x plays the role of a parameter, because the integration is not carried out with respect of x.

Eq. (21) is often hard to solve, but there is however an important case for which we can get an exact solution: it is the case of separable kernel $B(z+y)$, i.e. a kernel which can be written as a linear combination of the products of a function of z by a function of y. Assuming that

$$B(z+y) = \sum_{m=1}^{N} F_m(z)G_m(y) \tag{23}$$

Eq. (21) becomes

$$K(x,y) + \sum_{m=1}^{N} F_m(x)G_m(y) + \sum_{m=1}^{N} G_m(y)\int_x^{+\infty} dz F_m(z)K(x,z) = 0 \tag{24}$$

We see that the full kernel K may be represented in the separable form as well

$$K(x,y) = \sum_{m=1}^{N} L_m(x)G_m(y) \tag{25}$$

where the unknown functions can be obtained from the system of equations

$$L_m(x)+F_m(x)+\sum_{\rho=1}^{N}L_\rho(x)\int_x^{+\infty}dzF_m(z)G_\rho(z)=0 \tag{26}$$

Hence we get a set of N coupled *algebraic* equations for the unknown L_m, in which x plays the role of a simple parameter.

Such a separable case is obtained for all reflection less potentials, i.e. the potentials such that $R(k)=0$, because in this case the expression (22) reduces to

$$R(\xi)=\sum_{m=1}^{N}C_m^2e^{-\kappa_m\xi}$$

This happens when the initial condition is a multisoliton solution.

The relevant examples are given below.

(**b**) ***Single soliton initial condition***

As we know the KdV equation has the soliton solution (see, Eq. (16)).

Consider, for instance, the initial condition

$$u(x,t=0)=-2\operatorname{sech}^2 x$$

i.e. we are taking for speed $c=4$. This leads to the associated linear eigenvalue problem

$$\psi_{xx}+\left(2\operatorname{sech}^2 x+\lambda\right)\psi=0$$

Solutions of this Schrodinger equation is well known. For this particular ratio between the depth and the width of the potential well, there is only one bound state ($N=1$). It is

$$\psi(x)=\frac{1}{\sqrt{2}}\operatorname{sech} x, \qquad \textit{with}\ \ \lambda_1=-1, \quad \textit{i.e.}\quad \kappa_1=1$$

The asymptotic behavior $\psi(x)\sim\sqrt{2}e^{-x}$ gives $C_1(0)=\sqrt{2}.$ because the potential is reflection less, we get altogether

$$\kappa_1=1, \quad C_m(t)=\delta_{m1}\sqrt{2}e^{4t}; \qquad R(\xi)=2e^{8t-\xi}$$

Defining two functions F_1 and G_1, by $F_1(x)=\left(2e^{8t}\right)e^{-x}$ and $G_1(y)=e^{-y}$, we get $K(x,y)=e^{-y}L_1(x)$, where $L_1(x)$ is a solution of the equation

$$L_1(x)+2e^{8t}e^{-x}+L_1(x)\int_x^{+\infty}dz\left(2e^{8t}\right)e^{-z}e^{-y}=0$$

It leads to

$$L_1(x)=\frac{2e^{2t}e^{-x}}{1+e^{8t-2x}}; \qquad\qquad K(x,y)=-2\frac{e^{8t-x-y}}{1+e^{8t-2x}}$$

and therefore

$$\begin{aligned}u(x,t)&=-2\frac{dK(x,x)}{dx}=\\&=-8\frac{d}{dx}\frac{e^{8t-2x}}{\left(1+e^{8t-2x}\right)^2}=-2\sec h^2(x-4t)\end{aligned} \tag{27}$$

which is the soliton with c=4, as expected.

This simple example has shown that the initial condition $u(x,t=0)=-2\sec h^2 x$ is indeed a permanent profile soliton solution and we have determined its speed.

(c) ***Two- soliton solution***

Now we consider problem for which the initial profile is $u(x,t=0)=-6\sec h^2 x$. So that we must study the associated eigenvalue problem

$$\psi_{xx}+\left(6\sec h^2 x+\lambda\right)\psi=0 \tag{28}$$

or

$$\frac{d}{dT}\left\{\left(1-T^2\right)\frac{d\psi}{dT}+\left(6+\frac{\lambda}{1-T^2}\right)\right\}\psi=0,$$

Where $T=\tanh x$. This equation has bounded solutions for $\lambda=-\kappa^2\,(<0)$, if $\kappa_1=1$ or $\kappa_2=2$ of the form **(do exercise)**

$$\psi_1(x)=\sqrt{\frac{3}{2}}\tanh x\cdot\operatorname{sech} x\ , \qquad\qquad \psi_2(x)=\frac{\sqrt{3}}{2}\operatorname{sech}^2 x\ ,$$

both of which are have been made to satisfy the normalization condition. The asymptotic behaviors of these solutions are

$$\psi_1(x)\sim\sqrt{6}e^{-x}, \qquad \psi_2(x)\sim 2\sqrt{3}e^{-2x}; \qquad as\quad x\to+\infty$$

so that

$$C_1(0)=\sqrt{6}, \qquad C_2(0)=2\sqrt{3}$$

and then

$$C_1(t)=\sqrt{6}e^{4t}, \qquad C_2(t)=2\sqrt{3}e^{32t}$$

The choice of initial profile ensures that $b(k)=0$ for all k and so $b(k,t)=0$ for all t. The function F then becomes

$$F(X,t)=6e^{8t-X}+12e^{64t-2X}$$

and the Gelfand-Levitan-Marchenko equation is therefore

$$K(x,z;t)+6e^{8t-(x+z)}+12e^{64t-2(x+z)}+ \\ +\int_x^\infty K(x,y;t)\left\{6e^{8t-(x+z)}+12e^{64t-2(x+z)}\right\}dy=0 \tag{29}$$

The solution for K takes the form

$$K(x,z;t)=L_1(x,t)e^{-x}+L_2(x,t)e^{-2x} \tag{30}$$

Collecting the coefficients of exponents, we obtain the pair of equations

$$L_1+6e^{8t-x}+6e^{8t}\left(L_1\int_x^\infty e^{-2y}dy+L_2\int_x^\infty e^{-3y}dy\right)=0 \tag{31}$$

$$L_2+12e^{64t-2x}+12e^{64t}\left(L_1\int_x^\infty e^{-3y}dy+L_2\int_x^\infty e^{-4y}dy\right)=0\ , \tag{31'}$$

where $L_{1,2}$ are the functions of x. After evaluation of definite integrals, these two equations become

$$L_1+6e^{8t-x}+3L_1e^{8t-2x}+2L_2e^{8t-3x}=0$$

$$L_2+12e^{64t-2x}+4L_1e^{64t-3x}+e^{64t-4x}=0\ , \tag{32}$$

which can be solved to yield

$$L_1(x,t)=6\left(e^{72t-5x}-e^{8t-x}\right)/D; \qquad L_2(x,t)=-12\left(e^{64t-2x}+e^{72t-4x}\right)/D\ , \tag{33}$$

where

$$D(x,t)=1+3e^{8t-2x}+3e^{64t-4x}+e^{72t-6x} \tag{34}$$

The solution of the KdV equation can now be expressed as

$$u(x,t) = -2\frac{\partial}{\partial x}\left(L_1 e^{-x} + L_2 e^{-2x}\right) = 12\frac{\partial}{\partial x}\left\{\left(e^{8t-2x} + e^{72t-6x} - 2e^{64t-4x}\right)/D\right\} \qquad (35)$$

which can be simplified to give

$$u(x,t) = -12\frac{3 + 4\cosh(2x-8t) + \cosh(4x-64t)}{\left\{3\cosh(x-28t) + \cosh(3x-36t)\right\}^2} \qquad (36)$$

It is the two-soliton solution.

Let us comment, why it is a "two-soliton" solution?

Since the solution is valid for all positive and negative t, we may examine the development of the profile both before and after the formation of the initial profile, specified at $t = 0$. The wave profile, plotted as a function of x, at 5 different times, is shown **in Figure** bellow.

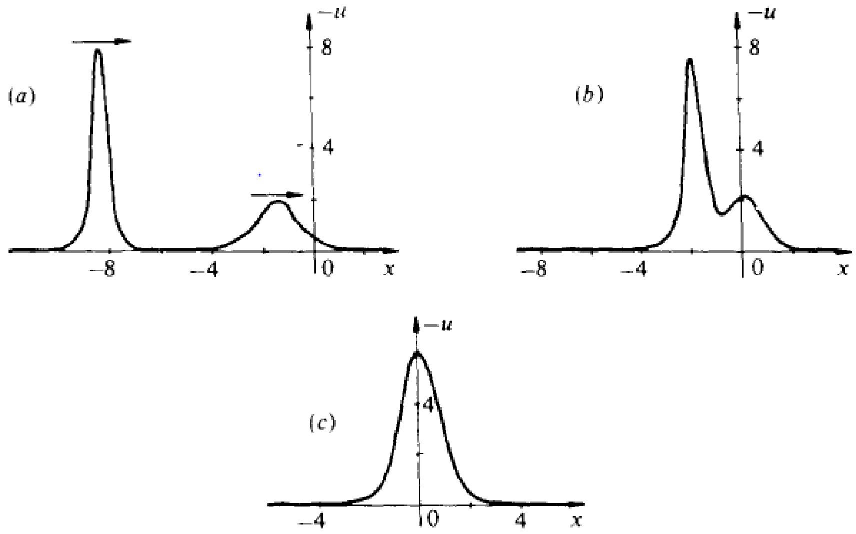
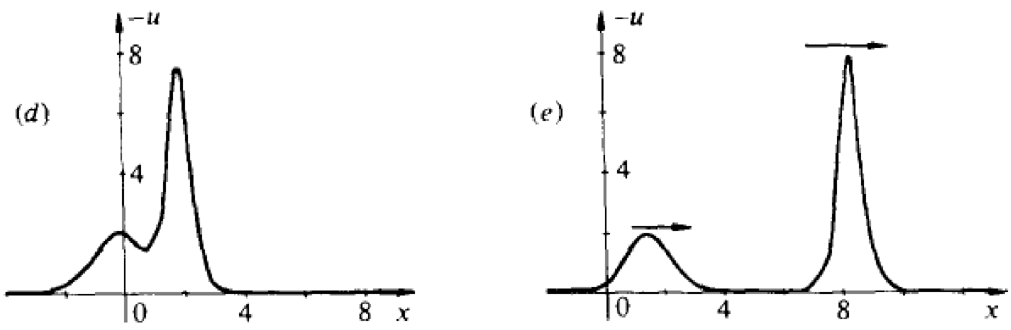


Fig. 11. The two soliton solution with $u(x,0) = -6\,\mathrm{sech}^2\, x$: see (c) (a) $t = -0.5$; (b) $t = -0.1$;



The same with $(d) = 0.1$; (e) $t = 0.5$

The solution depicts two waves, where the taller one catches the shorter, coalesces to form a single wave – our initial profile at $t=0$ – and then reappears to the right and moves away from the shorter one as t increases.

This interaction might see, at first sight, to be a purely linear process but this is not so. A more careful examination of the plots shows that the taller wave has moved *forward*, and the shorter one backward, relative to them positions they would have reached if the interaction was indeed linear.

The character of this solution is also made evident by examine the asymptotic behavior of $u(x,t)$, as $t \to \pm\infty$. For example, if we introduce $\xi = x-16t$, then the solution (36) can be expressed as

$$u(x,t) = -12\frac{3+4\cosh(2\xi+24t)+\cosh 4\xi}{\{3\cosh(\xi-12t)+\cosh(3\xi+12t)\}^2} \qquad (37)$$

It can be expanded as $t \to \pm\infty$ at ξ fixed. The asymptotic limit ensures that we follow the development of the wave which moves at a speed 16 (if such a one exists). We thus obtain

$$u(x,t) \sim -8\,\mathrm{sech}^2\left(2\xi \mp \frac{1}{2}\log 3\right), \qquad as\ \ t \to \pm\infty, \quad \xi = x+16t \quad (38)$$

and a similar procedure can be adopted for the wave which moves at the speed 4 let $\eta = x-4t$, then

$$u(x,t) \sim -2\sec h^2\left(\eta \pm \frac{1}{2}\log 3\right), \qquad as \quad t \to \pm\infty \qquad (39)$$

In fact these two asymptotic forms can be combined to produce a uniformly valid solution, since the error term are exponentially small, where

$$u(x,t) \sim -8\,\mathrm{sech}^2\left(2\xi \mp \frac{1}{2}\log 3\right) - 2\,\mathrm{sech}^2\left(\eta \pm \frac{1}{2}\log 3\right), \qquad \text{as } t \to \pm\infty \qquad (40)$$

The solution is therefore compound of two solitary waves at infinity, with phase shifts now explicit. From the last solution we see that the taller wave moves forward by an amount $x = \frac{1}{2}\log 3$, and the shorter one moves back by $x = \log 3$.

Finally, solution here contains no other component (such as, for example, an oscillatory dispersion wave)

(d) *N-soliton solution*

Above method can be generalized by introducing the matrix formulation. The initial profile is now taken to be

$$u(x,0) = -N(N+1)\sec h^2 x \tag{41}$$

So we have N discrete eigenvalues and no continuous spectrum (i.e. $b(k)=0$ for all k). These eigenvalues are $\lambda = -\kappa^2$, where $\kappa = \kappa_n = n$, for $n = 1, 2, ..., N$ and the discrete eigenfunctions take the asymptotic form:

$$\psi_n(x) \sim c_n e^{-nx} \ , \qquad \text{as} \quad x \to +\infty$$

If we use the associated Legendre functions, defined as

$$P_N^n(T) = (-1)^n \left(1-T^2\right)^{n/2} \frac{d^n}{dT^n} P_N(T) \qquad\qquad P_N(T) = \frac{1}{N!2^N} \frac{d^N}{dT^N} \left(T^2-1\right)^N$$

we have

$$\psi_n(x) \propto P_N^n(\tanh x)$$

and then $\qquad c_n(t) = c_n(0)\exp\left(4n^3 t\right)$

The function F in the Gelfand-Levitan- Marchenko equation is

$$F(\mathrm{X},t) = \sum_{n=1}^{N} c_n^2(0)\exp\left(8n^3 t - nX\right)$$

and therefore

$$K(x,z;t) = \sum_{n=1}^{N} c_n^2(0)\exp\left\{8n^3 t - n(x+z)\right\} +$$

$$+\int_x^\infty K(x,y;t)\sum_{n=1}^{N} c_n^2(0)\exp\left\{8n^3 t - n(y+z)\right\} dy = 0 \tag{32}$$

For K now take the form

$$K(x,z;t) = \sum_{n=1}^{N} L_n(x,t)e^{-nz} \tag{33}$$

Proceeding now by standard manner, we obtain the final result

$$u(x,t) \sim -2\sum_{n=1}^{N} n^2 \,\mathrm{sech}^2 \left\{n\left(x-4n^2 t\right) \mp x_n\right\}; \qquad as \;\; t \to \pm\infty \tag{34}$$

where x_n is a phase, given by

$$\exp(2x_n) = \prod_{m=1,m\neq n}^{N} \left|\frac{n-m}{n+m}\right|^{\mathrm{sgn}(n-m)}$$

Thus the asymptotic solution represents separate solitons, ordered according to their speeds, as $t \to +\infty$, the tallest (and therefore fastest) soliton is at the front followed by progressively shorter soliton behind.

All N solitons interact at $t = 0$ to form a single sech^2 pulse which was specified as the initial profile at that instant. Some plots of three-soliton solution ($N = 3$; $u(x,0) = -12\,\text{sech}^2\, x$) are given in Fig.12, where the emerging solitons are of amplitudes 18, 8 and 2.

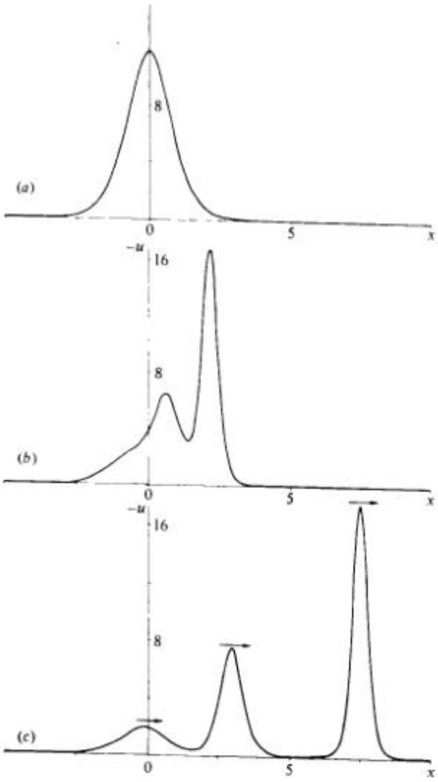


Fig. 12 **The three-soliton solution with** $u(x,0) = -12\,\text{sech}^2\, x \quad [(b)\ t = 0.05; \quad (c)\ \ t = 0.2]$

(In the last two figures $-u$ **is plotted against t)**

Lecture 5

4. Further properties of the KdV equation

(a) The role of ***conservation laws***

The dynamics of continuous media can be equivalently described by Lagrange (or Hamilton) formalism. For KdV equation, as it involves a third derivative, one must consider Lagrangians depending on higher derivatives. It is well known that the Lagrangian density $\mathscr{L}(q, q_t, q_x, q_{xx}, x, t)$ leads to the Euler-Lagrange Equation

$$\frac{d}{dt}\left(\frac{\partial \mathscr{L}}{\partial q_t}\right) + \frac{d}{dx}\left(\frac{\partial \mathscr{L}}{\partial q_x}\right) - \frac{d^2}{dx^2}\left(\frac{\partial \mathscr{L}}{\partial q_{xx}}\right) = \frac{\partial \mathscr{L}}{\partial q} \qquad (35)$$

In order to derive the KdV equation one can introduce the corresponding Lagrangian. It has a form

$$L = \int dx \mathscr{L} = \int dx \frac{1}{2}\left[q_x q_t + 2q_t^3 - q_{xx}^2 \right] \tag{36}$$

Then it follows the following equation for q

$$q_{xt} - 6q_x q_{xx} + q_{xxxx} = 0 \tag{37}$$

and defining $u = q_x$, we obtain the KdV equation

$$u_t - 6uu_x + u_{xxx} = 0 \tag{38}$$

Now the invariance of L under some continuous transformation produces the Noether's conservation laws.

In general, these equations for conservation laws have the following form

$$I_t + J_x = 0 \tag{39}$$

such that $J \to const$, when $|x| \to \infty$. Therefore by this equation one has

$$\int_{-\infty}^{\infty} dxI = const \tag{40}$$

a constant of motion. I and J are called conserved density and the flux, respectively.

The KdV equation can be performed in the form of conservation equation. Indeed, rewrite it as

$$u_t + \left(u_{xx} + 3u^2\right)_x = 0 \tag{41}$$

which implies that $I_1 = u$ is a conserved density. If, for example, u denotes the density of medium (gas or fluid), it follows the travelling mass conservation

$$\int_{-\infty}^{\infty} dxu = const. \tag{42}$$

Multiplying the KdV equation by u, one obtains

$$\frac{1}{2}\left(u^2\right)_t + \left(uu_{xx} - \frac{1}{2}u_x^2 + 2u^3 \right)_x = 0$$

Therefore $I_2 = u^2$ is also a conserved density and

$$\int_{-\infty}^{\infty} u^2 dx = const \tag{43}$$

Similarly, multiply KdV equation by $3u^2$ and then apply $u_x \dfrac{\partial}{\partial x}$ to the same equation, add two derived equations, we obtain

$$\left(u^3+\frac{1}{2}u_x^2\right)_t+\left(-\frac{9}{2}u^4+3u^2u_{xx}-6uu_x^2+u_xu_{xxx}-\frac{1}{2}u_{xx}^2\right)=0$$

We see immediately that

$$I_3=u^3+\frac{1}{2}u_x^2 \qquad (44)$$

is also a conserved density.

For the KdV equation these first three conserved densities are relatively easy to obtain (guess). It is not so for higher constants. A systematic method of finding them is available using so called Miura-Gardner transformation (see, e.g. P.C. Drazin et al., p.94).

In fact there exists an infinite number of conservation laws. The importance of the existence of an infinite number conservation laws is that they are believed to be essential for the elastic collision property of solitons to be established. The solitons have to and can maintain their identities after collision because of many constraints required by the infinite number of conservation laws.

(b) ***Lax formulation for KdV and other soliton equations***

We described above the KdV equation, which has many special properties. It is clear that, indeed, other equations with similar properties do exist: The KdV equation does not stand alone in this class of evolution equations. Lax in 1968 developed arguments, which introduces far deeper and more fundamental ideas than we have met in the inverse scattering method.

Suppose that we wish to solve the initial- value problem for u, where $u(x,t)$ satisfies some nonlinear evolution equation of the form

$$u_t=N[u], \qquad (*)$$

with $u(x,0)=f(x)$. We assume that $u\in Y$ for all t, and that $N:\ Y\to Y$ is some nonlinear operator, which is independent of t but may involve x, or derivatives with respect to x, and Y is some appropriate function space.

Next, we suppose that the evolution equation above can be expressed in the form

$$L_t=ML-LM$$

where L and M are some operators in x, which operate on elements of a Hilbert space, H, and which may depend upon $u(x,t)$ (By L_t we mean the derivative with respect to the parameter t, as it appears *explicitly* in the operator L; for example, if

$$L = -\frac{\partial^2}{\partial x^2} + u(x,t),$$

then $L_t = u_t$). The Hilbert space, H, is a space, with an inner product, (ϕ,ψ), which is complete; we assume that L is self-adjoint, so that $(L\phi,\psi) = (\phi, L\psi)$ for all $\phi,\psi \in H$

Now we introduce the *eigenvalue (or spectral) equation*

$$L\psi = \lambda\psi \quad \text{for } t \geq 0 \quad \text{and} \quad -\infty < x < \infty,$$

where $\lambda = \lambda(t)$. Differentiating with respect to t, we see that

$$L_t\psi + L\psi_t = \lambda_t\psi + \lambda\psi_t,$$

which becomes

$$\begin{aligned} \lambda_t\psi &= (L-\lambda)\psi_t + (ML - LM)\psi = \\ &= (L-\lambda)\psi_t + M\lambda\psi - LM\psi = \\ &= (L-\lambda)(\psi_t - M\psi) \end{aligned}$$

The inner product of ψ with this equation gives

$$(\psi,\psi)\lambda_t = (\psi,(L-\lambda)(\psi_t - M\psi))$$

Since $(L-\lambda)$ is self-adjoint, and so

$$(\psi,\psi)\lambda_t = (0,(L-\lambda)(\psi_t - M\psi)) = 0$$

$$\text{Or} \qquad \lambda_t = 0$$

Thus each value of operator L *is a constant*. With $\lambda_t = 0$, we obtain

$$(L-\lambda)(\psi_t - M\psi) = 0$$

So that $\psi_t - M\psi$ is an eigenfunction of the operator L with eigenvalue λ. Hence

$$\psi_t - M\psi \propto \psi,$$

And we can always define M with the addition of the product of the identity operator and an appropriate function of t; this will not alter equation (*). Thus we have the *time-evolution equation* for ψ,

$$\psi_t = M\psi, \qquad \text{for} \quad t > 0 \qquad (**)$$

In other words we have the following *theorem*:

If the evolution equation

$$u_t - N[u] = 0$$

can be expressed as the Lax equation

$$L_t + [L, M] = 0,$$

and if $\qquad L\psi = \lambda\psi,$

then $\lambda_t = 0$ and ψ evolves according to equation (**).

(b-I) *KdV in Lax form*

If $L = \sum_i c_i(x,t)\partial^i$ is an ordinary differential operator in the variable x whose coefficients also depend on the time parameter t, then $L_t = \sum_i \frac{\partial c_i}{\partial t}\partial^i$ is its time derivative, an indication of how it will evolve infinitesimally in time. Lax recognized the significance of the fact that the KdV equation can be written in the form $L_t = [M, L]$ for an appropriate operator M.

Theorem: Let $L = \partial^2 + u(x,t)$ and $M = \partial^3 + \frac{3}{2}u(x,t)\partial + \frac{3}{4}u_x(x,t)$. The question of whether the function $u(x,t)$ is a solution of the KdV equation, $u_t = \frac{3}{2}uu_x + \frac{1}{4}u_{xxx}$, is equivalent to the question of whether L and M satisfy the Lax equation $L_t = [M, L]$.

Proof: The left side of the Lax equation is simply $u_t(x,t)$. Thus, for $[M, L]$ to be equal to it all of the terms with positive powers of ∂ must cancel out. In fact, this is the case since

$$ML = \partial^5 + \frac{5}{2}u\partial^3 + \frac{15}{4}u_x\partial^2 + \frac{3}{2}\left(u^2 + 2u_{xx}\right)\partial + \frac{9}{4}uu_x + u_{xxx}$$

$$LM = \partial^5 + \frac{5}{2}u\partial^3 + \frac{15}{4}u_x\partial^2 + \frac{3}{2}\left(u^2 + 2u_{xx}\right)\partial + \frac{3}{4}(uu_x + u_{xxx})$$

The coefficients on all of the positive powers of ∂ are the same in these two products regardless of the choice of function $u(x,t)$. We have that

$$[M, L] = \frac{3}{2}uu_x + \frac{1}{4}u_{xxx}$$

Clearly, this is equal to $L_t = u_t$ if and only if $u(x,t)$ happens to be a solution to the KdV equation.

This result is more miraculous than it may at first appear. Note that it is not possible to write just any differential equation in the form $L_t = [M, L]$ for suitable differential operators L and M. If we can write a given differential equation in this way, then we say it has a *Lax Form* and that the operators L and M form its *Lax Pair*. (For instance, the operators L and M from above theorem are the Lax Pair for the KdV equation).

To illustrate how rare this is, and to demonstrate why the coefficients in the above equation are a natural choice, let us try to find a different equation by starting with a slightly more general form for M.

Consider a suitable example: Suppose again that
$L = \partial^2 + u(x,t)$; $M = \partial^3 + \alpha(x,t)\partial + \beta(x,t)$. What must be true about the commutator $[M, L]$ for the equation $L_t = [M, L]$ to be sensible? What is the most general nonlinear evolution equation that can have Lax operators of this type and why is it not much of an improvement over the previous example?

Solution : Since L_t is a zero order operator (there are no positive powers of ∂) for the Lax equation to be sensible, the commutator must also be a zeroth order operator. Then, in order to see what sorts of equations we can generate in this way, we need to find the most general choice of coefficients α, β that eliminate all positive powers of ∂ in $[M, L]$.

We compute the product of the operators in each order:

$$LM = \partial^5 + (\alpha + u)\partial^3 + (2\alpha_x + \beta)\partial^2 + (\alpha u + 2\beta_x + \alpha_{xx})\partial + \beta u + \beta_{xx}$$

$$ML = \partial^5 + (\alpha + u)\partial^3 + (3u_x + \beta)\partial^2 + (\alpha u + 3u_{xx})\partial + u_{xxx} + u + \alpha u_x$$

The coefficients of ∂^5 and ∂^3 are already equal for any choice of these unknown functions. However, for the coefficients of ∂^2 to be equal we must have that $\alpha = \frac{3}{2}u + c_1$ for some constant c_1. Similarly, equating the coefficients of the ∂ terms in the two products we conclude that $\beta = \frac{3}{4}u_x + c_2$.

Thus, for $L_t = [M, L]$ to make sense as a Lax equation assuming L and M have these very general forms, we are forced into assuming these values for α and β with only choice of the constants c_1 and c_2 as freedom. But then, since $[M, L] = \left(\frac{3}{2}u + c_1\right)u_x + \frac{1}{4}u_{xxx}$ the more general equation we can write in this way is

$$u_t = \frac{3}{2}uu_x + \frac{1}{4}u_{xxx} + c_1 u_x$$

Which is equivalent to our KdV equation (3.1) but with the solution just shifted by a constant, $u(x,t) \to u(x,t) - \frac{2c_1}{3}$. Note then that the coefficients $3/2$ *and* $1/4$ appear seemingly out of nowhere without us making any specific assumptions and in this sense are the natural coefficients for the KdV equation.

Lecture 6

(b-II) **Other soliton equations**

We saw above that if we choose L and M to be ordinary differential operators of orders 2 and 3 respectively, then essentially the only equation we can write in the Lax form $L_t = [M, L]$ is the KdV equation itself. However, the KdV equation is not the *only* differential equation with the Lax form. As we will see, we can find many more by assuming other forms for the operators L and M.

What does it tell us about a differential equation when we learn that it has a Lax form? It is a rather good clue that the equation shares those amazing properties of the KdV equation: being exactly solvable and having particle-like soliton solutions. So, let us proceed and find other KdV-like equations in the sense that they share these important and rare properties.

What if L is still a Schrodinger operator but M has order 5? With $L = \partial^2 + u(x,t)$ we have that the left side of the Lax equation $L_t = [M, L]$ is sure to be just u_t. If we can find M so that the right side has order zero then this will still be an evolution equation. The most natural generalization is to move on the higher order differential operators, for instance, of order 5.

Our aim is to choose values for the coefficients of

$$M = \partial^5 + \alpha_4(x,t)\partial^4 + \alpha_3(x,t)\partial^3 + \alpha_2(x,t)\partial^2 + \alpha_1(x,t)\partial + \alpha_0$$

so, that $[M, L]$ is an operator of order zero. What nonlinear PDE for u we get from the Lax equation?

For simplicity, we will ignore the dependence on t in the following computations. Hence, it should be understood that all derivatives are taken with respect to x.

Proceeding to the same line as before, one can define each α_i -s and write the Lax equation as

$$u_t = \frac{1}{16}\left(30u^2u_x + 20u_xu_{xx} + 10uu_{xxx} + u_{xxxxx}\right)$$

Like the KdV equation it is an evolution equation for a function $u(x,t)$. Like the KdV equation, this equation has n-soliton solutions for any positive number n, it has many solution which can be written exactly in terms of ordinary exponential and rational functions. In other words it is *also* a soliton equation.

One can find a soliton equation by considering only a slight variation of the Lax operators for the KdV equation. Let us again consider differential operators of orders 2 and 3, but this time we will let L have of order 3 and M have order 2, and we get a different soliton equation.

Question: If $L = \partial^3 + \alpha\partial + \gamma$ and $M = \partial^2 + \beta$, how can the coefficients be chosen to be functions of x so that the Lax equation $L_t = [M, L]$ *make sense?* What nonlinear PDE for $\alpha(x,t)$ does it imply?

The **answer** is the following equation takes place:

$$\alpha_{tt} = -\frac{4}{3}\alpha_x^2 - \frac{4}{3}\alpha\alpha_{xx} - \frac{1}{3}\alpha_{xxxx}$$

This is a form of the nonlinear *Boussinesq Equation*, another soliton equation which arguably is more interesting than the KdV equation itself. It can be demonstrate that the n-soliton solutions of this equation are a bit more complicated than the corresponding solutions to the KdV equation. Moreover, Boussinesq studied and not published this equation before Korteweg and de Vries did their work on waves on translation on canals. So, in a sense, it is only historical coincidence that the KdV equation is considered to be the canonical example of a soliton equation (Not only does it just happen to be the one which was studied by Zabusky and Kruskal, but it is also a coincidence that it is named after Korteweg and de Vries since some have argued that it also should be named after Boussinesq who studied it first).

(c) *Matrix differential operators*

Many other soliton equations can be derived from Lax equations involving differential operators with *matrix* coefficients. Our rules for multiplication differential operators still apply in this case, except that the coefficient functions are $n \times n$ matrices and no longer commute with each other.

We consider here only one **example:** Suppose $L = a\partial + U(x)$ and $M = V(x)$ are matrix differential operators of order 1 and 0, respectively, for some constant a. Compute the commutator $[M, L]$.

$$\begin{aligned}[M,L] &= V(x)\cdot(a\partial + U(x)) - (a\partial + U(x))\cdot V(x) = \\ &= aV(x)\partial + V(x)U(x) - a\partial\cdot V(x) - U(x)V(x) = \\ &= aV(x)\partial - aV(x)\partial - aV'(x) + V(x)U(x) - U(x)V(x) = \\ &= aV'(x) + [V(x), U(x)]\end{aligned}$$

Let us apply this in the special case where the entries of $U \, and \, V$ depend on an unknown function $u(x,t)$. If $a = 4i$,

$$U = \begin{pmatrix} -4 & 2iu_x \\ -2iu_x & 4 \end{pmatrix}, \quad and \quad V = \begin{pmatrix} -\frac{i}{4}\cos u & \frac{i}{4}u_{xt} \\ \frac{i}{4}u_{xt} & \frac{i}{4}\cos u \end{pmatrix}$$

Then you can check that

$$L_t - [M, L] = \begin{pmatrix} u_x (u_{xt} - \sin u) & u_x \cos u - u_{xxt} \\ u_x \cos u - u_{xxt} & u_x (\sin u - u_{xt}) \end{pmatrix}$$

If L *and* M are Lax operators, then this would be equal to zero. Note that this is zero if either $u_x = 0$ (constant solution) or if

$$u_{xt}(x,t) = \sin u(x,t)$$

This is a *Sine-Gordon equation.* It is nonlinear, but nonlinearity takes the form of a trigonometric function applied to u. This is very important equation as it has many applications in science and geometry. We will return to this equation in future.

A major point of our consideration is to recognize that the Lax form gives us a way to recognize other differential equations which like the KdV equation deserve to be called "soliton equations".

(d) *The Nonlinear Schrodinger Equation*

The nonlinear Schrodinger equation is another soliton equation, but it is one which can only be used in the context of complex numbers. It is also notable as being the basis for the greatest commercial application of solitons: the use of solitons of light for communication.

The systems with small-amplitude plane wave solutions

$$u = Ae^{i(qx - \omega t)} + c.c.,$$

(where we denote by c.c. the complex conjugate of the expression that precedes this symbol) 'are drastically different from the soliton solutions that we investigated. Therefore one may ask what happens to these plane waves when their amplitudes grow enough to allow nonlinearity to enter into play. The answer is that the plane waves may spontaneously self-modulate as shown **in Figure** 18, below.

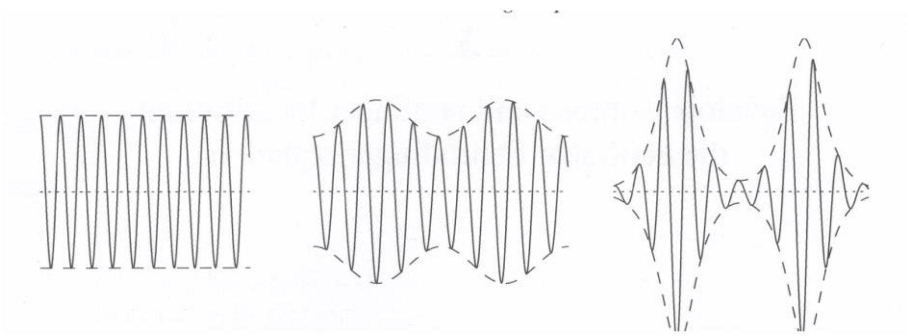


Fig.13 **Self-modulation of a plane wave. The dashed line shows the envelope of the wave which is displayed by the solid line. The three figures show three successive stages in the evolution of the initial plane wave**

This modulation arising due to the overtones induced by nonlinearity, can go as far as the splitting of the wave into “wave packets” which behave like solitons. These solitons are made of a carrier wave modulated by an envelope signal and this is why they are called *envelope solitons.*

Using the simple example of the pendulum chain (see, Fig.below) we shall derive the equation which describes them, which is extremely general in physics, since it appears naturally for most of the weakly dispersive and weakly nonlinear systems which are described by a wave equation in the small-amplitude.

(e) *Kadomtsev-Petviashvili equation*

All of the soliton equations we have considered thus far have depended on only *two* variables, one for space and one for time. The most obvious new feature is that it is a partial differential equation in three variables: x, y, t .An important example of the former was given by Kadomtsev and Petviashvili (KP), which appeared first in the stability study of the KdV solitons to transverse perturbations.

In order to determine the limit of validity of the KdV equation, it is also necessary to study stability of its solitary waves with respect to transverse perturbations. Assuming that a characteristic length in the transverse direction is large with respect to the spatial extend of the KdV equation, Kadomtsev and Petviashvili got their equation for a function $u(x, y, t)$

$$u_{yy} = \frac{4}{3}u_{xt} - 2u_x^2 - 2uu_{xx} - \frac{1}{3}u_{xxxx}$$

It looks like an entirely new equation, but as a next example shows, it actually is closely related to the KdV equation, which we have already studied in detail.

It is easy to observe that the KdV equation is “hidden” inside the KP equation by rewriting it as

$$u_{yy} = \frac{4}{3}\frac{\partial}{\partial x}\left(u_t - \frac{3}{2}uu_x - \frac{1}{4}u_{xxx} \right)$$

Note that the expression in the parentheses on the right-hand side is equal to zero precisely when u is a solution of the KdV equation. If the function $u(x, y, t)$ also happens to be independent of y, then u is a solution of KP equation (u_{yy} is also zero)

Since the left and right-handed sides of the equation are each equal to zero when $u(x, y, t)$ is a solution of KdV equation and is independent of y, we conclude that such a functions are also solutions of KP equation.

This means that we already know many solutions to the KP equation. However, this is only a small subset of the solutions of the KP equation. The KP equation should look familiar in another way. It also contains the Boussinesq equation “hidden within it” in the same way.

A single soliton solution of the KP equation looks like

$$u(x,y,t) = 2a^2 \operatorname{sech}^2\left\{a\left[x + by - \left(3b^2 + 4a^2\right)t + x_0\right]\right\}$$

Which travels in an arbitrary direction in the (x, y) plane, as well as multisoliton. A 2-soliton is shown in Fig. 14

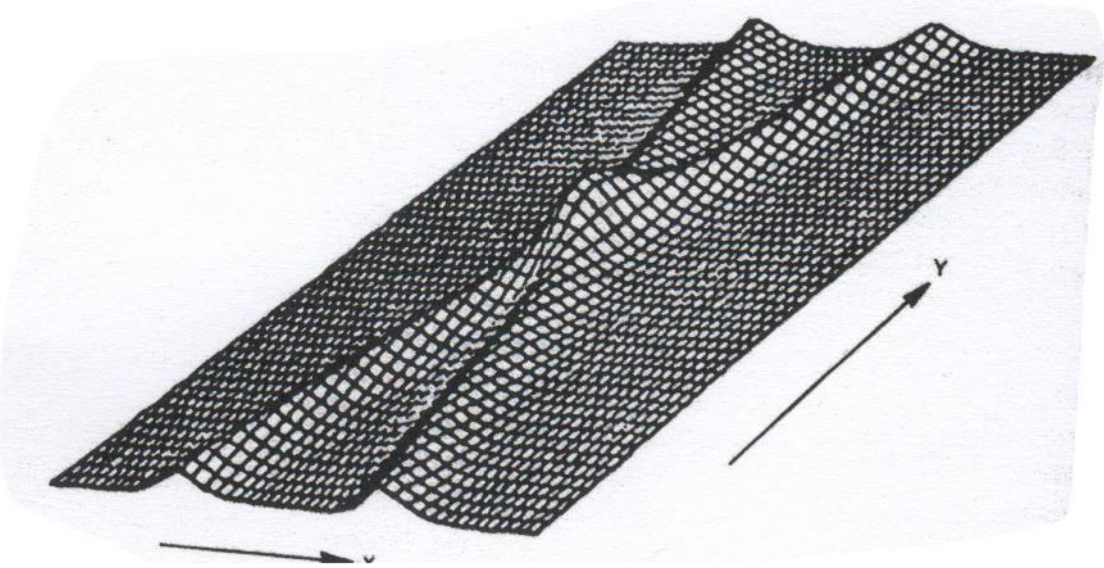


Fig.14 **A snapshot of a 2-soliton solution of the KP equation**

which resembles some real nonlinear waves observed in the shallow wave water off the Oregon coast. When the sign of u_{yy} is reversed, one obtains the so-called KP2 equation, which has the soliton solution

$$u(x,y,t) = 4\frac{\left(a^2y^2 - X^2 + a^{-2}\right)}{\left(a^2y^2 + X^2 + a^{-2}\right)},$$

where $X = x + a^{-1} - 3a^2t$. But such soliton is unstable.

The KP equation describes water surface waves and ion-acoustic waves in a plasma. Although the original motivation for the KP equation was the study of "ion acoustic wave propagation in plasmas", most readers will find it more intuitive to consider ocean waves as an application. Like the KdV equation the KP equation is certainly not an entirely realistic hydrodynamic model. For instance, it does not treat the z and y-directions equivalently; oscillations in the y-direction tend to be smoother. Still, one can see waves on the ocean which look like solutions of the KP equation.

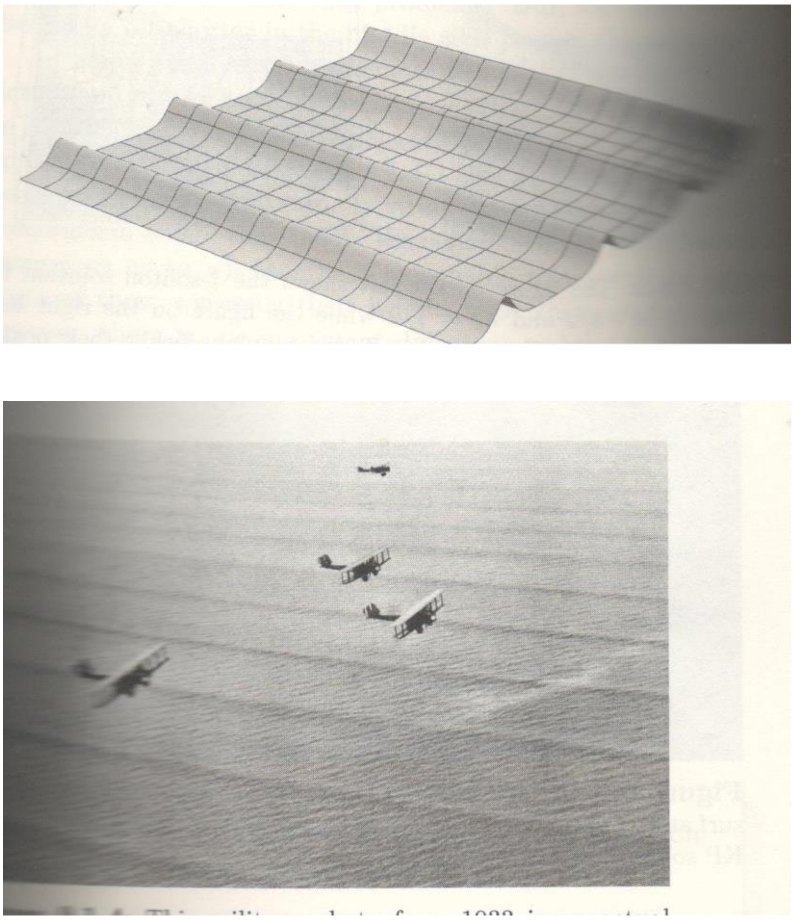


Fig.15 a) A "snapshot" of an exact solution of the KP equation

b) A military photo from 1933 is an actual snapshot of the ocean near Panama, shows a wave pattern very similar to the one in fig. a)

Some concluding remarks:

The soliton itself is a dramatic new concept in nonlinear science. Here, at last, on the classical level, is the entity that field theorists had been postulating for years, a local travelling wave pulse, a lump-like, coherent structure, the solution of a field equation with remarkable stability and particle-like properties. It is intrinsically nonlinear and owes its existence to a balance of two forces; one is linear and acts to disperse the pulse, the other is nonlinear and acts to focus it. Before the soliton physicists had often talked about wave packets and photons, which are solutions of the linear time-dependent Schrodinger equation. But such packets would always disperse on a time scale inversely proportional to the square of the spread of the packet in wave number space. Nonlinearity is essential for stopping and balancing the dispersion process.

What is remarkable is that so many of the equations, derived as asymptotic stability conditions under very general and widely applicable premises, are also soliton equations. But one of the key properties of a soliton equation is that it has an infinite number of conservation laws and associated symmetries.

What do we mean by a soliton equation? All we have said so far is that a soliton is a solitary, travelling wave pulse of a nonlinear partial differential equation with remarkable stability and particle-like properties.

The soliton story begins with the observation of a particular solitary wave on a canal. Presumably, the reason John Scott Rassell found the wave so interesting was not purely academic interest, but the desire to utilize it for improvements to ship designs. When the soliton concept was fully formed in the late of 20^{th} century, there was similarly interest in the individual solitons and their practical uses.

For instance, soliton theory has been applied to the study of tsunamis, rogue waves and internal waves, all hydrodynamic solitary waves which are of great interest due to damage they can potentially cause.

Because of the soliton's stability, it has also found application in communication where *optical solitons* (solitons of light traveling in a fiber optic cables) are used to transmit signals reliably over long distances.

Additionally, because of their very stability upon interaction that prompted Zabusky and Kruskal to name them "solitons" in the first place, soliton solutions to differential equations have application in particle physics. There have even been application of the soliton concept to biology, where the soliton dynamics are seen as having a role in DNA transcription or energy transfer.

Lecture 7

Part II. Topological solitons in relativistic field theoretical models

At the Introduction we have listed the certain requirements to soliton solutions of nonlinear equations. As we have seen there is no unique definition of solitons. In field theoretical models, which are used in particle physics, the names *solitary waves and solitons* refer to certain special solutions of non-linear wave equation. There is attempt to remain two principal features of linear wave equations: (i) we can construct a localized wave packet that will travel with uniform velocity $\pm c$ without *distortion in shape* (ii) Since for linear wave equation like (1) given two localized wave packet solutions $f_1(x-ct)$ and $f_2(x+ct)$ their sum $f_3(x,t) = f_1(x-ct) + f_2(x+ct)$ is also a solution. At large negative time $t \to -\infty$ this sum consists of the two packets widely separated and approaching each other essentially undistorted. At finite t they collide. But after collision they will asymptotically (as $t \to +\infty$) separate into the same two packets retaining their original shape and velocities.

These two features – the shape and velocity retention of a single packet and the asymptotic shape and velocity retention of several packets even after collision, do not take place for nonlinear systems in general. Typical wave equations in many branches of physics are much complicated: They can contain nonlinear terms, dispersive terms and several coupled wave fields with space dimensionality equal to 1, 2 or 3. The question is: can such equations, despite their complexity, nevertheless yield at least some solutions which enjoy this attractive features (i) and perhaps (ii)?

We have seen above that in hydrodynamics may happen that *both dispersive* and *non-linear* terms can balance each other effects in such a way that some special solutions do essentially enjoy feature (i). This can happen in one, two or three space dimensions, and such solutions are called solitary waves. If in some cases the feature (ii) is also exhibited, these solutions are called *solitons.*

The definition in particle physics is in terms of the energy density, rather than the wave fields themselves, since the former is more significant in this field. This means that we are restricted ourselves to those field equations that have an associated energy density $\varepsilon(\boldsymbol{x},t)$ being functions of fields $\phi_i(\boldsymbol{x},t)$. Its space integral is the conserved total energy functional $E[\phi]$. A large class of equations, including field equations in particle physics satisfy this. Since physical systems have energy bounded from below we can also, without lost of generality, set the minimal value reached by E equal to zero. In this framework we shall use the adjective '"localized" for those solutions to the field equations, whose energy density $\varepsilon(\boldsymbol{x},t)$ at any finite time is localized in space, i.e. it is finite in some finite region of space and falls to zero at spatial infinity sufficiently fast as to be integrable. Note that for those systems with $E[\phi_i]=0$ if and only if $\phi_i(x,t)=0$ a localized solution as defined above also has the fields themselves localized in space.

For instance, consider the following model

$$E[\phi]=\int_{-\infty}^{+\infty} dx\left[\frac{1}{2c^2}\left(\frac{\partial\phi}{\partial t}\right)^2+\frac{1}{2}\left(\frac{\partial\phi}{\partial x}\right)^2+\frac{1}{4}\phi^4\right] \tag{45}$$

It is minimized by $\phi(x,t)=0$. Localized solution of this system, if any, would asymptotically go to $\phi(x,t)=0$ as $x\to\pm\infty$ for any given t. The derivatives $\left(\frac{\partial\phi}{\partial x}\right)$ *and* $\left(\frac{\partial\phi}{\partial t}\right)$ must also vanish in this limit.

By contrast, the energy functional

\ $$E[\phi]=\int_{-\infty}^{\infty} dx\left[\frac{1}{2c^2}\left(\frac{\partial\phi}{\partial t}\right)^2+\frac{1}{2}\left(\frac{\partial\phi}{\partial t}\right)^2+\frac{1}{4}\left(\phi^2-1\right)^2\right] \tag{46}$$

is minimized by $\phi(x,t)=\pm1$ and now a localized solution must approach $\phi=\pm1 \quad as \quad x\to\pm\infty$ at any instant.

Given localization in this sense of energy density, we define a *solitary wave* as localized non-singular solution of any nonlinear field equation (or coupled equations, when several fields are involved) whose energy density, as well as being localized, has a space-time dependence of the form

$$\varepsilon(\boldsymbol{x},t)=\varepsilon(\boldsymbol{x}-\boldsymbol{u}t) \tag{47}$$

where $\boldsymbol{u}$ is some velocity vector.

In other words, the energy density should move undistorted with constant velocity. This differs from the requirement that the fields themselves have such a "travelling wave" space-time dependence.

Note that this equation defines solitary waves in one or more space dimensions. Further, any static (time-independent) <u>localized</u> solution is automatically a solitary wave, with the velocity $\boldsymbol{u}=0$. Many of the solitary waves will be obtained as static solutions. However, for systems with relativistic (or

Galilean) invariance, moving solutions are trivially obtained by boosting – transformation to a moving coordinate frame.

Let us now turn to solitons: these are solitary waves with an added requirement that generalizes a feature (ii). Consider some nonlinear equations. Let them have a solitary wave solution whose energy density is some localized function $\varepsilon_0\left(\boldsymbol{x}-\boldsymbol{u}t\right)$. Consider any other solution of this system which in the far past consists of N such solitary waves, with arbitrary velocities and positions. Then energy density of this solution will have the form

$$\varepsilon\left(\boldsymbol{x},t\right) \to \sum_{i=1}^{N}\varepsilon_0\left(\boldsymbol{x}-\boldsymbol{a}_i-\boldsymbol{u}_i t\right), \quad as \ \ t\to-\infty \tag{48a}$$

Given this configuration at $t=-\infty$, it will then evolve in time as governed by the nonlinear equations. Suppose this evolution is such that

$$\varepsilon\left(\boldsymbol{x},t\right) \to \sum_{i=1}^{N}\varepsilon_0\left(\boldsymbol{x}-\boldsymbol{a_i}-\boldsymbol{u_i}t+\boldsymbol{\delta}_i\right), \quad as \ \ t\to+\infty \tag{48b}$$

where $\boldsymbol{\delta}_i$ are some constant vectors. Then such a solitary wave is called a *soliton.* In other words, solitons are those solitary waves whose energy density profiles are asymptotically (as $t\to\infty$) restored to their original shapes and velocities. The vectors $\boldsymbol{\delta}_i$ represent the possibility that the solitons may suffer a bodily displacement compared with their precollision trajectories. This displacement should be the sole residual effect of collisions if they are to be solitons. Obviously this is a remarkable property for solutions of a nonlinear field equation to have.

While all solitons are solitary waves, the converse is clearly not true. The added requirement (ii) on solutions is very stringent. The bulk of localized solutions discussed in the physics literature seem only to be solitary waves.

(a) ***Some solitary waves in two dimensions*** (one space +one time)

We shall concentrate on static solutions in the simplest context – scalar field in two (one space +one time) dimensions. Consider first a single scalar field $\phi\left(x,t\right)$ whose dynamics is governed by the Lorentz-invariant Lagrangian density

$$\mathscr{L}\left(x,t\right)=\frac{1}{2}\left(\dot{\phi}\right)^2-\frac{1}{2}\left(\phi'\right)^2-U\left(\phi\right) \tag{49}$$

where henceforth a dot or a prime represents differentiation with respect to time t or the space variable x, respectively, and the velocity of light is set equal to one. The potential $U\left(\phi\right)$ is any positive semi-definite function of ϕ, reaching a minimum value of zero for some value or values of ϕ. Corresponding wave equation has the form

$$\Box\phi \equiv \ddot{\phi} - \phi'' = -\frac{\partial U}{\partial \phi}(x,t) \tag{50}$$

Nonlinear terms depend on the choice of $U(\phi)$. The equation conserves the total energy functional given by

$$E[\phi] = \int_{-\infty}^{\infty} dx \left[\frac{1}{2}(\dot{\phi})^2 + \frac{1}{2}(\phi')^2 + U(\phi) \right] \tag{51}$$

Let the absolute minima of $U(\phi)$, which are also its zeros, occur at M points,

$$U(\phi) = 0, \quad for \quad \phi = g^{(i)}, \quad i = 1, 2, ..., M \geq 1 \tag{52}$$

Then the energy functional is also minimized when the field $\phi(x,t)$ is constant in space-time and takes any one of these values. That is,

$$E[\phi] = 0, \quad if \ and \ only \ if \quad \phi(x,t) = g^{(i)}; \quad i = 1, 2, ..., M \tag{53}$$

As we are interested in static solutions, the equation of motion reduces to

$$\phi''(x) \equiv \frac{\partial^2 \phi}{\partial x^2} = +\frac{\partial U}{\partial \phi}(x) \tag{54}$$

A solitary wave must have finite energy and localized energy density. In view of (51) its field must approach one of the values $g^{(i)}$, as $x \to \pm\infty$. Subject to these boundary conditions, one solves the equation (54). Since this is an ordinary second order differential equation, it can easily be solved by quadrature for any $U(\phi)$.

Before we write down the solution of equation (54) in explicit form, we note a mechanical analogy which is useful also in certain situations. Formally, (54) has the form of Newton's law for a particle with coordinate ϕ moving in "time" x in the potential $\left[-U(\phi)\right]$. For a static solution its energy E is given by

$$E = \int_{-\infty}^{\infty} dx \left[\frac{1}{2}\left(\frac{d\phi}{dx}\right)^2 + U(\phi) \right]$$

Upon multiplying (54) by ϕ' and integrating once, we have

$$\int \phi'\phi'' dx = \int \frac{dU}{d\phi}\phi' dx \qquad or \qquad \frac{1}{2}(\phi')^2 = U(\phi) \tag{55}$$

Since both ϕ' and $U(\phi)$ vanish at $x \to -\infty$, the integration constant is zero. Equation, derived above, is just a virial theorem for the "analogue-particle".

We consider first the case of unique minimum at $\phi = \phi_1$, where $U(\phi_1) = 0$. The analogue- particle sees a potential $[-U(\phi)]$ as in Fig.16a, with a maximum at $\phi = \phi_1$ in the past and far future $(x = \pm\infty)$. Once the particle takes off from $\phi = \phi_1$ in either direction, it will not return. Its kinetic energy will never be zero again since its total energy will always be larger than its potential energy $[-U(\phi)]$. Consequently the particle never stop and turn back towards ϕ_1. In terms of the static field solution $\phi(x)$ this means that once we fix the boundary condition as $\phi = \phi_1$ and $\phi' = 0 \quad at \ x = -\infty$, the same condition at $x = +\infty$ will not be satisfied by a non-trivial non-singular solution, without explicitly solving Eq. (54) and independent of the details of $U(\phi)$, we see that if $U(\phi)$ has a unique absolute minimum, there can be no static solitary wave, the trivial solution $\phi(x) = \phi_1$ for all x, is permitted.

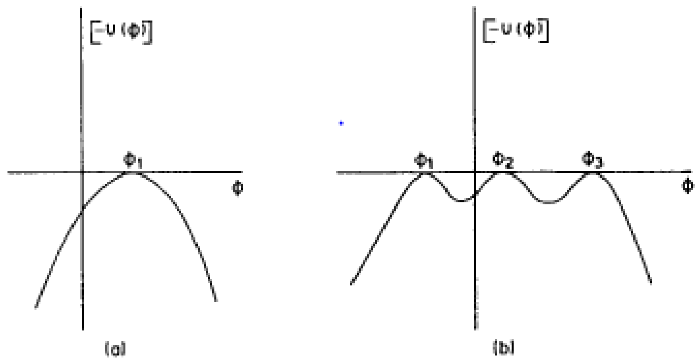


Fig.16 **(a) The potential $[-U(\phi)]$ of the "analogue-particle" when $U(\phi)$ has a unique minimum at There are no non-trivial static solutions here.**

(b) The case where $U(\phi)$ has three discrete degenerate minima. Here 4 non-trivial solutions are possible

Let $U(\phi)$ have two or more degenerate minima, where it vanishes. Fig. (b) Corresponds to an example where $U(\phi)$ has three minima at ϕ_1, ϕ_2, ϕ_3. The boundary conditions now state that the particle must leave any of these points at $x \to -\infty$ and end up at $x = +\infty$ at any one of them. This is now possible. It can take off from the top of the hill ϕ_1 at $x = -\infty$ and roll up to the top of the hill ϕ_2 asymptotically as $x \to +\infty$. Or, it can begin at ϕ_2 and end up at ϕ_3. Or, it can undergo the reverse of these two motions. There are the only four non-trivial possibilities for this example. It cannot, for instance, leave ϕ_1, go up to ϕ_2 and either return back to ϕ_1 or go to ϕ_3. Indeed, note that at ϕ_2 both $U(\phi)$ and $dU/d\phi$ vanish. Consequently, from (54) and virial theorem, both the "velocity" (ϕ') and the "acceleration" (ϕ'') vanish there. Further,

$$\phi''' = \frac{d}{dx}\left(\frac{dU(\phi)}{d\phi}\right) = \frac{d^2U}{d\phi^2}\phi' = 0$$

$$\phi^{(iv)} = \frac{d^2U}{d\phi^2}\phi'' + \frac{d^3U}{d\phi^3}(\phi')^3 = 0, \quad etc.$$

Thus, all derivatives $d^n\phi / dx^n$ vanish at ϕ_2. The particle, having left ϕ_1 can barely make it to ϕ_2 as $x \to +\infty$, where all derivatives of its motion vanish. It cannot return or proceed to ϕ_3.

Therefore, the mechanical analogy helps us conclude that if $U(\phi)$ has a unique absolute minimum, there can be no static solitary wave, and when $U(\phi)$ has n discrete degenerate minima, we can have $2(n-1)$ types of solutions, which connect any two neighboring minima, as x varies from $-\infty$ *to* $+\infty$. It is of course understood that trivial space-time independent solutions can exist in addition.

We can also explicitly solve the Eq. (50), because we have

$$\frac{d\phi}{dx} = \pm\left[2U(\phi)\right]^{1/2}$$

Upon integration one obtains explicit solution

$$x - x_0 = \pm\int_{\phi(x_0)}^{\phi(x)} \frac{d\varphi}{\left[2U(\varphi)\right]^{1/2}} \tag{56}$$

where the integration constant, x_0 is any point in space where the field has value $\phi(x_0)$.

The solution $\phi(x)$ can be obtained in principle explicitly, given at x_0 and a $\phi(x_0)$ by integrating (56) and inverting it. In practice, it may be possible to do this analytically only for some $U(\phi)$. As an illustration of this method let us consider the "kink" solution of the special model.

Lecture 8

(b) ***Kink***

The simplest topological object, the kink, arises in the theory of a single scalar field in (1+1) space-time. The action for this model is chosen in the form

$$S = \int d^2x\left[\frac{1}{2}\left(\partial_\mu\phi\right)^2 - U(\phi)\right] \tag{57}$$

where

$$U(\phi)=\frac{\lambda}{4}\left(\phi^2-\upsilon^2\right)^2, \qquad and \qquad \upsilon=m/\sqrt{\lambda} \tag{58}$$

This action is invariant under discrete transformation $\phi\to-\phi$, but this symmetry is spontaneously broken, since the classical vacuum (minima of $U(\phi)$) is

$$\phi=\pm\upsilon=\pm m/\sqrt{\lambda} \tag{59}$$

Consequently localized solutions must tend to $\pm m/\sqrt{\lambda}$ as $x\to\pm\infty$. In particular, static solutions can be of two types, as per earlier arguments. They can begin from $\phi=-\upsilon$ at $x=-\infty$ and end up with $\phi=+\upsilon$ at $x=\infty$, or vice versa. Specifically, the static solution of Eq. (56) is

$$x-x_0=\pm\int_{\phi(x_0)}^{\phi(x)}\frac{d\varphi}{\sqrt{\frac{\lambda}{2}\left(\varphi^2-m^2/\lambda\right)}} \tag{60}$$

After integration and inversion, we find the solutions

$$\phi(x)=\pm\left(m/\sqrt{\lambda}\right)\tanh\left[\left(m/\sqrt{\lambda}\right)\left(x-x_0\right)\right] \tag{61}$$

The solution with the plus sign plotted in Fig.17 (a) will be called the “kink” and one with minus sign the “antikink”. These solutions exhibit the translational invariance explicitly, since a change x_0 merely shifts the solution in space. The other symmetry $\phi\to-\phi$ together with $x\to-x$ are reflected in the relations which take on a particularly simple form when x_0 is chosen equal to zero

$$\phi_{kink}(x)=-\phi_{antikink}(x)=\phi_{antikink}(-x) \tag{62}$$

The energy density of the kink solution,

$$\varepsilon(x)=\frac{1}{2}(\phi')^2+U(\phi)=2U(\phi)=\frac{m^6}{2\lambda}\sec h^4\left[m\left(x-x_0\right)/\sqrt{2}\right] \tag{63}$$

is plotted in Fig.17(b) and is clearly localized near x_0.

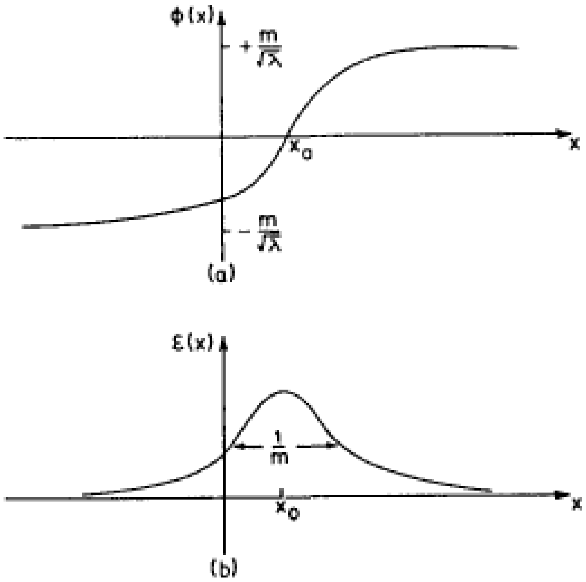


Fig.17 (a) Schematic plot of static kink solution (61), (b) The energy density of the kink (it is localized with a width $1/m$. The profile function is not localized, but the energy density does.

The total kink energy or the classical kink mass is given by

$$M_{cl} = \int_{-\infty}^{\infty} dx\varepsilon(x) = \frac{2\sqrt{2}}{3}\frac{m^2}{\lambda} \tag{64}$$

It is finite. The kink is a legitimate solitary wave. So is the antikink.

Are they solitons or no? One needs information on time dependent solutions involving several such waves. But in the present example a two-kink configuration cannot even exist with finite energy. Indeed, let consider scattering of two such solutions. The first kink must begin at $x = -\infty$ with $\phi = -m/\sqrt{\lambda}$ and tend to $\phi = +m/\sqrt{\lambda}$ on the right. If this were to be followed by a second kink, the latter would tend to $\phi = 2m/\sqrt{\lambda}$ as $x \to +\infty$. This would lead to a constant non-zero energy density as $x \to +\infty$ and hence to infinite total energy. A kink can of course be followed by an antikink, bringing the field ϕ back to $\phi = -m/\sqrt{\lambda}$. Here again numerical calculations indicate that a kink and an antikink approaching one another do not retain their shapes after collisions. Therefore, *the kink is a solitary wave but not a soliton.* It resembles a "lump" of matter in the sense that it is static, self-supporting localized packet of energy. The resemblance to an extended particle goes further, because the system is Lorentz invariant, given the static solution (63), one can Lorentz-transform it to obtain a moving kink solution. Because ϕ is a scalar field, we need only to transform the coordinate variables in (63). This gives

$$\phi_u(x,t) = \frac{m}{\sqrt{\lambda}}\tanh\left[\frac{m}{\sqrt{2}}\left(\frac{(x-x_0)-ut}{\sqrt{1-u^2}}\right)\right] \tag{65}$$

where $-1 < u < 1$ is the velocity of a kink. That this is a solution of the field equation

$$\ddot{\phi} - \phi'' = \lambda\phi^3 - m^2\phi \tag{66}$$

can be verified by substitution. Corresponding spatial width of the moving kink in (65) is $\sqrt{1-u^2}\,/\,m$, as would happen from Lorentz contraction for a lump of matter. Further, the energy of the time dependent solution (65) is

$$\begin{aligned} E[\phi_u] &= \int_{-\infty}^{\infty} dx \left[\left(\frac{m^4}{4\lambda}\frac{u^2}{1-u^2} + \frac{m^4}{4\lambda}\frac{1}{1-u^2} + \frac{m^4}{4\lambda} \right) \sec h^4 \left(\frac{m}{\sqrt{2}} \frac{x - x_0 - ut}{\sqrt{1-u^2}} \right) \right] = \\ &= \frac{2\sqrt{2}}{3}\frac{m^3}{\lambda}\frac{1}{\sqrt{1-u^2}} = \frac{M_{cl}}{\sqrt{1-u^2}} \end{aligned} \tag{67}$$

where M_{cl} is the static kink energy (64). It is the Einstein mass-energy equation for a particle. So, in the quantum version of this model the kink solution leads to a particle state. Another important feature of kink solution is that it is singular as the nonlinearity parameter λ goes to zero. Thus it cannot be obtained by mere perturbation expansion starting from the linear equation, kink (63) is non-perturbative.

(c) *Topological indices*

We are interested in non-singular finite energy solutions, of which solitary waves and solitons are minimum of $U(\phi)$ at every point of spatial infinity, in order the energy E in (51) be finite. In one space dimension spatial infinity consists of two points, $x = \pm\infty$. Consider $x = +\infty$, for instance. Let at some given instant t_0,

$$\lim_{x\to\infty} \phi(x, t_0) \equiv \phi(\infty, t_0) = \phi_1$$

where ϕ_1 has to be one of the minima of $U(\phi)$. Then, as the time develops (starting from t_0), the field $\phi(x,t)$ will change continuously with t at every x as governed by the differential equation. In particular, $\phi(\infty, t)$ will be some continuous function of t. On the other hand, since the energy of that solution is conserved and remains finite, $\phi(\infty, t)$ must always be one of the minima of $U(\phi)$, which are a discrete set. It cannot jump from ϕ_1 to another of the discrete minima if it is to vary continuously with t Therefore $\phi(\infty, t)$ must remain stationary at ϕ_1. The same arguments apply to $x = -\infty$, where $\phi(-\infty, t) = \phi_2$, must also be time-independent and minimum of $U(\phi)$, but not necessarily the same as ϕ_1 in the case of degenerate minima.

We can therefore divide the space of all finite-energy non-singular solutions into sectors, characterized by two indices, namely, the time independent values of $\phi(x=\infty)$ and $\phi(x=-\infty)$. These sectors are topologically unconnected, in the sense that fields from one sector cannot be distorted continuously into another without violating the requirement of finite energy. In particular, since the evolution is an example of continuous distortion, a field configuration from any one sector stays within that sector as time evolves. Of course, when $U(\phi)$ has a unique minimum, there is only one permissible value for both $\phi(x=\infty)$ and $\phi(x=-\infty)$ therefore only one sector of solutions exists.

Consider the kink solution. The potential has two degenerate minima at $\phi = \pm m/\sqrt{\lambda}$. Consequently, all finite-energy non-singular solutions of this system, whether static or time-dependent, fall into four topological sectors . These are characterized by the pairs of indexes $\left(-m/\sqrt{\lambda}, m/\sqrt{\lambda}\right)$, $\left(m/\sqrt{\lambda}, -m/\sqrt{\lambda}\right)$, $\left(-m/\sqrt{\lambda}, -m/\sqrt{\lambda}\right)$, $\left(m/\sqrt{\lambda}, m/\sqrt{\lambda}\right)$ respectively, which represent the values of $\left(\phi(x=-\infty), \phi(x=+\infty)\right)$. Thus the kink, the antikink and the trivial constant solutions $\phi(x) = \mp\left(m/\sqrt{\lambda}\right)$ are members of the four sectors, respectively. When a kink from the far left and an antikink from the far right approach one another, the field configuration belong to the ($-m/\sqrt{\lambda}$, $-m/\sqrt{\lambda}$) sector. Even though we may not be able to calculate easily what happens after they collide, we can be sure that the resulting field configuration will always stay in the $\left(-m/\sqrt{\lambda}, -m/\sqrt{\lambda}\right)$ sector.

A quantity sometimes called the “topological charge” is often used in the literature. It can be defined here as

$$Q = \left(\frac{\sqrt{\lambda}}{m}\right)\left[\phi(x=\infty) - \phi(x=-\infty)\right] \tag{67}$$

with an associated conserved current,

$$k^{\mu} = \left(\sqrt{\lambda}/m\right)\varepsilon^{\mu\nu}\partial_{\nu}\phi \tag{68}$$

where covariant summation notation

$$\mu,\nu = 0,1 \qquad g_{\mu\nu} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{69}$$

has been used and $\varepsilon_{\mu\nu}$ is the antisymmetric tenzor. Clearly

$$\partial_{\mu}k^{\mu} = 0 \qquad \text{and} \qquad Q = \int_{-\infty}^{\infty} dx k_0 \tag{70}$$

This Q is just difference between the two indices $\left(\sqrt{\lambda}/m\right)\phi(\infty)$ and $\left(\sqrt{\lambda}/m\right)\phi(-\infty)$. We mention it here because it is the analogue of topological indices in more complicated systems, such as gauge theories in higher dimensions. The adjective "topological" is sometimes bestowed on solitary waves which have $Q \neq 0$. Waves with $Q = 0$ are "non-topological. Thus the kink and the antikink of the considered system are topological solutions, while the trivial solutions $\phi(x) = \pm\left(m/\sqrt{\lambda}\right)$ are "non-topological" solutions. One of our conclusions here is that for a single scalar field in two dimensions, non-trivial static solutions are necessarily topological.

The topological indices, as boundary conditions, are conserved because of finiteness of energy. In many cases these indices are closely related to a certain kind of breaking of some symmetry. Suppose the Lagrangian and $U(\phi)$ are invariant under some symmetry transformation acting on $\phi(x)$. If $U(\phi)$ had a unique minimum at some $\phi = \phi_0$, then ϕ_0 itself must remain invariant under that transformation. But in order to get non-trivial topological sectors, we need to have two or more degenerate minima. In that case while the full set of minima is invariant under the transformation, each individual minimum need not be so. For instance, our considered system, which permits four topological sectors, has a $U(\phi)$ invariant under $\phi \leftrightarrow -\phi$. But its two minima are not separately invariant. Rather, they are transformed into one another. This fact has great importance in the quantum theory as well as the statistical mechanics of the field system and is called "spontaneous symmetry breaking". At this stage we merely observe the relation of non-trivial topological sectors to the existence of several degenerate minima of the potential, which in turn is connected (often, but not always) to spontaneous symmetry breaking.

exercises

1. Find the size of the kink and compare it to the Compton wavelength of an elementary excitation.
2. For large $|x|$, the kink field differs very little from the vacuum. Find this difference for large $|x|$ and show that it satisfies to the Klein-Gordon equation and decreases exponentially.
3. Find the spectrum of small perturbations about the kink, i.e. the spectrum of the eigenvalues and eigenfunctions of the operator $-\frac{d^2}{dx^2} + V(x)$ with the potential $V(\phi) = \left.\frac{\partial^2 U}{\partial \phi^2}\right]_{\phi=\phi_{kink}}$, where $U(\phi)$ is given by the Eq. (58)

Lecture 9

(d) *The sine-Gordon system*

Let us consider the chain of coupled pendula drawn in Fig. 18. The pendula are moving around a common axis, and two neighboring pendula are linked by a torsional spring. Denote by ϕ_n the rotation of pendulum n with respect to its equilibrium position. The Hamiltonian of this system is the sum of three terms:

$$H = \sum_n \frac{I}{2}\left(\frac{d\phi_n}{dt}\right)^2 + \frac{C}{2}\left(\phi_n - \phi_{n-1}\right)^2 + mgl\left(1 - \cos\phi_n\right) \tag{71}$$

The first term is the kinetic energy associated with the rotation of the pendula, where I is the moment of inertia of the pendulum with respect to the axis. The second term describes the coupling energy between neighboring pendula, due to the torsional spring having torsion constant C, while the last contribution comes from the gravitational potential energy of the pendula, l being the distance of their centers of mass to the axis, m is the mass of a pendulum, and g - the acceleration of gravity.

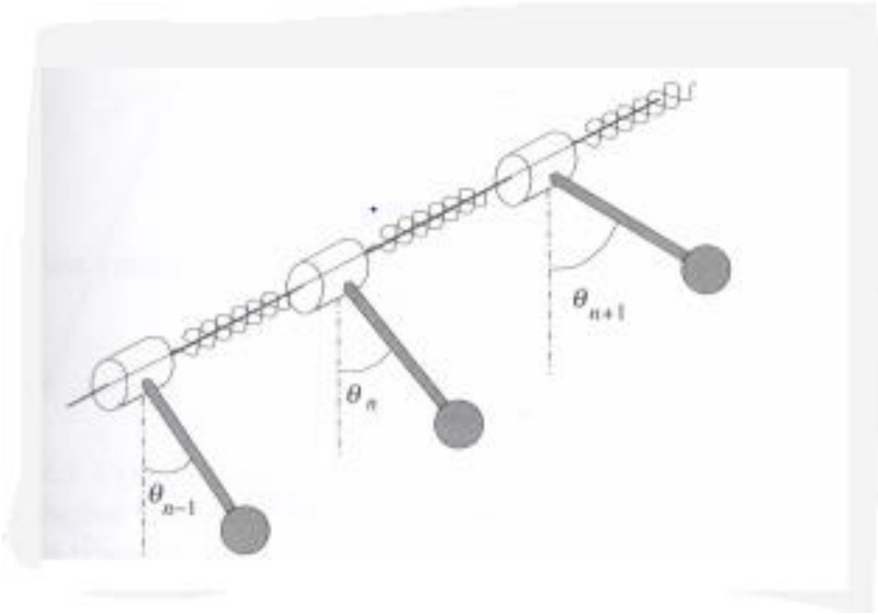


Fig.18 A chain of pendula sharing a common axis, coupled by torsional springs. In the continuum limit, the equations of motions of this device lead to the sine-Gordon equation.

Introducing the momentum $p_n = I\dot{\phi}_n$, which is the canonical conjugate to the ϕ_n , the equation of motion of the pendulum chain can be derived from Hamiltonian (71) with the Hamilton equations

$$\frac{d\phi_n}{dt} = \frac{\partial H}{\partial p_n}, \qquad \frac{dp_n}{dt} = \frac{\partial H}{\partial \phi_n}$$

They lead to the non-linear coupled differential equations

$$\ddot{\phi}_n - C\left(\phi_{n+1} + \phi_{n-1} - 2\phi_n\right) + mgl\sin\phi_n = 0 \tag{72}$$

Their exact solution is not known but an approximate solution can be derived from the continuum limit approximation, provided the coupling between adjacent pendula is strong enough to ensure that ϕ varies only slightly from one pendulum to the next.

Let us denote by a the distance between two pendula along the axis. We replace the discrete variables $\phi_n(t)$ by a function $\phi(x,t)$, where $\phi_n = \phi(x = na, t)$. The Taylor expansion of ϕ_{n+1} leads to

$$\phi_{n+1} + \phi_{n-1} - 2\phi_n \simeq a^2 \frac{\partial^2 \phi}{\partial x^2} + O\left(a^4 \frac{\partial^4 \phi}{\partial x^4}\right) \tag{73}$$

If we truncate the expansion to the lowest non-vanishing term, we obtain then the partial differential equation

$$\ddot{\phi} - c_0^2 \phi'' + \omega_0^2 \sin\phi = 0 \tag{74}$$

where $\omega_0^2 = \dfrac{mgl}{I}, \quad c_0^2 = \dfrac{Ca^2}{I}$ (square of frequency and speed) (75)

The equation (4) is known as the sine-Gordon (sG) equation. We'll see below that it is a *completely integrable equation*, which has exact soliton solution.

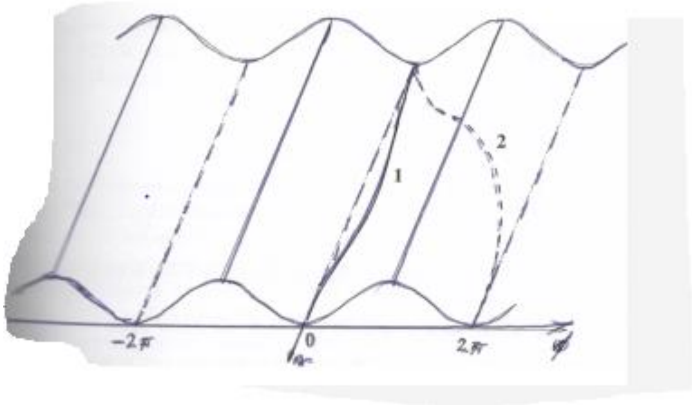


Fig.19 Topology of the potential energy landscape of the sine-Gordon model. Solid and dashed lines labeled 1 and 2 show the position of an imaginary massive elastic string, which would have the same motion as the pendulum chain in the continuum limit approximation

In order to completely figure out the potential energy of the system, one must also consider the harmonic coupling energy due to torsional springs connecting pendula. In the continuum limit the pendulum chain can be viewed as an elastic string which is massive and subjected to the undulations of the potential.

One notice, that the system has *several energetically degenerate ground states*. Indeed the ground state can be achieved with $\phi = 0$ or $\phi = 2\pi p$ (p being any integer). This was not for the KdV model – because the water in a canal only has one possible equilibrium level. This feature of the sG model suggests the existence of *several families of solutions:*

- Solutions in which the whole chain stays within a single potential valley (case 1 of Fig.)
- Solutions in which the chain moves from one valley to another one (case 2 of Fig. , which corresponds to a soliton solution)

More quantitively, solution can be distinguished by their behavior towards the boundaries $\pm\infty$:

$$\begin{aligned} &\lim_{\phi\to+\infty}\phi - \lim_{\phi\to-\infty}\phi = 0 && (in\ \ case\ \ 1) \\ &\lim_{\phi\to+\infty}\phi - \lim_{\phi\to-\infty}\phi = 2\pi p, \quad p \neq 0; && (in\ \ \ case\ \ 2) \end{aligned} \tag{76}$$

These two solution are said *topologically different* because their differences is a property of the solution as a whole. Indeed, if one looks at the solution for $|x| \to \infty$, a local view does not make any difference between the two: one sees pendula at rest in their minimal energy state. It is only by moving the whole pendulum chain that one can notice that there is a full turn from one end to other in case 2.

(e) ***Soliton solutions in sG system***

In order to derive the solutions of the sG equation, one must notice, that the equation is preserved by a Lorentz transform relative to speed c_0. Therefore it is sufficient to look for static solutions, from which solutions moving in velocity u can be derived with a Lorentz transform, as well as in kink case. However, as in KdV, soliton solutions can also be obtained by looking for permanent profile solutions moving at velocity u, i.e. solutions which only depends on a single variable $\xi = x - ut$. For such permanent profile solutions the sG equation becomes

$$u^2\phi''_{\xi\xi} - c_0^2\phi''_{\xi\xi} + \omega_0^2 \sin\phi = 0 \tag{77}$$

or

$$\phi''_{\xi\xi} = \frac{\omega_0^2}{c_0^2 - u^2}\sin\phi \tag{78}$$

Multiplying by ϕ'_ξ as in kink case and integrating with respect to ξ, we get

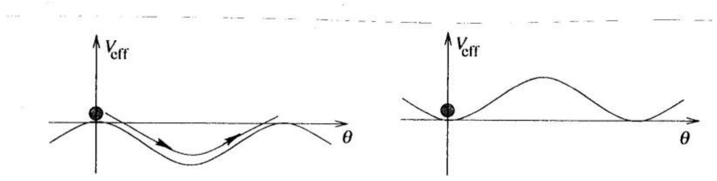
$$\frac{1}{2}\left(\frac{d\phi}{d\xi}\right)^2 = \frac{\omega_0^2}{c_0^2 - u^2}\cos\phi + C_1 \tag{I79}$$

The integration constant C_1 is determined by the boundary condition that we impose on the solution. Since we are looking for a soliton, i.e. spatially localized solution, we must have $\phi(\xi) \to 0 \pmod{2\pi}, \quad \mathit{for} \quad |\xi| \to \infty$, because at infinity the pendula must be in one of their ground states. For the same reason we impose $d\phi / d\xi \to 0, \quad \mathit{if} \quad |\xi| \to \infty$, which leads to $C_1 = \omega_0^2 / \left(c_0^2 - u^2\right)$, and therefore

$$\frac{1}{2}\left(\phi'_\xi\right)^2 - \frac{\omega_0^2}{c_0^2 - u^2}\left(1 - \cos\phi\right) = 0 \tag{80}$$

As in the kink case, we can consider this relation as the sum of kinetic energy (in the “pseudo—time” ξ) and the potential energy (of an “analogue-particle”). Thus the solution $\phi(\xi)$ describes the motion of this particle, having zero total energy, in the potential

$$V_{eff}(\phi) = -\frac{\omega_0^2}{c_0^2 - u^2}\left(1 - \cos\phi\right) \tag{81}$$



$c_0^2 - u^2 > 0$ $\qquad\qquad$ $c_0^2 - u^2 < 0$

Fig. 20. Search for the possible solutions of the sG equation by studing a fictious particle mobile in the pseudo-potential $V_{eff}(\phi)$.

Fig. (20) shows that, for $c_0^2 - u^2 > 0$ there is a possible motion for an analogue- particle leaving $\phi = 0$ at rest. It can reach $\phi = 2\pi\,(or\ \phi = -2\pi)$ with a vanishing "velocity" ϕ'_ξ after an infinite fictitious "time" ξ . Consequently, for $c_0^2 - u^2 < 0$, a particle initially at rest in $\phi = 0$ cannot move. This analysis shows that solitons can only travel at speeds *smaller* than c_0. Moreover it indicates that there are no permanent profile solutions which start and end in the same potential valley.

For $c_0^2 - u^2 > 0$, the solution can be obtained from Eq. (80) and we get

$$\phi = 4\arctan\exp\left[\pm\frac{\omega_0}{c_0}\frac{\xi - \xi_0}{\sqrt{1 - u^2 / c_0^2}}\right], \qquad \xi = x - ut \qquad (82)$$

The arbitrary integration constant ξ_0 determines the position of the soliton at time $t = 0$. The solution exhibits the characteristic expression associated with the Lorentz invariance, as well as the validity condition $u^2 < c_0^2$. The solutions "soliton" (with a + sign) and "ant soliton" (- sign) are plotted in Fig.21.

The soliton interpolates between two different states of the system which have the same energy. Solitons and antisolitons differ by their *topological charge* defined by:

$$Q = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{\partial\phi}{\partial x}dx = \frac{1}{2\pi}\left[\lim_{t\to\infty}\phi(x,t) - \lim_{t\to-\infty}\phi(x,t)\right] \qquad (83)$$

Which is equal +1 for a soliton and -1 for an antisoliton. The conservation of the topological charge explains the exceptional stability of topological solitons. They are much more stable than the non-topological solitons of KdV equation. In an infinite medium, perturbations can modify the speed of a soliton, or even bring it to rest, but they cannot kill it because it would imply a change of the topological charge.

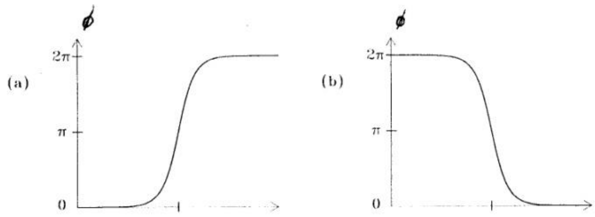


Fig.21 **Soliton (a) and antisoliton (b) solutions of the sG equation**

(f) ***Energy of soliton***

In the continuum limit approximation the contribution of a single cell, divided by the cell spacing, gives the energy density – the Hamiltonian per unit length

$$\mathcal{H}\left(x,t\right)=\frac{1}{a}\left[\frac{I}{2}\left(\dot{\phi}\right)^{2}+\frac{Ca^{2}}{2}\left(\phi'\right)^{2}+mgl\left(1-\cos\phi\right)\right] \qquad (84)$$

If we are interested in the expression of $\mathcal{H}$ for a soliton, it is again convenient to introduce the variable $\xi=x-ut$ such that $\phi\left(x,t\right)=\phi\left(\xi\right)$. Then

$$\mathcal{H}\left(\xi,t\right)=\frac{1}{a}\left[\frac{Iu^{2}}{2}\left(\phi'_{\xi}\right)^{2}+\frac{Ca^{2}}{2}\left(\phi'_{\xi}\right)^{2}+mgl\left(1-\cos\phi\right)\right] \qquad (85)$$

Using the definition of c_0^2 and Eq. (80), all the terms can be written as a function of ϕ'_{ξ}, and substitution the obtained solution (82), leads to

$$\mathcal{H}=\frac{Ic_0^2}{a}\left(\phi'_{\xi}\right)^{2}=\frac{4I\omega_0^2}{a\left(1-u^2/c_0^2\right)}\sec h^2\frac{\omega_0\left(x-ut\right)}{\sqrt{c_0^2-u^2}}$$

This expression containing $\sec h^2$ does indeed describe an energy density *localized* around the center of the soliton. Because of this property soliton often is called as “quasi-particle”

Because $$\int_{-\infty}^{\infty}dx\cdot\operatorname{sech}^2 x=2$$

an integration over space gives the energy of the soliton

$$E=\frac{8I\omega_0 c_0}{a\sqrt{1-u^2/c_0^2}} \qquad (86)$$

which has a standard “relativistic” expression, with respect to the speed c_0 for a particle of mass

$$M=8I\omega_0/c_0a. \qquad (87)$$

It is interesting to note that the energy of the soliton can be calculated even if we do not know the analytical expression of the solution. Let us restart from the energy density. The energy is

$$E = \frac{Ic_0^2}{a}\int_{-\infty}^{\infty}\left(\frac{d\phi}{d\xi}\right)\left(\frac{d\phi}{d\xi}\right)d\xi$$

and can be calculated by replacing one of the two factors by its analytic expression deduced from the solution (82). We get

$$E = \frac{Ic_0^2}{a}\frac{\sqrt{2}\omega_0}{\sqrt{c_0^2-u^2}}\int_{-\infty}^{\infty}\sqrt{1-\cos\phi}\left(\frac{d\phi}{d\xi}\right)d\xi = \frac{Ic_0^2}{a}\frac{\sqrt{2}\omega_0}{\sqrt{c_0^2-u^2}}\int_{0}^{2\pi}\sqrt{1-\cos\phi}\,d\phi$$

The integration over ϕ is easily to carry out and we deduce the above expression (86).

Lecture 10

(g) ***Other models, related to soliton equations***

1. FPU problem

We saw above that the sine-Gordon system has soliton solutions as against the kink model. The SG system has been in the study of a wide range of phenomena, including propagation of crystal dislocations, of splay waves in membranes, of magnetic flux in Josephson lines, Bloch wall motion in magnetic crystals, as well as two-dimensional models of elementary particles. There are specific applications in solid state physics, such as the Fermi-Pasta-Ulam (FPU) problem: how a crystal evolves toward thermal equilibrium by simulating a chain of particles of mass unity, linked by a quadratic interaction potential, but also by a weak nonlinear interaction.

The question of interest was: why do solid have finite heat conductivity? The solid is modeled by a one-dimensional lattice. In 1914 Debye had suggested that the finiteness of the thermal conductivity of a lattice is due to the anharmonicity of the nonlinear forces in the strings. If the force is linear (Hook's law), energy is carried unhindered by the independent fundamental normal modes of propagation. The effective thermal conductivity is infinite. no thermal gradient is required to push the heat through the lattice from one end to another and no diffusion equation obtains. Debye thought that if the lattice were weakly nonlinear, the normal modes (calculated from the linearized spring) would interact due to the nonlinearity and thereby hinder the propagation of energy. The net effect of many such nonlinear interactions (phonon collisions) would manifest itself in a diffusion equation with a finite transport coefficient.

This suggestion motivated Fermi, Pasta and Ulam to undertake a numerical study of the one-dimensional anharmonic lattice . They argued that a smooth initial state in which all the energy was in the lowest mode or the first few lowest modes would eventually relax to a state of statistical equilibrium due to nonlinear couplings. In that state, energy would be equidistributed among all modes on the average.

The model used by FPU to describe their one-dimensional lattice of length L consists of a row of $N-1$ identical masses each connected to the next and the end ones to fixed boundaries by N nonlinear springs of length L. Those springs when compressed or extended by an amount Δ exert a force

$$F = k\left(\Delta + \alpha\Delta^2\right)$$

where k is the linear spring constant and α, taken positive, measures the strength of nonlinearity. The equations governing the dynamics of this lattice are

$$\begin{aligned} &m u_{tt}^{i} = k\left(u_{i+1} - 2u_i + u_{i-1}\right)\left(1 + \alpha\left(u_{i+1} - u_{i-1}\right)\right), \qquad i = 1, 2, ..., N-1 \\ &u_0 = u_N = 0 \end{aligned}$$

This one-dimensional system is described by the Hamiltonian

$$H = \sum_{i=0}^{N-1} \frac{p_i^2}{2} + \sum_{i=0}^{N-1} K\left(u_{i+1} - u_i\right)^2 + \frac{K\alpha}{3}\sum_{i=0}^{N-1}\left(u_{i+1} - u_i\right)^3 \tag{88}$$

where u_i is the displacement along a chain of atom i with respect to its equilibrium position and p_i is its momentum. The coefficient α is the measure of nonlinearity. The two ends of chain were assumed to be fixed, i.e. $u_0 = u_N = 0$. Introducing new displacements and their frequencies by

$$A_k = \sqrt{\frac{2}{N}}\sum_{i=0}^{N-1} u_i \sin\left(ik\pi / N\right) \quad and \quad \omega_k^2 = 4K\sin^2\left(k\pi / 2N\right),$$

respectively, the Hamiltonian (88) reduces to the form

$$H = \frac{1}{2}\sum_{k}\left(\dot{A}_k^2 + \omega_k^2 A_k^2\right) + \alpha\sum_{k,l,m} c_{klm} A_k A_l A_m \tag{89}$$

The last term, due to nonlinearity, leads to a coupling between the “normal” modes. FPU studied numerically influence of this non-linear term on the normal modes. They noticed that the system, after remaining in as steady state for a while, had then departed from it. To their great surprise, after 157 periods of the mode $k = 1$, almost all energy was back in the lowest frequency mode. This mystery, that nonlinearity was seemingly nicer than expected, was known as the Fermi-Pasta-Ulam Problem. This highly remarkable result, known also as FPU paradox, shows that nonlinearity is not enough to guarantee the equipartition of energy. To understand it, it is necessary to stop thinking in terms of linear normal modes, and to consider the nonlinearity intrinsically. It also means that one should stop thinking in Fourier space and come back to real space. A “mode” is a localized excitation in Fourier space, but it is fully delocalized in real space. Conversely a soliton is localized in real space, but extended in Fourier space.

The solution of the FPU paradox was found ten years later by Zabusky and Kruskal in terms of solitons. They studied the equations of motion derived from the Hamiltonian (88)

$$\ddot{u}_i = K\left(u_{i+1} + u_{i-1} - 2u_i\right) + K\alpha\left[\left(u_{i+1} - u_i\right)^2 - \left(u_i - u_{i-1}\right)^2\right] \tag{90}$$

Zabusky and Kruskal considered continuous limit and get the KdV equation.

2. **Frenkel-Kontorova (FK) model**

A model containing the essentials of the physics of a dislocation was proposed in 1939 by Frenkel and Kontorova. It describes the dynamics of a line of atoms above the slip plane

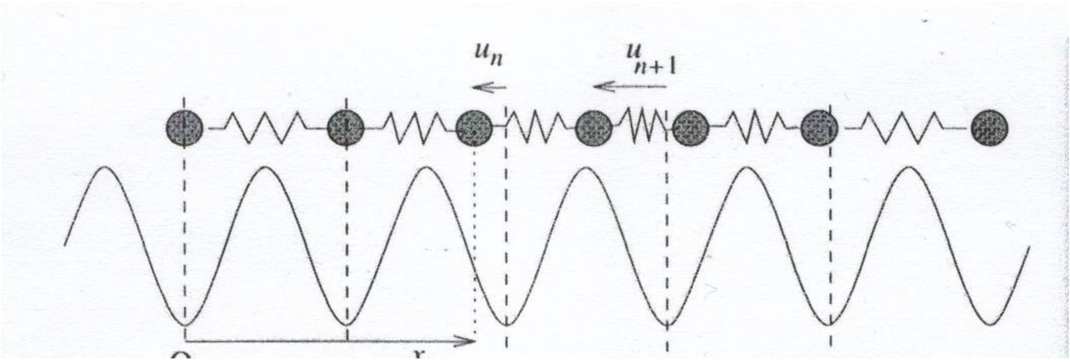


Fig.22 **The Frenkel-Kontorova model of an edge dislocation**

The position of the atom of index n is measured with respect to a fixed origin, chosen to be at the equilibrium position of one of particular atom in a perfect crystal (Fig.21). It can also be given by its displacement with respect to its equilibrium position, $u_n = x_n - na$, where a is the lattice spacing along the line of atoms. As atom is subjected to the potential $V\left(u_n\right)$ created by the atoms which are below the slip plane. This "substrate potential" has the periodicity of the lattice and the FK model chooses the simplest periodic function

$$V\left(u_n\right) = V_0\left(1 - \cos\frac{2\pi u_n}{a}\right) \tag{91}$$

The model must also take into account the interaction of the atoms along the line. The substrate potential $V\left(u_n\right)$ cannot ignore non-linearity because we want to describe motion which may be as large as the period of the potential. However, harmonic approximation is possible for the interaction potential between atoms $n-1$ and n because it depends on the relative displacement of two neighboring atom, which is small with respect to the lattice spacing, even in the core of a lattice displacement. It is written as

$$W\left(u_{n-1,}u_n\right) = \frac{C}{2}\left(u_n - u_{n-1}\right)^2 \tag{92}$$

So that the Hamiltonian of the model is

$$H = \sum_n \frac{p_n^2}{2m} + \frac{C}{2}\left(u_n - u_{n-1}\right)^2 + V_0\left(1 - \cos\frac{2\pi u_n}{a}\right) \tag{93}$$

This model contains the basic ingredients for soliton solutions, the nonlinearity of the substrate potential and the cooperativity coming from the interatomic interactions. It can be shown that indeed solitons exist in this system. Analogy with the sG model (71) is evident.

The equations of motion of the atoms, which derive from this Hamiltonian, are

$$m\ddot{u}_n = C\left(u_{n+1} + u_{n-1} - 2u_n\right) - \frac{2\pi V_0}{a}\sin\frac{2\pi u_n}{a} \tag{94}$$

As in FPU problem, we get a set of coupled nonlinear differential equations. This is a common situation in solid state physics. The system, although it has a simple form does not have any known analytic solution. As in the FPU case, approximations are required to solve it. As in the FPU case, we can use a continuum limit approximation by replacing the set of discrete variables $u_n(t)$ by the continuous function $u(x,t)$, so that $u_n(t) = u(x = na, t)$. Then expanding $u_{n\pm 1}(t)$ around $u_n(t)$ we get

$$u_{n+1}(t) = u\left((n \pm 1)a, t\right) = u(na,t) \pm a\frac{\partial u}{\partial x}(na,t) + \frac{a^2}{2}\frac{\partial^2 u}{\partial x^2}(na,t) + \ldots \tag{95}$$

If we truncate the expansion at order 2, the set of equations (94) becomes

$$\ddot{u} - \frac{Ca^2}{m}u'' + \frac{2\pi V_0}{ma}\sin\frac{2\pi u}{a} = 0 \tag{96}$$

We are faced to the sG equation

$$\ddot{\phi} - c_0^2\phi'' + \omega_0^2 \sin\phi = 0 \tag{97}$$

where we have used the definitions

$$\phi(x,t) = 2\pi u(x,t)/a, \qquad c_0^2 = Ca^2/m, \qquad \omega_0^2 = 4\pi^2 V_0/\left(ma^2\right) \tag{98}$$

This result suggests therefore that the solutions of the sG equation (97) could describe the dislocations in solid state.

(h) *A particle physics approach to the sG system*

The starting point is a Lagrangian density for the single scalar field in (1+1) dimensions:

$$\mathscr{L}(x,t) = \frac{1}{2}\left(\partial_\mu\phi\right)\left(\partial^\mu\phi\right) + \frac{m^4}{\lambda}\left[\cos\left(\frac{\sqrt{\lambda}}{m}\right)\phi - 1\right] \tag{99}$$

The field equation arising from (99) is the sine-Gordon equation:

$$\Box\phi + \frac{m^3}{\sqrt{\lambda}}\sin\left(\frac{\sqrt{\lambda}}{m}\right)\phi = 0 \tag{100}$$

Let us change the variables as follows

$$\bar{x} = mx, \qquad \bar{t} = mt \qquad and \qquad \bar{\phi} = \sqrt{\lambda / m}\phi$$

The equation of motion becomes

$$\frac{\partial^2\bar{\phi}}{\partial\bar{t}^2} - \frac{\partial^2\bar{\phi}}{\partial\bar{x}^2} + \sin\bar{\phi}\left(\bar{x},\bar{t}\right) = 0 \tag{101}$$

and the conserved energy is

$$E = \frac{m^3}{\lambda}\int d\bar{x}\left[\frac{1}{2}\left(\frac{\partial\bar{\phi}}{\partial\bar{t}}\right)^2 + \frac{1}{2}\left(\frac{\partial\bar{\phi}}{\partial\bar{x}^2}\right)^2 + \left(1-\cos\bar{\phi}\right)\right] \tag{102}$$

This Lagrangian and the field equation are unchanged under the discrete symmetries

$$\bar{\phi}\left(x,t\right) \to -\bar{\phi}\left(x,t\right)$$

and

$$\bar{\phi}\left(\bar{x},\bar{t}\right) \to \bar{\phi}\left(\bar{x},\bar{t}\right) + 2N\pi, \qquad N = ...-2,-1,0,1,2,... \tag{103}$$

Consistent with these symmetries, the energy vanishes at the absolute minima of

$$U\left(\bar{\phi}\right) = 1-\cos\bar{\phi}\ , \tag{104}$$

which are

$$\bar{\phi}\left(\bar{x},\bar{t}\right) = 2N\pi \tag{105}$$

All finite energy configurations can be divided into an infinite number of topological sectors, each characterized by a conserved pair of integer indices $\left(N_1, N_2\right)$, corresponding to the asymptotic values $2N_1\pi$ *and* $2N_2\pi$ that the field must approach as $\bar{x}$ tends to $-\infty$ *and* $+\infty$, respectively. The topological charge is

$$Q \equiv N_1 - N_2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} d\bar{x}\,\frac{\partial\bar{\phi}}{\partial\bar{x}} \tag{106}$$

Let us begin with static localized solutions. In one space dimension, static solutions must connect only neighboring minima of $U\left(\bar{\phi}\right)$. That is, they must carry $Q = \pm 1$. Explicit solutions are easily obtained using Eq. (56);

$$\bar{x} - \bar{x}_0 = \pm\int_{\bar{\phi}(\bar{x}_0)}^{\bar{\phi}(\bar{x})}\frac{d\varphi}{\sqrt{2U\left(\varphi\right)}} = \pm\int_{\bar{\phi}(\bar{x}_0)}^{\bar{\phi}(\bar{x})}\frac{d\varphi}{2\sin\varphi/2} \tag{107}$$

This integral can be easily performed by substitution $t = \tan\left(\varphi/4\right)$ and equals to $\sqrt{2}\ln\tan\left(\varphi/4\right)$ therefore we get

$$\bar{\phi}\left(x\right) = 4\tan^{-1}\left[\exp\left(\bar{x}-\bar{x}_0\right)\right] = \phi_{sol}\left(\bar{x}-\bar{x}_0\right) \tag{108a}$$

and

$$\bar{\phi}\left(x\right) = -4\tan^{-1}\left[\exp\left(\bar{x}-\bar{x}_0\right)\right] = \phi_{antisol}\left(\bar{x}-\bar{x}_0\right) = -\bar{\phi}_{sol} \tag{108b}$$

The solution with the plus sign here goes from $\bar{\phi} = 0$ to $\bar{\phi} = 2\pi$ (Fig.22) or equivalently from 2π to 4π etc. It corresponds to $Q = 1$, and is often called the soliton of the system. The other solution has

$Q=-1$ and called the antisoliton. Each has energy $M_S = 8m^2/\lambda$ (calculate it). Moving soliton solution can be obtained on Lorentz-transforming (108a), i.e. replacing $(\bar{x}-\bar{x}_0)$ by $(\bar{x}-\bar{x}_0-ut)/\sqrt{1-u^2}$. The solution (108a) is roughly similar in shape to the "kink", although the function, in detail, is different.

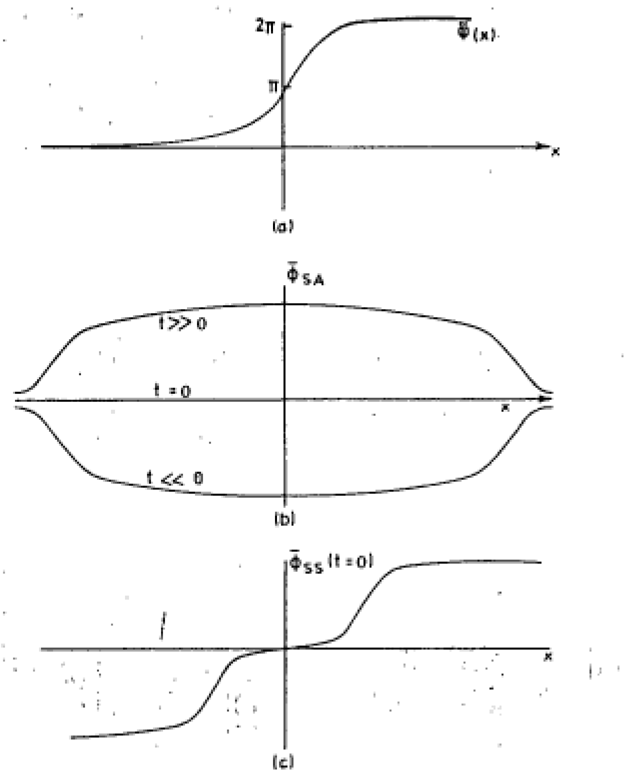


Fig.23. **A sketch of the sG soliton, (108a); (b) Three profiles of the soliton – antisoliton scattering solution $\bar{\phi}_{SA}(x,t)$ at t large negative, $t=0$ and t large positive; at $t=0$ $\bar{\phi}_{SA}$ vanishes; (c) A sketch of the soliton – soliton solution at $t=0$.**

However, unlike the kink, we assert that the solution (108a) is a genuine soliton as per the stringent requirement, given earlier. Similarly the $Q=-1$ solution (108b) is also a genuine soliton. It is called an antisoliton here partly to distinguish it from the $Q=1$ solution and partly because it is related to the latter by the symmetry $\bar{\phi} \leftrightarrow -\bar{\phi}$.

In fact this system permits a third type of soliton called the doublet or breather. Altogether then, the field equation (101) yields not just one but three different types of solitons.
It is easy to verify by substitution that the following function

$$\bar{\phi}_{SA}(x,t) = 4\tan^{-1}\left(\frac{\sinh\left(u\bar{t}/\sqrt{1-u^2}\right)}{u\cosh\left(\bar{x}/\sqrt{1-u^2}\right)}\right) \tag{109}$$

satisfies to the Eq. (101). Its asymptotic behavior in time can be extracted quite easily, by using the relation $\tan^{-1}(1/z) = \pi/2 - \tan^{-1}(z)$, to yield

$$\bar{\phi}_{SA}(x,t) \underset{t\to-\infty}{\rightarrow} 4\tan^{-1}\left[\exp\left(\frac{\bar{x}+u(\bar{t}+\Lambda/2)}{\sqrt{1-u^2}}\right)\right]-4\tan^{-1}\left[\exp\left(\frac{\bar{x}-u(\bar{t}+\Lambda/2)}{\sqrt{1-u^2}}\right)\right]=$$

$$=\bar{\phi}_{sol}\left(\frac{\bar{x}+u(\bar{t}+\Lambda/2)}{\sqrt{1-u^2}}\right)+\bar{\phi}_{antisol}\left(\frac{\bar{x}-u(\bar{t}+\Lambda/2)}{\sqrt{1-u^2}}\right) \quad (110a)$$

$$\Lambda \equiv \frac{(1-u^2)}{u}\ln u$$

The solution therefore corresponds to a soliton and antisoliton far apart and approaching one another with relative velocity $2u$, in the distant past.

Similarly one can check that

$$\bar{\phi}_{SA}(x,t) \underset{t\to\infty}{\rightarrow} \bar{\phi}_{sol}\left(\frac{\bar{x}+u(\bar{t}-\Lambda/2)}{\sqrt{1-u^2}}\right)+\bar{\phi}_{antisol}\left(\frac{\bar{x}-u(\bar{t}-\Lambda/2)}{\sqrt{1-u^2}}\right) \quad (110b)$$

We see that in distant future the solution $\bar{\phi}_{SA}$ corresponds to the same soliton-antisoliton pair, with the same shapes and velocities! The only change from the initial configuration (109) is the time delay Λ, which remains as the sole residual effect of the collision between the soliton and the antisoliton. As they approach one another, they tend to annihilate each other until at $\bar{t}=0$, the field vanishes everywhere (Fig.22(b)). But it re-emerges for positive $\bar{t}$, and asymptotically grows and separates into the same pair as if the collision had never taken place, except for the time delay.

There is a similar two soliton exact solution,

$$\bar{\phi}_{SS}(x,t)=4\tan^{-1}\left(\frac{u\sinh\left(\bar{x}/\sqrt{1-u^2}\right)}{\cosh\left(u\bar{t}/\sqrt{1-u^2}\right)}\right) \quad (111)$$

which is depicted in Fig.22(c). At any instant $\bar{t}$ it goes from -2π *to* $+2\pi$ as $\bar{x}$ goes from $-\infty$ *to* $+\infty$, and consequently belongs to the sector $Q=2$.

Finally, by the $\bar{\phi}\leftrightarrow-\bar{\phi}$ symmetry, $\bar{\phi}_{AA}\equiv-\bar{\phi}_{SS}$ is the antisoliton-antisoliton solution. These exact solutions (109) and (111) indicate that what we have termed the soliton and the antisoliton of this system may both be genuine solitons. But (109) and (111) correspond to cases where only two of these objects collide. It may be verified that

$$\bar{\phi}_{\upsilon}(\bar{x},\bar{t})=4\tan^{-1}\left[\frac{\sin(\upsilon t/\sqrt{1+\upsilon^2})}{\upsilon\cosh\left(\bar{x}/\sqrt{1+\upsilon^2}\right)}\right] \quad (112)$$

is also a solution of the field equations. Considering that $u=i\upsilon$ represented the asymptotic velocities of the soliton and the antisoliton in $\bar{\phi}_{SA}$, one can interpret the doublet in (112) as a “bound” solution of a soliton-antisoliton pair. The doublet is clearly a periodic solution with period

$$\bar{\tau}=\left(2\pi\sqrt{1+\upsilon^2}\right)/\upsilon$$

The soliton and antisoliton oscillate with respect to one another with this period.

(i) ***Backlund transformations***

Another important feature of this system is the presence of Becklund transformations. As applied to the sG system, these transformations provide a way of generating N-soliton solutions, starting from solutions with fewer solitons. Furthermore, only first order differential equations need to be solved. The essence of Becklund transformation is the following: Suppose that we have two uncoupled partial differential equations in two independent variables x and t, for the two functions u and υ, the two equations are expressed as

$$P(u)=0 \qquad and \qquad Q(\upsilon)=0 \tag{113}$$

where P *and* Q are two operators, which are in general nonlinear. Let $R_i=0$ be a pair of relations

$$R_i\left(u,\upsilon,u_x,\upsilon_x,u_t,\upsilon_t,...;\mathrm{x},\mathrm{t}\right)=0, \qquad i=1,2 \tag{114}$$

between the two functions u *and* υ. Then $R_i=0$ is a *Backlund transformation* if it is integrable for υ when $P(u)=0$ and if the resulting υ is a solution of $Q(u)=0$ and vice versa. If $P=Q$, so that u *and* υ satisfy the *same* equation, then $R_i=0$ is called an *auto-Becklund transformation.* Of course, this approach to the solution of the equations $P(u)=0$ and $Q(u)=0$ is normally only useful if the relations $R_i=0$ are, in some sense, simpler than the original equations (113).

For the sG system corresponding Becklund transformation is formulated as follows: Let us write the sG equation in Light-cone variables $x_\pm=\frac{1}{2}(x\pm t)$, corresponding derivatives are $\partial_\pm=\partial/\partial x_\pm$, in terms of which the sG equation becomes

$$\partial_-\partial_+u=\sin u \quad \text{or} \quad u_{-+}=\sin u \tag{115}$$

Consider now the following pair of equations

$$\frac{1}{2}(u+\upsilon)_+=a\sin\left(\frac{u-\upsilon}{2}\right) \qquad \frac{1}{2}(u-\upsilon)_+=\frac{1}{a}\sin\left(\frac{u+\upsilon}{2}\right) \tag{116}$$

where a is a non-zero arbitrary real constant called the Backlund parameter.

Relations (116) are known as a Becklund transformation and may be thought of as determining the field u, given the field υ. Now let us construct the cross-derivatives

$$\begin{aligned}\frac{1}{2}(u+\upsilon)_{-+}&=\sin\left(\frac{u+\upsilon}{2}\right)\cos\left(\frac{u-\upsilon}{2}\right)\\ \frac{1}{2}(u-\upsilon)_{+-}&=\sin\left(\frac{u-\upsilon}{2}\right)\cos\left(\frac{u+\upsilon}{2}\right)\end{aligned} \tag{117}$$

The compatibility condition means, that the order of derivatives is not essential, $\partial_+\partial_-f=\partial_-\partial_+f$. Then, by adding and subtracting equations in (117) and using simple trigonometric relations one can derive the following equations

$$u_{+-}=\sin u, \quad and \quad \upsilon_{+-}=\sin\upsilon \tag{118}$$

So, both u *and* υ satisfy the sine-Gordon equations (115). Therefore equations (116) are an auto-Becklund transformation for equation (115).

Thus, the Becklund transformation is a mapping between solutions of the sine-Gordon equation and can be used to generate new solutions from known solutions. When applied repeatedly, the Becklund transformation can give the breather and the multisoliton solutions of sine-Gordon equation. It can also be applied to the KdV and other equations. The main difficulty is in finding functions P and Q.

As an example, remember that the sG equation has a trivial solution, $u(x,t)=0$ for all x *and* t and use this trivial solution to generate a non-trivial one. Let us choose $\upsilon = 0$, then the Becklund transformation, (116) becomes

$$u_+ = 2a\sin\left(\frac{1}{2}u\right) \qquad\qquad u_- = \frac{2}{a}\sin\left(\frac{1}{2}u\right)$$

These two equations may be integrating to give

$$2ax_+ = \int^u \frac{du}{\sin\left(\frac{u}{2}\right)} = 2\log\left|\tan\frac{u}{4}\right| + f(x_-), \qquad\qquad \frac{2x_-}{a} = 2\log\left|\tan\frac{u}{4}\right| + g(x_+)$$

respectively, where f and g are arbitrary functions. Thus, for consistency, we must have

$$\tan\left(\frac{u}{4}\right) = C\exp\left(ax_+ + x_-/a\right)$$

or

$$u = 4\arctan\left\{C\exp\left(ax_+ + x_-/a\right)\right\} \tag{119}$$

In terms of original coordinates $\{\bar{x},\bar{t}\}$, this gives

$$\bar{\phi}(\bar{x},\bar{t}) = 4\tan^{-1}\left[\exp\left(\frac{\bar{x}-u\bar{t}}{\sqrt{1-u^2}}\right)\right], \tag{120}$$

with $\quad u = \dfrac{1-a^2}{1+a^2}$.

This is just the one-soliton solution (108a), described in a frame where it moves with velocity u. Thus, starting from the no-soliton solution, the Becklund transformation generates the one-soliton solution The real power of the Becklund transformation is that it leads to a purely algebraic method of constructing multisoliton solutions, evading the task of having to explicitly integrate Eqs. (116), which may be tricky for a complicated seed solution υ.

Lecture 11

Scaling arguments and theorems on the absence of solitons

Till now we considered field theories only in 1+1 dimensions. Naturally, it is interesting to see what happens in more dimensions. We are interested only in time independent field configurations with finite energy.

The vacuum, which is spatially constant and has the minimal energy of all fields, belongs to trivial case. More generally, we may ask if there are any non-trivial stationary points of the energy.

By applying scale arguments it is possible to show that there are no non-trivial static solutions of the field equations in a number of models in $(d+1)$-dimensional space-time with $d>1$. These arguments apply not only to a stable solutions of the soliton type but also to unstable static solutions.

A simple and important <u>*non-existence*</u> theorem is due to Derrick (1964). He noted that in many models the variation of the energy functional for static fields with respect to a spatial rescaling is never zero for any non-vacuum field configuration. But a field configuration which is a stationary point of the energy should be stationary against all variations including spatial rescaling. Therefore, in such theories there can be no static finite energy solutions of the field equation, except the vacuum. In particular, there are no topological solitons.

More precisely: in R^d a spatial rescaling is a map $\boldsymbol{x} \to \lambda \boldsymbol{x}, \quad with \quad \lambda > 0.$ Let $\psi(\boldsymbol{x})$ be a field configuration, with any kind of field or multiplet of fields, and let $\psi^{(\lambda)}(\boldsymbol{x}),\ 0<\lambda<\infty$, be the 1-parameter family of field configurations, obtained from $\psi(\boldsymbol{x})$ by applying the map $\boldsymbol{x} \to \lambda \boldsymbol{x}$. We shall clarify how $\psi^{(\lambda)}(\boldsymbol{x})$ is related to $\psi(\boldsymbol{x})$ below. In any case of the field configuration, $\psi^{(\lambda)}(\boldsymbol{x})$ let's denote by $E\left(\psi^{(\lambda)}\right)=E(\lambda)$ the corresponding energy, as a function of scaling parameter λ. Then the Derick theorem reads: *Suppose that for arbitrary, finite energy configuration* $\psi(\boldsymbol{x})$*, which is not the vacuum, the function* $E(\lambda)$ *has no stationary point. Then the theory has no static solutions of the field equation with finite energy, other than the vacuum.*

The usefulness of this non-existence theorem depends on defining $\psi^{(\lambda)}$ in an appropriate way so that it is easy to determine $E(\lambda)$.

For better understanding of this theorem let us consider some appropriate examples.

Let us consider first the theory of n scalar fields $\varphi^a,\ a=1,2,...,n,$ in $(d+1)$-dimensional space-time. We shall write the Lagrangian in a quite general form

$$L=\frac{1}{2}F_{ab}(\varphi)\partial_\mu\varphi^a\partial^\mu\varphi^b - V(\varphi) \tag{121}$$

where $F_{ab}(\varphi)$ and $V(\varphi)$ are certain functions of the scalar fields φ^a. Let us assume that $\varphi_c^a(\boldsymbol{x})$ is a static solution of the classical field equations with finite energy. It is an extremum of the energy functional

$$E[\varphi]=\int d^d x\left[\frac{1}{2}F_{ab}(\varphi)\partial_i\varphi^a\partial_i\varphi^b + V(\varphi)\right], \tag{122}$$

we shall assume that for all φ the matrix $F_{ab}(\varphi)$ defines a positive-definite quadratic form, i.e. all the eigenvalues of this matrix are positive for all φ. When

$$F_{ab}\partial_i\varphi^a\partial_i\varphi^b \geq 0, \tag{123}$$

where equality holds only for the fields which are not depend on $\boldsymbol{x}$ (classical vacuum). In addition, we shall suppose that $V(\varphi)$ is bounded from below and choose the zero- point energy level such that the value of $V(\varphi)$ at the its absolute minimum (classical vacuum) is equal to zero

$$V\left(\varphi^{(\upsilon)}\right)=0.$$

Then

$$V(\varphi)\geq 0 \tag{124}$$

and equality holds only for the classical vacuum. Then any other configuration of fields will have positive energy.

If $\varphi_c^a(\boldsymbol{x})$ is a static solution of the field equations with a finite energy, then the energy functional must be extremal for $\varphi^a=\varphi_c^a$ with respect to any variations of the field which vanish at spatial infinity. Let us consider a field configuration of the form

$$\varphi_\lambda(\boldsymbol{x})=\varphi_c(\lambda\boldsymbol{x}) \tag{125}$$

For small λ, the difference

$$\varphi_\lambda(\boldsymbol{x})-\varphi_c(\boldsymbol{x})\equiv\varphi_c(\lambda\boldsymbol{x})-\varphi_c(\boldsymbol{x})$$

is a small variation of the field. It vanishes at spatial infinity, since $\varphi_c(\boldsymbol{x})$ tends to a constant as $|\boldsymbol{x}|\to\infty$ (otherwise the gradient contribution to the energy would diverge). Consequently, the energy functional calculated on the configuration (125)

$$E(\lambda)=E\left[\varphi_\lambda(\boldsymbol{x})\right]$$

must have a extremum at $\lambda=1$

$$\left.\frac{dE}{d\lambda}\right|_{\lambda=1}=0 \tag{126}$$

We shall see that in a number of cases this cannot hold.

Let us calculate the energy for this configuration, (125)

$$E(\lambda)=\int d^dx\left[\frac{1}{2}F_{ab}\left(\varphi_c(\lambda\boldsymbol{x})\right)\left(\frac{\partial}{\partial x^i}\varphi_c^a(\lambda\boldsymbol{x})\right)\left(\frac{\partial}{\partial x^i}\varphi_c^b(\lambda\boldsymbol{x})\right)+V\left(\varphi_c(\lambda\boldsymbol{x})\right)\right]$$

We make the change of variables in the integral

$$\boldsymbol{y}=\lambda\boldsymbol{x}$$

So that $d^dx=\lambda^{-d}d^dy,\qquad \frac{\partial}{\partial x^i}=\lambda\frac{\partial}{\partial y^i}$.

We obtain

$$E(\lambda)=\lambda^{-d}\int d^dy\left[\frac{1}{2}F_{ab}\left(\varphi_c(\boldsymbol{y})\right)\lambda^2\left(\frac{\partial}{\partial y^i}\varphi_c^a(\boldsymbol{y})\right)\left(\frac{\partial}{\partial y^i}\varphi_c^b(\boldsymbol{y})\right)+V\left(\varphi_c(\boldsymbol{y})\right)\right]$$

or

$$E(\lambda)=\lambda^{2-d}E_2+\lambda^{-d}E_0, \tag{127}$$

where

$$E_2 = \int d^d x \frac{1}{2} F_{ab}(\varphi_c) \partial_i \varphi_c^a \partial_i \varphi_c^b \quad (128)$$

and

$$E_0 = \int d^d x V(\varphi_c) \quad (129)$$

The last two factors are expressed solely in terms of the original solution $\varphi_c^a(\boldsymbol{x})$. They are the gradient and potential terms, respectively, in the energy of this configuration. by virtue of conditions (123) and (124), we have

$$E_2 > 0, \qquad E_0 > 0$$

and they do not depend on λ. The extremity condition of energy (126) gives

$$(2-d)E_2 - dE_0 = 0 \quad (130)$$

together with the positivity, this condition leads to serious constraints on the existence of classical solutions in scalar theories, as follows

1. d>2. From (130) we have

$$E_2 = E_0 = 0$$

This means that $\partial_i \varphi_c^a = 0$ and φ_c^a is the absolute minimum of the potential $V(\varphi)$, i.e. the only solution is the classical vacuum.

2. d=2 , condition (130) gives

$$E_0 = 0.$$

If the potential is non-trivial, then this condition also means that the only static solution is the classical vacuum. The only class of $(2+1)$-dimensional scalar models where the existence of non-trivial classical solutions is possible is that of models with

$$V(\varphi) = 0 \quad for \quad all \quad \varphi \ .$$

i.e. there is no potential term in the Lagrangian (in this case the kinetic energy term must have a complicated structure).

3. for d=1, condition (130) gives the virial theorem

$$E_2 = E_0$$

and does not impose constraints on the choice of model.

The physical reason of the absence of static solitons in $(d+1)$-dimensional scalar theories with $d > 1 \quad (and \ \ d = 2 \quad for \quad V(\varphi) \neq 0)$ is the following: If $\varphi_c^a(\boldsymbol{x})$ is some configuration of scalar fields, then, the energy of an adjacent configuration $\varphi^a(\lambda \boldsymbol{x})$ is less than the energy of the original field, at $\lambda > 1$ .The configuration $\varphi^a(\lambda \boldsymbol{x})$ differs in size from $\varphi^a(\boldsymbol{x})$ by a factor λ^{-1}. In other words it is energetically favorable that a particle-like configuration becomes unboundedly shrunken. Thus finite energy topological solitons in purely scalar theories with an energy of type (127) are possible only in one dimension, but not in higher dimensions. Appropriate examples were kink and sine-Gordon solutions.

Note that the vacuum solution evades Derick's theorem in all dimensions, because, by definition, the vacuum is a field that is constant in space and where the potential takes its minimal value, so $E_2 = E_0 = 0$.

There is a possibility to evade the theorem in two dimensions, if the potential term is absent $E_0 = 0$. In this case $E(\lambda) = E_2$ is independent of λ. We can do it at the cost of adding terms with higher derivatives to the Lagrangian. For example, if we add a terms with fourth order in derivatives (and, hence to the static energy). In this case the above scale arguments give the relation

$$(4-d)E_4 - (2-d)E_2 - dE_0 = 0 \tag{131}$$

where E_4 is the contribution to the energy of the field $\varphi_c(\boldsymbol{x})$ from terms with four derivatives (of the type $\int d^d x (\partial_i \varphi)^4$). For d=3 condition (131) can be satisfied for positive E_0, E_2, E_4, i.e. a soliton may exist. Such a situation is realized in the Skyrme model.

Application of the Derick theorem is not constrained only by scalar field theories. Much more rich results follows after inclusion of the gauge fields. Consider some well-known cases.

Let begin by the gauge field A_μ interacting with the scalar field multiplet φ, transforming according to a unitary representation T (generally speaking, reducible) of the gauge group G. The Lagrangian in this theory has the form

$$L = \frac{1}{2g^2} Tr\left(F_{\mu\nu}\right)^2 + \left(D_\mu \varphi\right)^+ \left(D^\mu \varphi\right) - V(\varphi), \tag{132}$$

where

$$\begin{aligned} F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu + \left[A_\mu, A_\nu\right] \\ D_\mu \varphi &= \left[\partial_\mu + T\left(A_\mu\right)\right]\varphi \end{aligned} \tag{133}$$

We shall use the matrix form of the gauge fields.

These quantities in the gauge $A_0 = 0$ have the form

$$\begin{aligned} F_{0i} &= 0 \\ D_0 \varphi &= 0 \end{aligned} \tag{134}$$

The energy functional for the considered fields has the form

$$E(A_i, \varphi) = \int d^d x \left[-\frac{1}{2g^2} TrF_{ij}F_{ij} + \left(D_i\varphi\right)^+ \left(D_i\varphi\right) + V(\varphi)\right]. \tag{135}$$

Here all three terms are positive (as before, we assume that $V(\varphi)$ is non-negative and equal to zero only for the classical vacuum).Suppose $\boldsymbol{A}_c(\boldsymbol{x})$ and $\varphi_c(\boldsymbol{x})$ are classical solutions. We apply again a scale transformation and we chose it so that F_{ij} and $D_i\varphi$ transform homogeneously. It leads us to the following family of fields

$$\begin{aligned}\varphi_\lambda(\boldsymbol{x}) &= \varphi_c(\lambda \boldsymbol{x})\\ A_\lambda(\boldsymbol{x}) &= \lambda A_c(\lambda \boldsymbol{x})\end{aligned} \quad . \qquad (136)$$

Then the covariant derivative with respect to $\boldsymbol{x}$ for the new configuration is equal to

$$\boldsymbol{D}_{\boldsymbol{x}}^{(\lambda)}\varphi_\lambda(\boldsymbol{x}) = \left[\frac{\partial}{\partial \boldsymbol{x}} + T\left(A_\lambda(\boldsymbol{x})\right)\right]\varphi_\lambda(\boldsymbol{x}) = \lambda \boldsymbol{D}_{\boldsymbol{y}}\varphi_c(\boldsymbol{y})$$

where $\boldsymbol{y} = \lambda x$, and

$$\boldsymbol{D}_{\boldsymbol{y}}\varphi_c(\boldsymbol{y}) = \left[\frac{\partial}{\partial \boldsymbol{y}} + T\left(\boldsymbol{A}_c(\boldsymbol{y})\right)\right]\varphi_c(\boldsymbol{y})$$

is the covariant derivative with respect to $\boldsymbol{y}$ for the original configuration.

The strength tensor for a new field is equal to

$$F_{ij}^{(\lambda)}(\boldsymbol{x}) = \frac{\partial}{\partial x^i}A_\lambda^j(\boldsymbol{x}) - \frac{\partial}{\partial x^j}A_\lambda^i(\boldsymbol{x}) + \left[A_\lambda^i(\boldsymbol{x}), A_\lambda^j(\boldsymbol{x})\right] = \lambda^2 F_{ij}^{(c)}(\boldsymbol{y})$$

where

$$F_{ij}^{(c)} = \frac{\partial}{\partial y^i}A_c^i(\boldsymbol{y}) - \frac{\partial}{\partial y^j}A_c^i(\boldsymbol{y}) + \left[A_c^i(\boldsymbol{y}), A_c^j(\boldsymbol{y})\right]$$

is the strength tensor of the original configuration with coordinates $\boldsymbol{y}$. Then the energy functional for the configuration (136) is

$$E(\lambda) = \lambda^{4-d}\tilde{E}_4 + \lambda^{2-d}\tilde{E}_2 + \lambda^{-d}E_0$$

where

$$\tilde{E}_4 = \int d^d y\left(-\frac{1}{2g^2}TrF_{ij}^{(c)}(\boldsymbol{y})F_{ij}^{(c)}(\boldsymbol{y})\right)$$

$$\tilde{E}_2 = \int d^d y\left(D_y\varphi_c\right)^+\left(D_y\varphi_c\right)$$

are contributions of gauge field and the covariant derivative, respectively. The extremity condition on $E(\lambda) \quad at \quad \lambda = 1$ gives

$$(4-d)\tilde{E}_4 + (2-d)\tilde{E}_2 - dE_0 = 0. \qquad (137)$$

This condition is far weaker than (130) – it does not prohibit the existence of non-trivial classical solutions for $d = 2 \;\; and \;\; d = 3$. The case d=4 is also interesting, as it is required that scalar fields be completely absent from the theory or the value of scalar fields would be the vacuum value everywhere in space. Condition (137) prohibits the existence of non-trivial static classical solutions - in theories with scalar fields for $d \geq 4$ and in purely gauge theories for $d \neq 4$ - in particular, there are no solitons in physically interesting $(3+1)$ - dimensional space-time.

Lecture 12

(a) ***Application of the Derick theorem: (2+1)-dimensional scalar model***

According to the Derick theorem the only class of $(2+1)$-dimensional scalar models where the existence of non-trivial classical solutions is possible, is that of models with

$$V(\varphi)=0 \quad for \quad all \quad \varphi \;.$$

i.e. there is no potential term in the Lagrangian (in this case the kinetic energy term must have a complicated structure). However this makes the model too simple. The equation obeyed by static solution, as derived from the Lagrangian (121), would be

$$\nabla^2\varphi=0$$

whose only non-singular solutions are constants. A non-trivial situation appears if we consider several fields and constrain them by some non-linear condition. For example, the $O(N)$ model consists of N real scalar fields $\boldsymbol{\varphi}(\boldsymbol{x},t)=\left[\varphi^a(\boldsymbol{x},t), a=1,2,3\right]$, where $\boldsymbol{x}=(x,y)$ is a two-dimensional vector in ordinary space. These fields obey the constraint equation at all $(\boldsymbol{x},t)$

$$\varphi^a(x)\varphi^a(x)=\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}=\boldsymbol{1} \tag{138}$$

Thus these fields belong to the sphere S^2 of unit radius in the internal space.

We chose the Lagrangian of the model in the form

$$L=\frac{1}{2}\partial_\mu\varphi^a\partial^\mu\varphi^a=\frac{1}{2}\left(\partial_\mu\boldsymbol{\varphi}\right)\cdot\left(\partial^\mu\boldsymbol{\varphi}\right) \tag{139}$$

Note that both Lagrangian (139) and constraint (138) are invariant under global $O(3)$ rotations in internal space.

Although the Lagrangian (139) is quadratic in the fields, the field equations are nonlinear, since the non-linear constraint (138) is imposed upon the fields. To obtain these equations we use the standard Lagrange multiplier method, i.e. we write

$$S[\boldsymbol{\varphi}]=\frac{1}{2}\int d^2x dt\left[\left(\partial_\mu\boldsymbol{\varphi}\right)\cdot\left(\partial^\mu\boldsymbol{\varphi}\right)+\lambda(\boldsymbol{x},t)(\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}-1)\right] \tag{140}$$

Resulting field equation is

$$\partial_\mu\partial^\mu\boldsymbol{\varphi}+\lambda\boldsymbol{\varphi}=(\Box+\lambda)\boldsymbol{\varphi}=0\,. \qquad (141$$

Multiplying this equation from the left by $\boldsymbol{\varphi}$, we obtain the Lagrange multiplier

$$\lambda=\lambda\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}=-\boldsymbol{\varphi}\cdot\Box\boldsymbol{\varphi} \tag{142}$$

Upon inserting this relation into the Eq. (141), we derive the final equation for static field

$$\nabla^2\boldsymbol{\varphi}-\left(\boldsymbol{\varphi}\nabla^2\boldsymbol{\varphi}\right)\boldsymbol{\varphi}=0 \tag{143}$$

We see that the equation is highly non-linear, therefore the non-trivial static solutions are expected. Let us study this problem in more detail. The energy functional for these configurations has the form

$$E=\frac{1}{2}\int\left(\partial_i\varphi^a\right)\left(\partial_i\varphi^a\right)d^2x \tag{144}$$

Since the static energy is quadratic in spatial derivatives, and since space is two-dimensional, a spatial rescaling does not change the energy. The model in fact is conformally invariant. This does not rule out static solutions, but it means that each solution lies in a 1-parameter family of solutions related by rescallings.

Consider first the ground state – the field configuration with least energy. It is clear that the least value of the energy is zero, which is realized for constant (non-dependent on $\boldsymbol{x}$) in space field, because

$\partial_i \boldsymbol{\varphi} = \boldsymbol{0}$ and $\boldsymbol{\varphi} = \boldsymbol{\varphi}^{(0)}$, which is any unit vector in internal space. While $\boldsymbol{\varphi}^{(0)}$ must be x -independent in an $E = 0$ solution, it could be point in any direction in internal space, as long as it is a unit vector, Thus we have a degenerate continuous family of $E = 0$ solutions, corresponding to the different directions in which $\boldsymbol{\varphi}^{(0)}$ could point. As usual, owing to global $O(3)$ symmetry, we can choose as ground state any constant vector, and the global $O(3)$ symmetry is broken. The usual choice is

$$\varphi^a = -\delta^{a3}$$

which corresponds to the south pole of the sphere S^2.

Next, we proceed to solutions with *non-zero but finite energy*. It is clear from (144) that they must satisfy the condition (in polar coordinates) in x-space

$$r\left\| grad\boldsymbol{\varphi} \right\| \to 0, \quad as \quad r \to \infty \tag{145}$$

or

$$\lim_{r\to\infty} \boldsymbol{\varphi}(x) = \boldsymbol{\varphi}^{(0)} \tag{146}$$

where $\boldsymbol{\varphi}^{(0)}$ is again some unit vector in internal space. Note that as we tend to infinity in coordinate space in different directions, $\boldsymbol{\varphi}(x)$ must approach the same limit $\boldsymbol{\varphi}^{(0)}$. Otherwise $\boldsymbol{\varphi}(x)$ will depend on the coordinate angle θ even at $r = \infty$ and the polar component of the gradient $\frac{1}{r}\frac{\partial\varphi}{\partial\theta}$ will not satisfy to (145). We conclude that $\boldsymbol{\varphi}(x)$ approaches the same value $\boldsymbol{\varphi}^{(0)}$ at all points at infinity. As far as all spatially equivalent infinite "points" can be identified, and the space becomes topologically equivalent to the two dimensional sphere, the physical coordinate plane R_2 is essentially compacted into a spherical surface $S_2^{(phys)}$. That is, the plane R_2 may be folded into a spherical surface, with the circle at infinity reduced to the north pole of the sphere. Meanwhile, the internal space is also a spherical surface of unit radius, because of (138) .Then any finite-energy state configuration $\boldsymbol{\varphi}(x)$ is just a mapping of $S_2^{(phys)}$ into $S_2^{(\text{int})}$. This mapping is characterized by a topological number $n = 0, \pm1, \pm2, \ldots$, called the degree of the mapping. The set of configurations of fields φ^a divides into disjoint subsets (sectors): The sector with $n = 0$ contains the vacuum, while in the sector with $n = 1$ the topological soliton can be sought.

It is useful to derive an explicit formula for this mapping as a function of $\varphi^a(\boldsymbol{x})$. For this, consider a mapping of a region near the point $\boldsymbol{x}$ to the region near the point $\varphi^a(\boldsymbol{x})$ (see, Fig.24).

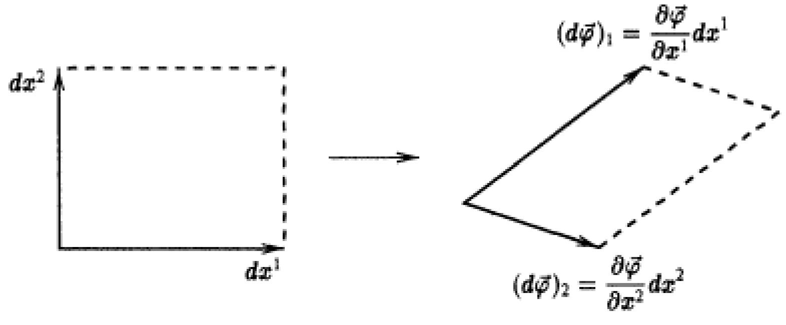


Fig.24 **The area of the region obtained by this mapping is equal to** $d\boldsymbol{\sigma} = (d\boldsymbol{\varphi})_1 \times (d\boldsymbol{\varphi})_2$

The vector $d\boldsymbol{\sigma}$ may be either parallel or antiparallel to the vector $\boldsymbol{\varphi}$ (because $d\boldsymbol{\sigma}$ is a region on the sphere $S_2^{(\text{int})}$ and $\boldsymbol{\varphi}$ is orthogonal to that sphere). If the mapping has degree ("winding number") n, then the sphere $S_2^{(\text{int})}$ is covered n times, i.e.

$$Q = \frac{1}{4\pi}\left(surface\ \ area,\ swept\ \ by\ \ \ mapping\right) = n$$

Hence the area of an element of the surface should be taken with the plus sign if the orientation of $(d\boldsymbol{\varphi})_1$ and $(d\boldsymbol{\varphi})_2$ is the same as that of $(d\mathbf{x})_1$ and $(d\mathbf{x})_2$, and the minus sign otherwise. The sign is derived correctly if we write

$$Q = \frac{1}{4\pi}\int \boldsymbol{\varphi}\cdot d\boldsymbol{\sigma}$$

which gives

$$Q = \frac{1}{4\pi}\int d^2x\boldsymbol{\varphi}\cdot\left[\frac{\partial\boldsymbol{\varphi}}{\partial x^1}\times\frac{\partial\boldsymbol{\varphi}}{\partial x^2}\right] = \frac{1}{8\pi}\int d^2x\varepsilon^{abc}\varepsilon_{ij}\varphi^a\partial_i\varphi^b\partial_j\varphi^c \qquad (147)$$

The topological number does not change under smooth variations of the fields φ^a, which do not affect spatial infinity. It is associated not with the properties of the field at spatial infinity, but with the fields in the whole space.

One can reconsider this result from another, but equivalent, point of view:
Since the Jacobian of the change of variables from $\{x_1, x_2\}$ to $\{\xi_1, \xi_2\}$, where $\{\xi_1, \xi_2\}$ are the polar angles in the internal sphere $S_2^{(\text{int})}$, which we introduce instead of three Cartesian variables φ^a, subject to $\sum_{a=1}^{3}\varphi_a^2 = 1$, can be extracted with the help of relation between surface area elements

$$dS_a^{(\text{int})} = d^2\xi\left(\frac{1}{2}\varepsilon_{ij}\varepsilon_{abc}\frac{\partial\varphi^b}{\partial\xi_i}\frac{\partial\varphi^c}{\partial\xi_j}\right) \qquad (148)$$

Now, according to (147)

$$Q = \frac{1}{8\pi}\int \varepsilon_{ij}\varepsilon_{abc}\varphi^a\frac{\partial\varphi^b}{\partial x^i}\frac{\partial\varphi^c}{\partial x^j}d^2x$$

$$=\frac{1}{8\pi}\int \varepsilon_{ij}\varepsilon_{abc}\varphi^a \frac{\partial\varphi^b}{\partial\xi_r}\frac{\partial\xi_r}{\partial x^i}\frac{\partial\varphi^c}{\partial\xi_s}\frac{\partial\xi_s}{\partial x^j}d^2x=\frac{1}{8\pi}\int \varepsilon_{rs}\varepsilon_{abc}\varphi^a\frac{\partial\varphi^b}{\partial\xi_r}\frac{\partial\varphi^c}{\partial\xi_s}d^2\xi \qquad (149)$$

Using the Jacobian

$$\varepsilon_{rs}d^2\xi=\varepsilon_{ij}\frac{\partial\xi_r}{\partial x^i}\frac{\partial\xi_s}{\partial x^j}d^2x$$

in (149) together with (148), we obtain

$$Q=\frac{1}{4\pi}\int dS_a^{(\mathrm{int})}\cdot\varphi^a=\frac{1}{4\pi}\int dS^{(\mathrm{int})}=n \qquad (150)$$

It clearly follows that n gives the number of times the internal sphere is traversed as we span the coordinate space R_2 as compacted into $S_2^{(phys)}$.

After this topological consideration let us return to solution in the sector with topological number equal to n. In order to find an explicit solution, we use a very useful technique, which has an analogy in certain more complicated models.

Let us consider the quantity

$$F_i^a=\partial_i\varphi^a\pm\varepsilon^{abc}\varepsilon_{ij}\varphi^b\partial_j\varphi^c \qquad (151)$$

In the explicit vector form it looks like

$$\boldsymbol{F}=\partial_i\boldsymbol{\varphi}\pm\varepsilon_{ij}\boldsymbol{\varphi}\times\partial_j\boldsymbol{\varphi} \qquad (151a)$$

It is clear, that

$$\int \boldsymbol{F}_i\cdot\boldsymbol{F}_i d^2x\geq 0 \qquad (152)$$

where equality holds only if

$$\partial_i\varphi^a\pm\varepsilon^{abc}\varepsilon_{ij}\varphi^b\partial_j\varphi^c=0 \qquad (153)$$

Fields, satisfying such equations are called a self-dual.

Upon expanding (152) we derive

$$\begin{aligned}&\int d^2x\Big[(\partial_i\boldsymbol{\varphi})\cdot(\partial_i\boldsymbol{\varphi})+\varepsilon_{ij}\left(\boldsymbol{\varphi}\times\partial_j\boldsymbol{\varphi}\right)\varepsilon_{ik}\left(\boldsymbol{\varphi}\times\partial_k\boldsymbol{\varphi}\right)\Big]\geq\\ &\pm 2\int d^2x\Big[\varepsilon_{ij}\boldsymbol{\varphi}\cdot\left(\partial_i\boldsymbol{\varphi}\times\partial_j\boldsymbol{\varphi}\right)\Big]\end{aligned} \qquad (154)$$

The two terms on the left side are actually equal to each other since because of (138),

$$\boldsymbol{\varphi}\cdot\partial_j\boldsymbol{\varphi}=0 \qquad (155)$$

and it follows that

$$\begin{aligned}&\varepsilon_{ij}\varepsilon_{ik}\left(\boldsymbol{\varphi}\times\partial_i\boldsymbol{\varphi}\right)\left(\boldsymbol{\varphi}\times\partial_k\boldsymbol{\varphi}\right)=\\ &\quad=\delta_{jk}\Big[\left(\boldsymbol{\varphi}\cdot\boldsymbol{\varphi}\right)\left(\partial_j\boldsymbol{\varphi}\cdot\partial_k\boldsymbol{\varphi}\right)-\left(\boldsymbol{\varphi}\cdot\partial_j\boldsymbol{\varphi}\right)\left(\boldsymbol{\varphi}\cdot\partial_k\boldsymbol{\varphi}\right)\Big]=\left(\partial_j\boldsymbol{\varphi}\right)\left(\partial_j\boldsymbol{\varphi}\right)\end{aligned}$$

where we have used Eqs.(155) and (138) and also the well-known relation from vector algebra

$$\left[\boldsymbol{a}\times\boldsymbol{b}\right]\cdot\left[\boldsymbol{c}\times\boldsymbol{d}\right]=\det\begin{pmatrix}\boldsymbol{a}\cdot\boldsymbol{c} & \boldsymbol{b}\cdot\boldsymbol{c}\\ \boldsymbol{a}\cdot\boldsymbol{d} & \boldsymbol{b}\cdot\boldsymbol{d}\end{pmatrix}$$

therefore we obtain

$$2\int d^2x\left(\partial_i\boldsymbol{\varphi}\right)\cdot\left(\partial_i\boldsymbol{\varphi}\right)\geq\pm2\int d^2x\varepsilon_{ij}\boldsymbol{\varphi}\cdot\left(\partial_i\boldsymbol{\varphi}\times\partial_j\boldsymbol{\varphi}\right)$$

or

$$E\geq4\pi\left|n\right| \tag{156}$$

This inequality sets a lower bound for the energy of any static configuration in a given n-sector. But we know that the static field equations are derived from the extremum condition of static energy functional together with constraint. Since any configuration cannot transfer from one sector to another by continuous variation, one can find extremums in each separate sectors for the given n, for a given sector the energy becomes minimum for the equality in (152). It means that the equality (153) takes place

$$\partial_i\boldsymbol{\varphi}=\pm\varepsilon_{ij}\boldsymbol{\varphi}\times\partial_j\boldsymbol{\varphi} \tag{157}$$

Any field configuration, satisfying this condition together with constraint (138), automatically will minimize E in some n-sector and therefore satisfy the energy extremum condition in a given sector in form of field equations (143). Indeed

$$\begin{aligned}\nabla^2\boldsymbol{\varphi}&=\partial_i\partial_i\boldsymbol{\varphi}=\pm\partial_i\left(\varepsilon_{ij}\boldsymbol{\varphi}\times\partial_j\boldsymbol{\varphi}\right)=\pm\varepsilon_{ij}(\partial_i\boldsymbol{\varphi})\times\left(\partial_j\boldsymbol{\varphi}\right)=\\&=\varepsilon_{ij}\left(\varepsilon_{ik}\boldsymbol{\varphi}\times\partial_k\boldsymbol{\varphi}\right)\times\partial_j\boldsymbol{\varphi}=\delta_{jk}\left[\partial_k\boldsymbol{\varphi}\left(\boldsymbol{\varphi}\cdot\partial_j\boldsymbol{\varphi}\right)-\boldsymbol{\varphi}\left(\partial_j\boldsymbol{\varphi}\cdot\partial_k\boldsymbol{\varphi}\right)\right]=\\&=\boldsymbol{\varphi}\left(\boldsymbol{\varphi}\cdot\nabla^2\boldsymbol{\varphi}\right)\end{aligned}$$

which is just the field equation. in the last step we have used

$$\boldsymbol{\varphi}\cdot\partial_i\boldsymbol{\varphi}=0,\quad\Rightarrow\quad\partial_i\boldsymbol{\varphi}\cdot\partial_i\boldsymbol{\varphi}+\boldsymbol{\varphi}\nabla^2\boldsymbol{\varphi}=0,\quad\rightarrow\partial_i\boldsymbol{\varphi}\cdot\partial_i\boldsymbol{\varphi}=-\boldsymbol{\varphi}\nabla^2\boldsymbol{\varphi}$$

which follows directly from differentiating the constraint (155).

Therefore, we have derived the equation of motion for any field configuration, i.e. any solution of the relation (153) will satisfy to the equation of motion also. But the opposite does not happen, it is in principle possible to have a solution of equation of motion, which does not obey to (153).One could in principle have solutions of Eq. (143) which do not satisfy (157). These would not represent absolute minima of E in the corresponding n-sector, but some higher valued extrema of E, such as local minima. In practice, one tried solving the Eq. (157), because it is a first-order differential equation while (143) is a second-order equation.

(b) ***Explicit solution for soliton fields***

It is important to note that, unlike the original field equation (143), Eq. (153), is a first order and is easier to solve. Let us use an ansatz which is invariant under $SO(2)$ spatial rotations, complimented by $SO(2)$ rotations around the third axis in the space of the field

$$\varphi^\alpha(\boldsymbol{x}) = n^\alpha \sin f(r)$$
$$\varphi^3(\boldsymbol{x}) = \cos f(r) \tag{158}$$

where $n^\alpha = \frac{x^\alpha}{r}, \quad \alpha = 1,2$. The condition $\varphi^a \varphi^a = \varphi^\alpha \varphi^\alpha + (\varphi^3)^2 = 1$ is automatically satisfied.

The derivatives of the fields are equal to

$$\partial_i \varphi^\alpha = \frac{1}{r}\left(\delta^{i\alpha} - n^i n^\alpha\right)\sin f + n^i n^\alpha f' \cos f$$
$$\partial_i \varphi^3 = -n^i f' \sin f$$

Moreover

$$\varepsilon_{ij}\varepsilon^{3\alpha\beta}\varphi^\alpha \partial_j \varphi^\beta = -\varepsilon_{ij}\varepsilon_{\alpha\beta} n_\alpha \sin f\left[\frac{1}{r}\left(\delta^{i\alpha} - n^i n^\alpha\right)\sin f + n^i n^\alpha f' \cos f\right] =$$
$$= -\varepsilon_{ij}\varepsilon_{\alpha\beta}\delta^{j\beta} n_\alpha \frac{1}{r}\sin^2 f = n_i \frac{1}{r}\sin^2 f$$

Eq. (153) with $\alpha = 3$ takes the form

$$-n_i f' \sin f + n_i \frac{1}{r}\sin^2 f = 0$$

or
$$f' = \frac{1}{r}\sin f \tag{159}$$

Eq. (153) with $\alpha = 1,2$ reduces to this equation. The solution of this equation with boundary condition which ensures that $\varphi^\alpha = -\delta^{\alpha 3} \quad as \quad r \to \infty$,

$$f(\infty) = \pi$$

has the form

$$f = 2\arctan\frac{r}{r_0}$$

so that

$$\varphi^\alpha = 2\frac{x_\alpha r_0}{r_0^2 + r^2}$$
$$\varphi^3 = \frac{r_0^2 - r^2}{r_0^2 + r^2}$$

where r_0 is an arbitrary constant (soliton size). The fact that the soliton size may be arbitrary, actually follows from the scale (Derick) considerations.

We could connect the above found solution to the method of stereographic projection. Indeed, let us take in accordance by stereographic projection the points of internal sphere $S_2^{(\text{int})}$ and points of surface with Cartesian coordinates ω_1 and ω_2 on this plane by following manner

$$\omega_\alpha = 2\frac{\varphi^\alpha}{1-\varphi^3}, \quad \alpha = 1,2$$

and consider the complex variables $\omega = \omega_1 + i\omega_2$ and $\varphi = \varphi_1 + i\varphi_2$. Then

$$\partial_1\omega = \frac{\partial\omega}{\partial x_1} = \left[2(1-\varphi_3)\partial_1\varphi + \varphi\partial_1\varphi_3\right]/(1-\varphi_3)^2 =$$

$$= \frac{2}{(1-\varphi_3)^2}\left(\partial_1\varphi_1 + \varphi\overleftrightarrow{\partial}\varphi_3\right)$$

where we have used the traditional notation for antisymmetric derivative

$$a\overleftrightarrow{\partial}b = a(\partial b) - (\partial a)b$$

We have from Eq. (151), that

$$\partial_i\varphi_1 = \pm\varepsilon_{ij}\varphi_2\overleftrightarrow{\partial}_j\varphi_3 \qquad \partial_i\varphi_2 = \pm\varepsilon_{ij}\varphi_3\overleftrightarrow{\partial}_j\varphi_1$$

Therefore

$$\partial_1\varphi = \mp i\varphi\overleftrightarrow{\partial}_2\varphi_3 \qquad \partial_2\varphi = \pm i\varphi\overleftrightarrow{\partial}_1\varphi_3$$

Thus

$$\partial_1\omega = \mp i\partial_2\omega$$

which means that

$$\frac{\partial\omega_1}{\partial x_1} = \pm\frac{\partial\omega_2}{\partial x_2} \qquad \frac{\partial\omega_1}{\partial x_2} = \mp\frac{\partial\omega_2}{\partial x_1}. \tag{160}$$

Let us remember that $x_{1,2}$ are the Cartesian coordinates of our initial physical space, and $\omega_{1,2}$ belong to the plane in an “internal” space, on which the internal $S_2^{(\mathrm{int})}$ sphere has been projected stereographically. Eq. (160) is all too familiar as the Cauchy-Riemann condition for ω being analytic function of z^* (for upper sign) or z (for lower sign), where $z = x_1 + ix_2$. Thus, any analytic function $\omega(z)$ or $\omega(z^*)$ automatically solves (153) and therefore also the field equation when written in terms of original variables φ_a *and* x. Furthermore, while ω must be analytic in either z or z^*, it need not be an entire function – isolated poles in $\omega(z)$ is permitted.

Let us write down the expressions for energy and topological number in terms of $\omega = \omega(z)$. It is evident that

$$E = \int d^2x \frac{\left|\frac{d\omega}{dz}\right|^2}{\left(1+|\omega|^2/4\right)} \qquad \text{and} \qquad |Q| = \frac{E}{4\pi}. \tag{161}$$

A prototype solution for arbitrary positive n is given by

$$\omega(z) = \left[(z-z_0)/\lambda\right]^n \tag{162}$$

where n is any positive integer, and λ is any real number, while z_0 is any complex number. Above ω represents a point in field space, while z stands for a point in coordinate space. Clearly (162) allows n roots z for a given $\omega(z)$. Therefore it must corresponds to the n-sector. This may be verified by substitution (162) into (161). We have

$$Q=\frac{1}{4\pi}\int d^2x\frac{n^2\left|z-z_0\right|^{2n-2}\lambda^{2n}}{\left(\lambda^{2n}+\frac{1}{4}\left|z-z_0\right|^{2n}\right)^2}$$

Using

$$z-z_0=\rho e^{i\vartheta} \qquad and \qquad d^2x=\rho d\rho d\vartheta$$

the integration yields $Q=n$. Hence $E=4\pi Q=4\pi n$ is finite. Clearly, then, these are explicit solitary-wave solutions for any positive integer n.

The constants λ and z_0 refer to the size and location of the soliton solution. The fact that the solution exists for arbitrary λ and z_0 and the fact that neither Q nor E depend on these constants is a reflection of scale and translational invariance: $x\to\lambda x, \quad x\to x-a$ does not change the energy functional $E[\phi]$ in (144).

Lecture 13

Skyrme model and skyrmion

In what follows we consider a soliton, which in topological sense is similar to the soliton considered in previous section, but appearing in nonlinear model with chiral symmetry, so-called the sigma model in physically interesting (3+1)-dimensional space-time. This model was found by Skyrme in 1961 and the corresponding soliton was named as skyrmion. During the long time this model was forgotten, but in 80-ies a revival interest has inspired to this model, thanks to fundamental papers of Witten (1983). Using N_c^{-1} expansion (where N_c is the number of colors) it was shown that low energy limit of QCD can be described by an effective meson Lagrangian and baryons in the large N_c world evidently are the soliton solutions of this Lagrangian. The Skyrme model is interesting and is a rather realistic model of nucleon, and the characteristic qualities of this soliton - spin, isospin etc., correspond to nucleon, and many quantitive features of nucleon, such as a charge radius and so on, are reproduced by this model with a rather satisfactory accuracy.

Because the Skyrme model is based on the chiral symmetry of strong interactions, below we at first consider fundamentals of the chiral symmetry and then construct the Skyrme model Lagrangian and consider its principal properties.

It is well known that the strong interactions of ordinary hadrons (nucleons, their resonances, pions, kaons, hyperons etc.) exhibit a symmetry under the rotations in the internal (isospin) space, which is generated by 3 isotopic charges $Q_A(A=1,2,3)$ and obey to the commutation relations of $SU(2)$ algebra

$$\left[Q^A,Q^B\right]=i\varepsilon^{ABC}Q^C \tag{161a}$$

According to general strategy of quantum field theory these charges are space integrals from the zeroth components of corresponding Noether's vector current

$$Q_A(t)=\int d^3xV_A^0\left(\boldsymbol{x},t\right).$$

If we have also the axial current, $A_\mu^5\left(x\right)$ (Weak interactions of hadrons), one can define its charge with full analogy 3-dimensional integral from its zeroth component,

$$Q^5\left(t\right)=\int d^3xA_0^5\left(\boldsymbol{x},t\right)$$

It is a pseudo scalar in ordinary space and the vector in isotopic space. Therefore we have

$$\left[Q^A,Q_5^B\right]=i\varepsilon_{ABC}Q_5^C \tag{161b}$$

One can enlarge of isospin algebra if we close axial charge commutator on the vector charges (In suitable models this statement may be confirmed (for example, if currents have the Noether's form of spinor theory, i.e. $J_\mu^A\left(x\right)=i\bar{\psi}\left(x\right)\gamma_\mu\gamma_5\tau^A\psi\left(x\right)$, and we use the equal time anticommuting relations for quantized Dirac fields). In particular, one can assert that

$$\left[Q_5^A,Q_5^B\right]=i\varepsilon_{ABC}Q^C \tag{161c}$$

Exercise: Derive relations (161,a,b,c) from the Noether's currents $J_\mu^A\left(x\right)=i\bar{\psi}\left(x\right)\gamma_\mu\gamma_5\tau^A\psi\left(x\right)$, proceed from definitions.

Hint: Use the identities

$$\left[AB,CD\right]\equiv -AC\left\{D,B\right\}+A\left\{C,B\right\}D-C\left\{D,A\right\}B+\left\{C,A\right\}DB$$

and

$$\left[\Gamma_a\tau_a,\Gamma_b\tau_b\right]\equiv\frac{1}{2}\left\{\Gamma_a,\Gamma_b\right\}\left[\tau_a,\tau_b\right]+\frac{1}{2}\left[\Gamma_a,\Gamma_b\right]\left\{\tau_a,\tau_b\right\}$$

where $\Gamma_{a,b}$ are the Dirac matrices.

Relations (161a, b, and c) form a closed algebra with 6 generators. To study the structure of this algebra, let us introduce new generators

$$Q_\pm^A=\frac{1}{2}\left(Q^A\pm Q_5^A\right) \tag{162}$$

Then above commutators decay as follows

$$\begin{aligned}&\left[Q_+^A,Q_+^B\right]=i\varepsilon_{ABC}Q_+^C\\&\left[Q_-^A,Q_-^B\right]=i\varepsilon_{ABC}Q_-^C\\&\left[Q_+^A,Q_-^B\right]=0\end{aligned} \tag{163}$$

It is a famous $SU(2)\times SU(2)$ algebra of Gell-Mann (1963). Though this algebra is a direct product of two algebras, the effective Lagrangians, generated from it are nontrivial because the parity operation connects both of them

$$PQ_{+}^{A}P^{-1} = Q_{-}^{A} \tag{164}$$

Historically chiral symmetry was directed to special role of pions in strong interaction physics, particularly to the realization of Goldstone theorem and the hypothesis of PCAC (partial conservation of axial current). Therefore construction of effective Lagrangians for strong interaction also be turn to π-mesons.

Representations of $SU(2)\times SU(2)$ *algebra*

Let us assume that the generators Q_{+}^{A} transform the irreducible tensor representation with isospin t as

$$U_{+}\phi_{A}U_{+}^{-1} = \phi_{B}\left[e^{-i\boldsymbol{\alpha}\cdot\boldsymbol{T}}\right]_{BA} \tag{165a}$$

where

$$U_{+} = e^{-i\boldsymbol{\alpha}\cdot\boldsymbol{Q}_{+}} \tag{165b}$$

and $\boldsymbol{T}$ composes $(2t+1)\times(2t+1)$ matrix representation of $\boldsymbol{Q}_{+}$ charges.

As we want establish field transformation properties under the group $SU(2)\times SU(2)$, consider one more operator of transformation

$$U_{-} = e^{-i\boldsymbol{\beta}\cdot\boldsymbol{Q}_{-}} \tag{166}$$

Because we have the direct product of two algebras, we must use the additional index (doted) for the fields transforming by U_{-}

$$\begin{aligned} U_{+}\phi_{A\dot{B}}U_{+}^{-1} &= \phi_{C\dot{B}}\left[e^{-i\boldsymbol{\alpha}\cdot\boldsymbol{T}}\right]_{CA} \\ U_{-}\phi_{A\dot{B}}U_{-}^{-1} &= \phi_{A\dot{D}}\left[e^{-i\boldsymbol{\beta}\cdot\boldsymbol{T}'}\right]_{\dot{D}\dot{B}} \end{aligned} \tag{167}$$

where $\boldsymbol{T}'$ correspond to $(2t'+1)\times(2t'+1)$ matrix representation for $\mathbf{Q}_{-}$ charges.

We say that $\phi_{A\dot{B}}$ forms (t,t') representation of the group $SU(2)\times SU(2)$.

Because in the $SU(2)$ space there is a such matrix C that

$$C\boldsymbol{T}^{*}C^{-1} = -\boldsymbol{T},$$

the second transform law is equivalent to

$$U_{-}\phi_{A\dot{B}}U_{-}^{-1} = \left[e^{i\boldsymbol{\beta}\cdot\boldsymbol{T}'}\right]_{\dot{B}\dot{D}}\phi_{A\dot{D}} \tag{167a}$$

As we have mentioned above the chiral symmetry was intended for π-mesons. But the careful examination shows that in case of linear representation (167) it is impossible to construct the theory only for π-fields. Therefore we must include the other fields as well. The minimal possibility is to use the representation $\left(1/2,1/2\right)$, which has the following law of transformation

$$e^{-i\alpha\cdot\mathbf{Q}_+}M_{A\dot{B}}e^{i\alpha\cdot\mathbf{Q}_+}=\left(e^{i\alpha\cdot\frac{\tau}{2}}\right)_{AC}M_{C\dot{B}} \tag{168}$$

also

$$e^{-i\beta\cdot\mathbf{Q}_-}M_{A\dot{B}}e^{i\beta\cdot\mathbf{Q}_-}=M_{A\dot{C}}\left(e^{-i\beta\cdot\frac{\tau}{2}}\right)_{\dot{C}\dot{B}} \tag{169}$$

Now if we use the combined transformation, i.e. transform these relations once again, and take $\boldsymbol{\beta}=\boldsymbol{\alpha}$ and $\boldsymbol{\beta}=-\boldsymbol{\alpha}$, one gets

$$e^{-i\alpha\cdot\mathbf{Q}}Me^{i\alpha\cdot\mathbf{Q}}=e^{i\alpha\cdot\frac{\tau}{2}}Me^{-i\alpha\cdot\frac{\tau}{2}} \tag{170a}$$

and

$$e^{-i\alpha\cdot\mathbf{Q}_5}Me^{i\alpha\cdot\mathbf{Q}_5}=e^{i\alpha\cdot\frac{\tau}{2}}Me^{i\alpha\cdot\frac{\tau}{2}} \tag{170b}$$

It is seen that the axial transformation (by Q_5) differs from vector one (Q) by the sign in exponent. Because M is 2×2 matrix, it can be written in the following (quaternionic) form

$$M=\sigma\cdot I+i\boldsymbol{\tau}\cdot\boldsymbol{\pi} \tag{171}$$

Then it follows for infinitesimal transformations that

$$\delta\left(\sigma I+i\boldsymbol{\tau}\cdot\boldsymbol{\pi}\right)=-i\boldsymbol{\tau}\cdot\boldsymbol{\alpha}\times\boldsymbol{\pi}$$

$$\delta'\left(\sigma I+i\boldsymbol{\tau}\cdot\boldsymbol{\pi}\right)=i\sigma\boldsymbol{\alpha}\cdot\boldsymbol{\tau}-\boldsymbol{\alpha}\cdot\boldsymbol{\pi}I$$

Here $\delta\left(\delta'\right)$ denotes the infinitesimal vector (axial) transformation, respectively. We have from here the transformation rules under both transformations

$$\begin{aligned}\delta\sigma&=0 &\qquad \delta'\sigma&=-\left(\boldsymbol{\beta}\cdot\boldsymbol{\pi}\right)\\ \delta\boldsymbol{\pi}&=-\left(\boldsymbol{\alpha}\times\boldsymbol{\pi}\right) &\qquad \delta'\boldsymbol{\pi}&=\boldsymbol{\beta}\sigma\end{aligned} \tag{172}$$

Moreover the parity transformation gives $PMP^{-1}=M^+$, and

$$e^{-i\alpha\cdot\mathbf{Q}_5}M^+e^{i\alpha\cdot\mathbf{Q}_5}=e^{-i\alpha\cdot\frac{\tau}{2}}M^+e^{-i\alpha\cdot\frac{\tau}{2}}$$

Therefore

$$e^{-i\alpha\cdot\mathbf{Q}_5}MM^+e^{i\alpha\cdot\mathbf{Q}_5}=e^{i\alpha\cdot\frac{\tau}{2}}MM^+e^{-i\alpha\cdot\frac{\boldsymbol{\tau}}{2}}$$

and

$$Tr\left(MM^+\right)$$

is chiral invariant. The explicit calculation gives

$$MM^+=M^+M=\left(\sigma^2+\boldsymbol{\pi}^2\right)I \tag{173}$$

Linear $SU(2)\times SU(2)$ *sigma model*

The well-known sigma model of Gell-Mann and Levy is based on $(1/2,1/2)$ linear representation. According the previous consideration The chiral invariant (renormalizable) Lagrangian may be written as

$$L_0=\frac{1}{2}\left(\partial_\mu\phi\right)^2+\frac{1}{2}\left(\partial_\mu\sigma\right)^2-\frac{1}{2}\mu^2\left(\sigma^2+\boldsymbol{\phi}^2\right)-\frac{\lambda}{4!}\left(\sigma^2+\boldsymbol{\phi}^2\right)^2 \tag{174}$$

Evidently one could add any degree of chiral invariant combination $\left(\sigma^2+\boldsymbol{\phi}^2\right)$, but Eq. (174) is renormalizable. One can clear up the role of σ field: The potential

$$V(\sigma,\phi)=\frac{1}{2}\mu^2\left(\sigma^2+\boldsymbol{\phi}^2\right)+\frac{1}{4!}\lambda\left(\sigma^2+\boldsymbol{\phi}^2\right)^2 \tag{175}$$

may have non-trivial minimum at $\sigma=\sigma_0\neq 0$ and the possibility of spontaneous breaking of symmetry. In result there appear 3 massless Goldstone bosons. For giving a nonzero mass it is traditional way to add a term, which explicitly breaks an underline symmetry, and is linear in the sigma field

$$L_{SB}=c\sigma \tag{176}$$

For derivation of currents and their divergences usually the Gell-Mann and Levy method is used.

Gell-Mann and Levy equations

Let us propose that the internal symmetry permits the following infinitesimal transformations

$$\begin{aligned}&\phi_A(x)\to\phi_A(x)+\delta\phi_A(x)=\phi_A(x)+iC_{ABC}\alpha_B\phi_C(x)\\&\partial_\mu\phi_A(x)\to\partial_\mu\phi_A(x)+iC_{ABC}\partial_\mu\left[\alpha_B(x)\phi_C(x)\right]\end{aligned} \tag{177}$$

Then , using the equation of motion one gets

$$\begin{aligned}\delta L&=\frac{\partial L}{\partial\phi_A(x)}\delta\phi_A(x)+\frac{\partial L}{\partial\left(\partial_\mu\phi_A(x)\right)}\partial_\mu\left(\delta\phi_A(x)\right)=\\&=\partial_\mu\left(\frac{\partial L}{\partial\left(\partial_\mu\phi_A(x)\right)}\right)=\partial_\mu\left(i\alpha_B(x)\frac{\partial L}{\partial\left(\partial_\mu\phi_A(x)\right)}C_{ABC}\phi_C(x)\right)\end{aligned} \tag{178}$$

If transformation could be global (non x-dependent), then from invariance of L the existence of conserved Noether's current follows

$$J_\mu^B(x)=-i\frac{\partial L}{\partial\left(\partial_\mu\phi_A(x)\right)}C_{ABC}\phi_C(x), \tag{179}$$

consequently the integral from the zeroth component of this current

$$Q^B = \int d^3x J_0^B(x) \tag{180}$$

is a symmetry generator. After such definition of current the variation of Lagrangian can be rewritten as

$$\delta L = -\alpha_B(x)\partial^\mu J_\mu^B(x) - J_\mu^B(x)\partial^\mu \alpha_B(x) \tag{181}$$

Thus we have

$$\begin{aligned} J_\mu^B(x) &= -\frac{\partial(\delta L)}{\partial\left(\partial^\mu \alpha_B(x)\right)} \\ \partial^\mu J_\mu^B(x) &= -\frac{\partial(\delta L)}{\partial \alpha_B(x)} \end{aligned} \tag{182}$$

They are Gell-Mann and Levy(GML) equations, very useful in many applications. For example, after inclusion of the L_{SB} term, the sigma model Lagrangian is no longer chiral invariant, but this term changed under axial transformation (172). Indeed

$$\delta' L = \sigma\partial^\mu \boldsymbol{\phi}\partial_\mu \boldsymbol{\beta} - \partial_\mu \sigma \boldsymbol{\phi}\cdot\partial^\mu \boldsymbol{\beta} - c\boldsymbol{\beta}\cdot\boldsymbol{\phi} \tag{184}$$

Then, according to GML equations, we have

$$\boldsymbol{J}_{5\mu} = \boldsymbol{\phi}\partial_\mu \sigma - \sigma\partial_\mu \boldsymbol{\phi} \qquad\qquad \partial^\mu \boldsymbol{J}_{5\mu} = c\boldsymbol{\phi} \tag{185}$$

The last equation is known as PCAC relation – the divergence of axial current is proportional to the pion field $\boldsymbol{\phi}$. According to PCAC one can fix the value of c,

$$c = \mu^2 \sigma_0 \tag{186}$$

Study of symmetry breaking

One can find the minimum of potential

$$V(\sigma,\phi) = \frac{1}{2}\mu^2\left(\sigma^2 + \phi^2\right) + \frac{\lambda}{4!}\left(\sigma^2 + \phi^2\right)^2 - c\sigma \tag{187}$$

The extremum with respect of σ (with respect of $\boldsymbol{\phi}$ minimum is trivial, $\boldsymbol{\phi}_0 = 0$, otherwise we destroy many non-trivial laws) gives the equation

$$\frac{\partial V(\sigma,\phi)}{\partial \sigma} = 0$$

Explicitly, this means

$$\mu^2\sigma_0 - \frac{\lambda}{6}\sigma_0^3 - c = 0 \tag{188}$$

We see, that there is a solution $\sigma_0 \neq 0$, but this solution survives or not in the limit of restored symmetry $c \to 0$, when we find two solutions

$$\sigma_0 = 0 \qquad or \qquad \sigma_0 = \sqrt{\frac{-6\mu^2}{\lambda}} \tag{189}$$

I.e. we have two solutions in the limit of vanishing symmetry breaking term. The case $\sigma_0 = 0$ correspond to the *exact (or Weyl) symmetry* - Lagrangian is symmetric and the ground state is also symmetric. In this case Lagrangian describes degenerate σ *and* $\boldsymbol{\phi}$ particles. If $\boldsymbol{\phi}$ is a pion, then it must have the scalar partner with same mass $\left(\mu_\pi = \mu_\sigma = \mu\right)$.

The second case is more interesting. It is clear that we must have $\lambda > 0$, otherwise potential energy will not have a minimum. Then it follows from the second solution that non-trivial ground state appears in case, when $\mu^2 < 0$ (i.e. if we have tachyon in this model). If so then the chiral symmetry will restore, but the ground state remains asymmetric – we have spontaneous breaking. In this case there appear massless pions. Indeed, the mass matrices as the second derivatives of potential energy in ground state have the forms

$$\begin{aligned} m_\phi^2 &= \left(\frac{\partial^2 V}{\partial \phi^2}\right)_0 = \mu^2 + \frac{\lambda}{6}\sigma_0^2 = 0 \\ m_\sigma^2 &= \left(\frac{\partial^2 V}{\partial \sigma^2}\right)_0 = \mu^2 + \frac{\lambda}{2}\sigma_0^2 = -2\mu^2 > 0 \end{aligned} \tag{190}$$

Therefore, when the chiral symmetry is broken spontaneously a pion is massless (the Goldstone theorem). If the symmetry is broken explicitly, the pion acquires the mass, proportional to $c \neq 0$

$$m_\phi^2 = \frac{c}{\sigma_0} \tag{191}$$

The linear sigma model had many applications as effective low energy theory of pions. There was various generalizations to other flavor symmetry, for example, $SU(3)$ and others. The weak point of linear sigma model is the absence of scalar particle in experiments, σ's prototype. While in studying mechanisms of chiral symmetry breaking the linear model played good laboratory, particularly pion-pion scattering lengths were calculated with rather good accuracy.

Lecture 14

Non-linear realization of pion fields

A non-linear realization contains only pion field, $\boldsymbol{\phi}$ which under the vector transformation transforms in usual manner, as linear transformation (172).

As regards of axial transformations, the general form is

$$\delta' \boldsymbol{\phi} = \boldsymbol{\beta}\, f_1\left(\phi^2\right) + \boldsymbol{\phi}\left(\boldsymbol{\beta}\cdot\boldsymbol{\phi}\right) f_2\left(\phi^2\right) \tag{192}$$

or in infinitesimal form

$$\left[Q_5^k,\phi_l\right]=i\delta_{kl}\,f_1\left(\phi^2\right)+i\phi_k\phi_l\,f_2\left(\phi^2\right) \tag{192a}$$

We must require satisfaction of the Jacobi identity

$$\left[Q_5^A,\left[Q_5^B,\phi_C\right]\right]-\left[Q_5^B,\left[Q_5^A,\phi_C\right]\right]=\left[\left[Q_5^A,Q_5^B\right],\phi_C\right]$$

After explicit calculation it is possible to find the relation

$$1+2\,f_1(x)\,f_1'(x)+2xf_2(x)\,f_1'(x)-f_1(x)\,f_2(x)=0\,, \qquad x\equiv\phi^2$$

From here we find the ratio

$$\frac{f_2(x)}{f_1(x)}=\frac{1+\dfrac{df_1^2(x)}{dx}}{f_1^2(x)-x\dfrac{df_1^2(x)}{dx}} \tag{193}$$

This relation, which is derived from the Jacobi identity, says that if we will know one of function f_1 *or* f_2, then the second one may be determined from the algebra. Therefore, suppose that $f_2(x)=0$. Then we have

$$1+\frac{df_1^2(x)}{dx}=0, \qquad \to \quad \frac{d}{dx}\left(x+f_1^2(x)\right)=0$$

So $\qquad x+f_1^2(x)=C, \qquad C=f_1^2(0)$

Therefore, the solution in this case is

$$f_1(x)=\left[f_1^2(0)-x\right]^{1/2} \tag{194}$$

It is possible to show that this solution is most general, i.e. one can always introduce a new field so that its transformation law contains only one function.

Indeed, if we introduce a new field $\boldsymbol{\pi}$, connected to the old one by relation

$$\boldsymbol{\phi}=\boldsymbol{\pi}g\left(\pi^2\right) \tag{195}$$

and at the same time assume, that the Jacobi identity is satisfied by a simple law

$$\delta'\boldsymbol{\pi}=\boldsymbol{\beta}\,f_3\left(\pi^2\right)\,, \tag{196}$$

then substituting it into (192), one finds

$$\delta'\boldsymbol{\phi}=\boldsymbol{\beta}\,f_1(x)+\boldsymbol{\pi}\left(\boldsymbol{\beta}\cdot\boldsymbol{\pi}\right)g^2(y)\,f_2(x), \qquad y\equiv\pi^2 \tag{197}$$

Now combine (195) and (196)

$$\begin{aligned}\delta'\boldsymbol{\phi}&=\delta'\boldsymbol{\pi}g\left(\pi^2\right)+2\boldsymbol{\pi}g'(y)\left(\boldsymbol{\pi}\cdot\delta'\boldsymbol{\pi}\right)\\&=\boldsymbol{\beta}\,f_3(y)g(y)+2\boldsymbol{\pi}\left(\boldsymbol{\beta}\cdot\boldsymbol{\pi}\right)f_3(y)g'(y)\end{aligned} \quad \text{= taking into account (196) =}$$

Comparing last two relations, we get

$$f_3(y)g(y)=f_1(x), \qquad 2\,f_3(y)g'(y)=g^2(y)\,f_2(x) \tag{198}$$

Then

$$\frac{f_2(x)}{f_1(x)}=2\frac{dg}{dy}\frac{1}{g^3(y)}$$

Remembering (195), from which we have $x = yg^2(y)$, after a simple calculations one can derive

$$\frac{dg}{dy} = \frac{g^2(y)}{1 - 2yg(y)\frac{dg}{dx}} \frac{dg}{dx}$$

using it in previous relation, we obtain

$$\frac{2}{g(y)} \frac{dg}{dx} = \frac{1 + \frac{d}{dx} f_1^2(x)}{x + f_1^2(x)} \tag{199}$$

Separating variables, one finds

$$g(y) = C\left[x + f_1^2(x)\right]^{1/2} \tag{200}$$

Using (198),

$$f_3(y) = \frac{f_1(x)}{C\left[f_1^2(x) + x\right]^{1/2}} \tag{201}$$

Now one can choose a condition $g(0) = 1$, then from (200) we find

$$C = \frac{1}{f_1(0)}$$

and

$$\boldsymbol{\pi}^2 = f_1^2(0) \frac{\boldsymbol{\phi}^2}{\boldsymbol{\phi}^2 + f_1^2(\boldsymbol{\phi}^2)} \tag{202}$$

Non-linear sigma model

Without restriction of generality one can suppose that the pion field transforms by only one arbitrary function,

$$\delta' \boldsymbol{\phi} = \boldsymbol{\beta} f_1(\boldsymbol{\phi}^2) \tag{203}$$

As we know from previous section, the justification of the Jacobi identities permits to find explicit expression of this function (194)

$$f_1(\boldsymbol{\phi}^2) = \left[f_1^2(0) - \boldsymbol{\phi}^2\right]^{1/2}$$

Let compare this law to the one for linear sigma model

$$\delta' \boldsymbol{\phi}(x) = \boldsymbol{\beta} \sigma(x)$$

We see that in order to construct the non-linear theory it is sufficient consider the σ field as

$$\sigma(x) \equiv f_1(\boldsymbol{\phi}^2)$$

or, remembering Eq. (194), one sees that we must take these two fields connected each other by relation

$$\sigma^2 + \boldsymbol{\phi}^2 = f_1^2(0) \tag{204}$$

Therefore the general form of the non-linear sigma model Lagrangian, which do not contain field derivatives not more than two degrees is

$$L=\frac{1}{2}\left(\partial_\mu \boldsymbol{\phi}\right)^2+\frac{1}{2}\left(\partial_\mu \sigma\right)^2 \tag{205}$$

where

$$\sigma(x)=\sqrt{f_1^2(0)+\boldsymbol{\phi}^2} \tag{204a}$$

If we add a term which breaks a chiral symmetry (176), one can write

$$L_{NLN}=\frac{1}{2}\left(\partial_\mu \boldsymbol{\phi}\right)^2+\frac{1}{2}\frac{\left(\boldsymbol{\phi}\cdot\partial_\mu \boldsymbol{\phi}\right)^2}{f_1^2(0)-\boldsymbol{\phi}^2}+m_\pi^2 f_\pi^2\left(f_1^2(0)-\boldsymbol{\phi}^2\right)^{1/2} \tag{205}$$

It contains only one upward parameter $f_1^2(0)$. One can fix it requiring that Lagrangian correctly reproduces the mass term of pion. Let us consider expansion till to fourth order

$$L\approx\frac{1}{2}\left(\partial_\mu \boldsymbol{\phi}\right)^2+\frac{\left(\boldsymbol{\phi}\cdot\partial_\mu \boldsymbol{\phi}\right)^2}{2f_1^2(0)}+m_\pi^2 f_\pi f_1(0)\left[1-\frac{\boldsymbol{\phi}^2}{2f_1^2(0)}-\frac{1}{8}\left(\frac{\boldsymbol{\phi}^2}{f_1^2(0)}+\ldots\right)\right]=$$

$$=\frac{1}{2}\left(\partial_\mu \boldsymbol{\phi}\right)^2+\frac{\left(\boldsymbol{\phi}\cdot\partial_\mu \boldsymbol{\phi}\right)^2}{2f_1^2(0)}-\frac{m_\pi^2 f_\pi}{2f_1(0)}\boldsymbol{\phi}^2-\frac{m_\pi^2}{8f_1^3(0)}\left(\boldsymbol{\phi}^2\right)^2+\ldots$$

It follows that we must take

$$f_1(0)=f_\pi \tag{206}$$

Finally we obtain the Lagrangian of non-linear sigma model

$$L=\frac{1}{2}\left(\partial_\mu \boldsymbol{\phi}\right)^2+\frac{1}{2}\frac{\left(\boldsymbol{\phi}\cdot\partial_\mu \boldsymbol{\phi}\right)^2}{f_\pi^2-\boldsymbol{\phi}^2}+m_\pi^2 f_\pi\left[f_\pi^2-\boldsymbol{\phi}^2\right]^{1/2} \tag{207}$$

This Lagrangian consists only known parameters, m_π and f_π. It had a wide application in low energy pion physics. This Lagrangian exhibits all relevant ideas about the breaking of global symmetries.

Physical meaning of symmetry breaking in considered models

Let we have the local conserved current

$$\partial^\mu j_\mu(x)=0$$

The strong definition of charge is

$$Q_R(t)=\int\limits_{|\boldsymbol{x}|\le R} d^3x j_0(\boldsymbol{x},t) \tag{208}$$

i.e. as the integral in bounded space.

Let there exists a time independent operator $A \neq A(t)$. It follows from the current conservation equation that

$$0 = \int_{|x| \le R} d^3x \left[\partial^\mu j_\mu(x), A\right]$$

or

$$\left[\partial_0 Q_R(t), A\right] + \left[\int_{\sigma(R)} d\boldsymbol{\sigma} \cdot \boldsymbol{j}, A\right] = 0$$

If we consider the second term to be zero for very large R, it remains

$$\lim_{R>L} \partial_0 \left[Q_R(t), A\right] = 0$$

or

$$\left[Q_R(t), A\right] = B, \ \ (R > L), \qquad where \ \ \frac{dB}{dt} = 0$$

i.e. B does not depend on time. If above surface integral is zero only in the limit $R \to \infty$, then we would have

$$\lim_{R\to\infty} \left[Q_R(t), A\right] = B, \qquad B \neq B(t) \tag{208}$$

This is a result of local current conservation.

If there exist a limit $Q(t) = \lim_{R\to\infty} Q_R(t)$, then instead of we would have

$$\left[Q(t), A\right] = B \tag{209}$$

However as will be clear below, the existence of this commutator will be more important for our purpose.

Spontaneous symmetry breaking is introduced as follows

$$\langle 0|B|0\rangle \neq 0 \tag{210}$$

I.e. for us the existence of such operator A is sufficient, commutator of which with $Q(t)$ is time independent. It follows from (209) that

$$\langle 0|\left[Q(t), A\right]|0\rangle = \langle 0|B|0\rangle \neq 0 \tag{211}$$

Therefore, if symmetry is spontaneously broken, then the vacuum state is not annihilated by action of $Q(t)$. This means that if we consider $Q(t)$ as a generator of some transformation, then its action does not give a vacuum

$$e^{i\lambda Q}|0\rangle \neq |0\rangle$$

This means that $\exp(i\lambda Q)$ is not a unitary operator.

Consider now action on vacuum of an operator

$$Q(t) = \int d^3x j_0(\boldsymbol{x}, t) \tag{212}$$

$$|Q\rangle = Q|0\rangle \tag{213}$$

We now show that a state is not normalizable. Indeed the norm is

$$\langle Q|Q\rangle = \langle Q|Q(t)|0\rangle = \int d^3x \langle 0|\, j_0(\boldsymbol{x},t) Q|0\rangle$$

It is evident that if Q does not annihilate the vacuum, this expression diverges. Let perform space translations

$$j_0(\boldsymbol{x},t) = e^{i\hat{\boldsymbol{P}}\cdot \boldsymbol{x}} j_0(0,t) e^{-i\hat{\boldsymbol{P}}\cdot \boldsymbol{x}}$$

As usual vacuum is translational invariant, and moreover $\left[Q, \hat{\boldsymbol{P}}\right] = 0$. Therefore we reduce the expression of the norm to

$$\langle Q|Q\rangle = \int d^3x \langle 0|\, j_0(0) Q|0\rangle = \infty$$

I.e. there exists no such Q which does not annihilate the vacuum.

There is a *Theorem of Fabri and Picasso*, according to which if there exist an operator, like (212) even in the sense of weak limit, it must annihilate vacuum.

This means that if the symmetry is broken, then Q does not exist, but it may exist as a commutator with other operators, as in (209).

Let us assume that there happens so, i.e. there exist a commutator such that the result is time independent [see, (208)]. Evidently, this requirement is safer. Let study what follows from (210):

$$\lim_{R\to\infty} \sum_n \left[\langle 0|Q_R(t)|n\rangle\langle n|A|0\rangle - \langle 0|A|n\rangle\langle n|Q_R(t)|0\rangle\right] = \langle 0|B|0\rangle \neq 0$$

Let perform the translations

$$\lim_{R\to\infty} \sum_n \int_R d^3x \left[\langle 0|\, j_0(0)|n\rangle\langle n|A|0\rangle e^{ip_n\cdot x} - \langle 0|A|n\rangle\langle n|\, j_0(0)|0\rangle e^{-ip_n\cdot x}\right]$$

and integrate in space and tend $R\to\infty$. It follows

$$= \sum_n (2\pi)^3 \delta^{(3)}(p_n)\left[\langle 0|\, j_0(0)|n\rangle\langle n|A|0\rangle e^{-ip_n^0 x_0} - \langle 0|A|n\rangle\langle n|\, j_0(0)|0\rangle e^{ip_n^0 x_0}\right] =$$

$$= \langle 0|B|0\rangle \neq 0 \tag{214}$$

B is time independent, but there remains time dependent exponents on the left-hand side. This relation must be valid at any time. So exponents must disappear on the left hand side also. It means that in intermediate states $|n\rangle$, between them necessarily exists the state, energy of which vanishes when 3-momentum vanishes. In other words, there must be zero mass states in the spectrum of $j_0(0)$ *or* Q. It is a maintenance of the Goldstone theorem (1961)

Symmetry realization in considered models

We considered above two realizations of chiral symmetry. It is important that these two models manifest diverse ways for realization of chiral symmetry. There is interesting theorem – the **Coleman** theorem, which orders ways of symmetry realization.

Theorem: This theorem figuratively expresses as: "*Invariance of the vacuum is invariance of the world*"

i.e. If the vacuum is invariant under some symmetry, then Hamiltonian (Lagrangian) is also invariant

Proof is very easy:

Done $Q|0\rangle = 0$ and we must show that it follows $\dot{Q} = 0$

Because $\dot{Q} = i[H,Q]$, then $\dot{Q}|0\rangle = iHQ|0\rangle - QH|0\rangle = iHQ|0\rangle$, (vacuum is translationary invariant). Now we write:

$$\int d^3x \dot{j}^0(\boldsymbol{x},t) = 0 \qquad \Rightarrow \qquad \int d^3x\left(\dot{j}^0 - \nabla \boldsymbol{J}\right)|0\rangle = 0$$

Here the additional divergent term $\nabla \boldsymbol{J}$ disappears owing to the Gauss theorem. So we have

$$\int d^3x \partial_\mu J^\mu(x)|0\rangle = 0$$

It follows that the local operator $\Theta(x) \equiv \partial_\mu J^\mu(x)$ annihilates vacuum

$\int d^3x \Theta(x)|0\rangle = \int d^3x \Theta(0)|0\rangle = 0$. Then it follows

$$\Theta(0)|0\rangle = 0$$

Therefore, according to the Federbush-Johnson theorem we would have $\Theta(x) \equiv 0$, i.e. $\partial_\mu J^\mu(x) = 0$.

So, from invariance of the vacuum under some transformation, we prove that the corresponding local current is conserved.

In the *linear sigma model* we have for transformed creation operator of the pion field

$$\left[Q_5^i, \phi_{p,a}^+\right] = i\delta_{ai}\sigma_p^+$$

i.e. neutral particle arises as partner of pion (degeneracy with π)

In the *non-linear sigma model* $\sigma(x)$ means following series

$$\sigma(x) = f_\pi - \frac{1}{2f_\pi}\phi^2(x) - \frac{1}{8f_\pi^4}\left(\phi^2(x)\right)^2 + \ldots$$

Therefore

$$Q_5^i\,|\,\boldsymbol{p},a\rangle = i\delta_{ai}\sigma_p^+|0\rangle \tag{215}$$

gives on the right hand side states, consisting pairs of pions, with total energy –momentum $\left(\omega_{\boldsymbol{p}}, \boldsymbol{p}\right)$. Evidently it is possible only if pions are massless. Indeed the model Lagrangian

$$L = \frac{1}{2}\left(\partial_\mu \boldsymbol{\phi}\right)^2 + \frac{1}{2}\left(\partial_\mu \sigma\right)^2$$

does not contain mass term. Mass term appears after expansion of symmetry breaking term $\left[f_\pi^2 - \phi^2\right]^{1/2}$. When symmetry is broken, both models give $\langle\sigma(x)\rangle_0 \equiv \sigma_0 \neq 0$ or equivalently

$$\langle 0|\left[Q_5^i, \phi_{p,a}^+\right]|0\rangle = i\delta_{ai}\sigma_0 \neq 0 \tag{216}$$

This means that Q_5^i does not annihilate vacuum. It is not surprising, because now axial current is not conserving and Q_5^i is no more integral of motion (Coleman theorem).

In case of non-linear model it follows from the expression of $\sigma(x)$, that even in the absence of explicit symmetry breaking term

$$Q_5^i|0\rangle \neq 0$$

I.e. symmetry is broken spontaneously and masslessness of pion is a result of the Goldstone theorem.

In the linear model spontaneous breaking is connected to the specific behavior of the potential – the chosen of spontaneous breaking way takes place by artificial choice of model parameters.

Therefore in this respect the non-linear realization is more economic, as it consists a minimal number of parameters, which are known from experiments! ($f_\pi \approx 93 MeV$)--------- ?

In considered above models the $SU(2)\times SU(2)$ chiral symmetry is broken spontaneously till $SU(2)_V$ symmetry.

Lecture 15

Chiral symmetry in the framework of Quantum Chromodynamics (QCD)

In QCD the structure of baryons and mesons are determined by the Color $SU(3)_C$ Lagrangian

$$L = -\frac{1}{4} F_{\mu\nu}^i F_{\mu\nu}^i + \bar{q}_a \left(i\not{D} - m \right) q_a \tag{217}$$

where $F_{\mu\nu}^i \quad (i = 1, 2, ..., 8)$ is the stress tensor of gluon fields from the adjoint representation of $SU(3)_C$ group (octet), and the quark's fields $q_a^\alpha \quad (\alpha = 1, 2, 3)$ are triplets of the same group (fundamental representation). Index a takes as much values as much quark flavors have we: $a = 1, 2, ..., N_f$. D_μ is the usual covariant derivative ($\left(\not{D} = \gamma^\mu D_\mu\right)$

$$D_\mu = I \cdot \partial_\mu + igA_\mu^i(x)\lambda^i$$

λ^i are the Gell-Mann matrices, $m_{ab} = m_a \delta_{ab}$ is a quark's mass matrix (for current quarks)

The main characteristic features of QCD are:
At short distances the color antiscreaning (asymptotic freedom) takes place, and at large distances we have infrared slavery – the gauge constant g increases and only neutral colorless states are observable (color confinement). Characteristic range is determined by cutting parameter $\Lambda_{QCD} \sim 200 MeV$, which fixes the scale of confinement mechanism.

Now, what is the quark mass matrix? The source of mass appearance lies outside the QCD. According to the standard model quark masses arise by Higgs mechanism, which breaks electroweak $SU(2)_L \times U(1)$ symmetry spontaneously till U (1). Therefore if one ignores the Higgs fields, quarks remains massless.

According to modern point of view current masses of $u = q_1 \quad and \quad d = q_2$ are very small in compare to Λ_{QCD}, the mass of $s = q_3$ quark is comparable with Λ_{QCD}, i.e. is also very small. Hence the approximation when these masses are neglected, is not very rough. As regards of heavy quarks (c, b, t) such an approximation is questionable.

Therefore, QCD gives a fresh point of view about global symmetries. If early the approximate equality of proton and neutron masses was the reason of isotopic symmetry, or equivalently u and d quarks masses must be equal, now we know that the isotopic symmetry arises because the masses of u and d quarks are very small, as compared to other quark masses. Moreover, if we neglect quark masses all in all, the chiral symmetry appears.

We are interested in the chiral symmetry in the QCD framework, therefore we neglect quark mass matrix, while this will not be a good approximation except the light quarks.

In this approximation the Lagrangian has the form

$$L^{(0)}{}_{QCD} = i\bar{q}_a \gamma^\mu D_\mu q_a - \frac{1}{4} F^i_{\mu\nu} F^i_{\mu\nu} \tag{218}$$

Let us define the left and right quarks

$$q_L = \frac{1-\gamma_5}{2} q, \qquad q_R = \frac{1+\gamma_5}{2} q \tag{219}$$

Then the quark sector of Lagrangian looks like

$$L^q_{QCD} = i\bar{q}_{La} \gamma^\mu D_\mu q_{La} + i\bar{q}_{Ra} \gamma^\mu D_\mu q_{Ra} \tag{220}$$

This part of Lagrangian and therefore (218) is invariant under unitary transformation on the index a for left and for right quarks separately. hence there appears a symmetry

$$U(N_f) \times U(N_f) \equiv U(N_f)_L \times U(N_f)_R = \{(U_L, U_R)\} \tag{221}$$

under which quarks transform as

$$q_L \to q'_L = U_L q_L, \qquad q_R \to q'_R = U_R q_R \tag{222}$$

For calculating the Noether's current, as we know already, transformation law must be made temporarily local. Then the divergence with respect to difference of infinitesimal parameter arises only from terms like $\bar{q}_i \gamma^\mu \partial_\mu q_i$ and the Gell-Mann-Levy equations give for currents

$$J^a_{L\mu} = \bar{q}_L \gamma^\mu \frac{\lambda^a}{2} q_L = \bar{q} \gamma^\mu \frac{1-\gamma_5}{2} \frac{\lambda^a}{2} q, \qquad a = 0, 1, 2, ..., N_f^2 - 1 \tag{223}$$

$$J^a_{R\mu} = \bar{q}_R \gamma^\mu \frac{\lambda^a}{2} q_R = \bar{q}\gamma^\mu \frac{1+\gamma_5}{2}\frac{\lambda^a}{2} q, \qquad a = 0,1,2,\ldots,N_f^2 - 1 \tag{224}$$

other terms in Lagrangian do not change, so we have conservation of $2N_f^2$ currents on the classical level,

$$\partial^\mu J^a_{L\mu} = 0 \qquad \partial^\mu J^a_{R\mu} = 0, \qquad \left(a = 0,1,2,\ldots,N_f^2 - 1\right) \tag{225}$$

Generators are defined as usual

$$Q^a_L = \int d^3x J^a_{L0}\left(\boldsymbol{x},t\right) \qquad Q^a_R = \int d^3x J^a_{R0}\left(\boldsymbol{x},t\right) \tag{226}$$

Therefore

$$U_L = e^{iQ_L}, \qquad U_R = e^{iQ_R} \tag{227}$$

These quantities correspond to our earlier $SU(2)\times SU(2)$ chiral charges as follows

$$\begin{aligned} &Q_L \to Q_-, && Q_R \to Q_+ \\ &Q \to \frac{1}{2}\left(Q_R + Q_L\right), && Q_5 \to \frac{1}{2}\left(Q_R - Q_L\right) \end{aligned} \tag{228}$$

Let us now consider various subgroups:

(i) $\quad U\left(1\right)_V$

The generator is $Q = Q_V$, i.e. it is an Abelian subgroup of transformation, which appears because among N_f^2 matrices λ^a, there is unit matrix also. There corresponds to this subgroup a multiplication on common phase factor

$$q_L \to e^{i\theta} q_L, \qquad q_R \to e^{i\theta} q_R, \qquad or \qquad q \to e^{i\theta} q \tag{229}$$

Corresponding conserved charge in QCD is a baryon number. This symmetry is exact. It means, that this symmetry does not break neither spontaneously, not by anomaly after quantization.

(ii) $\quad U\left(1\right)_A$

The generator is $Q_5 = Q_A$. It is so called "chiral" $U(1)$ subgroup

$$q \to e^{i\theta\gamma_5} q \tag{230}$$

in this case left and right quarks acquire opposite sign phases

$$q_L \to e^{-i\theta} q_L, \qquad q_R \to e^{i\theta} q_R \tag{231}$$

This symmetry is not exact – it is broken after quantization of QCD, by so-called axial anomaly. Corresponding current $J_{5\mu} = \bar{q}\gamma_\mu \gamma_5 q$ is not conserving

$$\partial^\mu J_{5\mu}\left(x\right) = -\frac{iN_f}{16\pi^2}\varepsilon_{\mu\nu\alpha\beta} F^i_{\mu\nu} F^i_{\alpha\beta} \tag{232}$$

(iii) $\quad G \equiv SU(N_f)_L \times SU(N_f)_R$

It is a $\{(L,R)\}$ subgroup, where L *and* R are matrices with determinant equal to 1. In quantum theory it is desirable that this group could be broken spontaneously till vector subgroup $H=\{(V,V)\}$, which transforms q_L *and* q_R in the same manner. This is a subgroup $SU(N_f)$: for 2 quarks it is $H=SU(2)$ for isospin, for 3 quarks – it is the eightfold way, $SU(3)$.
As G is spontaneously broken till H, according to Goldstone theorem we'll have N_f^2-1 massless bosons – pions in $SU(2)$, pseudo scalar octet in $SU(3)$.

Now our aim will be to construct the effective Lagrangian for this symmetry, i.e. the Lagrangian which describes the dynamics of Goldstone particles. We must require such properties which follow from the massless quark's QCD. they are:

(a) L_{eff} must be invariant under the group $G\equiv SU(N_f)_L\times SU(N_f)_R$. Therefore it has to be constructed by multicomponent Φ field.

(b) This Φ field should be transformed by action of $G\equiv SU(N_f)_L\times SU(N_f)_R$ and offers exactly N_f^2-1 degrees of freedom at any space-time point. It is the requirement of minimality – we want describe only Goldstone modes. Evidently, the model can be enlarged by inclusion of other fields, e.g. vector or axial bosons.

(c) We have to require that the subgroup of G, which remains invariant the arbitrary values of Goldstone field must be exactly H, and nothing more.

The last two requirements determine that the space of field values must be a factor-space G/H

In our case $G=\{(L,R)\}$ and $H=\{(V,V)\}$, therefore G/H manifold coincides with $SU(N_f)$ manifold. It follows that the field Φ can be identified with the U field, where $U(\boldsymbol{x},t)$ is an element of the $SU(N_f)$ group.
G group acts on U in the following way

$$U\to LUR^{+} \tag{233}$$

Because the dimension of $SU(\mathrm{N}_f)$ is N_f^2-1, the requirement (b) is satisfied, Now we can take as typical value for $U(\boldsymbol{x},t)$ the identity matrix I. a little group that remains invariant the unit matrix is $\{(V,V)\}$:

$$VIV^{+}=I$$

it is evident that I is not peculiar value somehow for $U(\boldsymbol{x},t)$, because its little group is

$$\{(U(x,t)VU^{+}(x,t),V)\}$$

It leaves unchanged any value of $U(\boldsymbol{x},t)$, because $UVU^{+}U\cdot V^{+}=U$.
It is also isomorphic to H.

Therefore the requirement (c) is also valid.

Construction of chiral invariant Lagrangian

Now we are going to construct the corresponding chiral invariant Lagrangian. As a rule the strategy is the following: One introduces the covariant vector called a Mauri-Cartan form

$$L_\mu = U^+ \partial_\mu U \tag{234}$$

Because the group acts as

$$U \to LUR^+, \qquad U^+ \to RU^+L^+$$

it follows

$$L_\mu \to RU^+L^+L\partial_\mu UR^+ = RU^+\partial_\mu UR^+$$

or

$$L_\mu \to RL_\mu R^+ \tag{235}$$

i.e. L_μ is invariant under $SU(N_f)_L$. It is called a Mauro-Cartan “left-invariant form”

We can define by analogy the “Right-invariant form” (co-vector)

$$R_\mu = \partial_\mu UU^+ \tag{236}$$

It is clear that

$$R_\mu \to LR_\mu L^+$$

So, it is $SU(N_f)_R$ invariant.

Sometimes the “right-invariant” form is defined as $R_\mu = U\partial_\mu U^+$. Obviously it is also $SU(N_f)_R$ invariant.

Now let us mention that arbitrary chiral invariant polynomial $f\left(L_\mu\right)$, constructed from L_μ, coincides with $f\left(R_\mu\right)$ and vice versa. This happens because L_μ *and* R_μ are connected by the relation

$$R_\mu = UL_\mu U^+ \tag{237}$$

Therefore

$$f\left(R_\mu\right) = f\left(UL_\mu U^+\right) = f\left(L_\mu\right) \tag{238}$$

Thus, we can construct the chiral invariant Lagrangian both as with the help of L_μ or R_μ. By historical reasons L_μ is used for this aim. The example of simplest Lagrangian is

$$Tr\left(L_\mu L^\mu\right) \tag{239}$$

Topology of the non-linear sigma model

Let us now return to $SU(2)\times SU(2)$ model. Above described geometrical picture fully corresponds to sigma model, in the form suggested by Sugawara. The above considered representation M plays a role of $U(x)$:

$$U(x)=\frac{1}{f_\pi}\left[\sigma(x)+i\boldsymbol{\tau}\cdot\boldsymbol{\pi}(x)\right] \qquad (240)$$

where fields are constraint by the relation (because $DetU=1$)

$$\sigma^2+\boldsymbol{\pi}^2=f_\pi^2 \qquad (241)$$

Let us construct a left-invariant form

$$L_\mu=U^+\partial_\mu U \qquad (242)$$

and a Lagrangian

$$L=\frac{f_\pi^2}{4}Tr\left[L_\mu L^\mu\right] \qquad (243)$$

which may be rewritten as (using $U^+U=1$)

$$L=\frac{f_\pi^2}{4}Tr\left[\partial_\mu U\partial^\mu U^+\right] \qquad (244)$$

Sometimes Eq. (243) is named as Sugawara form.
If we take the pion field to be weak, then

$$L_\mu\sim -R_\mu\sim\frac{i}{f_\pi}\boldsymbol{\tau}\cdot\partial_\mu\boldsymbol{\pi} \qquad (245)$$

There is also other very useful parametrization, which was first introduce by Gursey

$$U=e^{i\boldsymbol{\tau}\cdot\boldsymbol{\phi}(x,t)} \qquad (246)$$

Evidently

$$U(\boldsymbol{x},t)=\cos\phi+i\boldsymbol{\tau}\cdot\hat{\boldsymbol{\phi}}\sin\phi; \qquad \phi=|\boldsymbol{\phi}|, \qquad \hat{\boldsymbol{\phi}}=\boldsymbol{\phi}/\phi \qquad (246a)$$

In this notation only one, isovector field $\boldsymbol{\phi}(\boldsymbol{x},t)$ remains, which is connected to $(\sigma,\boldsymbol{\pi})$ fields by relations

$$\sigma=f_\pi\cos\phi, \qquad \boldsymbol{\pi}=f_\pi\hat{\boldsymbol{\phi}}\sin\phi \qquad (247)$$

As we know, the chiral group transforms U as follows

$$U\to U'=LUR^+=LUR^{-1} \qquad (248)$$

and the Lagrangian (243) is invariant. At the same time the matrix $U=1$ corresponds to $\boldsymbol{\phi}=0$ (or $\boldsymbol{\pi}=0$), which is the vacuum value and it is not invariant under these transformation, accepts the case, when $R=L$. It is exactly the spontaneous symmetry breaking till to subgroup $SU(2)_V$.
If we expand U around unit matrix, and limit ourselves to the first order term

$$U\approx 1+i\boldsymbol{\tau}\cdot\hat{\boldsymbol{\phi}}$$

the Lagrangian takes form

$$L_0 = \frac{f_\pi^2}{2}\partial_\mu \boldsymbol{\phi} \cdot \partial^\mu \boldsymbol{\phi} = \frac{1}{2}\partial_\mu \boldsymbol{\pi} \cdot \partial^\mu \boldsymbol{\pi}$$

which coincide to massless pion Lagrangian. Therefore we have to recognize a pion as a fluctuation of U field near the unity matrix, I .

Lecture 16

Topological properties of the non-linear sigma model

Parametrization, given above (246), considers the fundamental field $\boldsymbol{\phi}(x)$ as an angular variable (phase), which takes his values on 3-dimensional unit sphere inserted into the 4-dimensional space of internal symmetry $\{SU(2)\times SU(2) \sim SO(4)\}$. When x runs space-time points, $\boldsymbol{\phi}(x)$ moves on this S^3 sphere. It seems that $\boldsymbol{\phi}(x)$ is periodic as all angular variable. Therefore it is not determined uniquely by the physical state.

Let us consider a configuration with finite energy, which is derived when U is static, $U = U(\boldsymbol{x})$

$$E = \frac{f_\pi^2}{4}\int d^3x Tr\left(\nabla U \cdot \nabla U^+\right) \tag{249}$$

It is clear that to guarantee the finiteness we need that when $|\boldsymbol{x}| \to \infty$ in arbitrary direction $U = U(\boldsymbol{x})$ necessarily tends to the constant matrix. Such $U(\boldsymbol{x})$ determines the mapping of 3-dimensional configuration space R^3 onto to internal sphere, S^3. Therefore one can perform a topological classification of $U(\boldsymbol{x})$ configurations, i.e. all such mappings must be divided into various *topological sectors, with accordance of that how much times S^3 sphere will be covered when $\boldsymbol{x}$ takes all its values from R^3.* This number is called as *topological index* and is used for characterization of various *topological sectors.*

In our example of non-linear model, the solutions with finite energy are the results of space-topology of $U(\boldsymbol{x})$ fields. This topology from its point of view is sensitive on the boundary conditions imposing at the infinity. As we say above the sufficient condition for finite energy configurations is a requirement

$$U(\boldsymbol{x},t) \to U_0, \ \ as \ \ |\boldsymbol{x}| \to \infty \tag{250}$$

where U_0 is a constant matrix. Moreover the tending $U \to U_0$ must be sufficient for providing finiteness of energy [for example, $U_0^{-1}U = I + O\left(r^{-2}\right), \quad when \quad r \to \infty$]

From the invariance of Lagrangian under the chiral transformation follows that U_0 can be reduced to a unit matrix, for this it is sufficient perform a global rotation: $U \to U_0^{-1}U$. (See, above discussion about a little group). In case of such chiral rotation we do not loss any physical information. Therefore, one chooses the following boundary condition

$$U \to I, \qquad as \quad r \to \infty \tag{251}$$

i.e. U field tends to I, but not to some angular dependent limit. Therefore, we can think that all points of space infinity are identified to one point. Such an identification converts the 3-dimensional Euclidean space with coordinates $\boldsymbol{x}$ in the fixed moment of time into the 3-dimensional sphere. According to boundary condition, therefore, U field is now defined on this S^3 sphere.

This mapping in mathematics is realized by the stereographic projection –stereographic coordinates are introduced on S^3 in the way:

$$\xi_0 = \frac{1-r^2}{1+r^2}, \qquad \boldsymbol{\xi} = \frac{2\boldsymbol{x}}{1+r^2} \tag{252}$$

where $\xi_0^2 + \boldsymbol{\xi}^2 = 1$. The inverse transformation has a form

$$\boldsymbol{x} = \frac{\boldsymbol{\xi}}{1+\xi_0}$$

Coordinates $\xi_\mu\left(\xi_0, \boldsymbol{\xi}\right)$ draw S^3 sphere. But this last is a compact manifold, on the contrary of R^3. This change of topology comes about because the infinite points of R^3 are mapped into one point – the south pole of S^3 $\left(\xi = \left(-1, 0\right)\right)$. This is permitted by means of the boundary condition for U (251). $\boldsymbol{x} \to \boldsymbol{\xi}$ transformation generates on S^3 well-defined functions (counterexample, $f\left(\boldsymbol{x}\right) = \left|\boldsymbol{x}\right|$ is continuous in R^3, but after transformation it no more is continuous on S^3, because the singularity arises in the South Pole).

Now our aim is to study the topology of the configuration space.

An arbitrary $SU(2)$ matrix has a form

$$n_0 + i\boldsymbol{\tau} \cdot \boldsymbol{n}, \quad where \quad n_0^2 + \boldsymbol{n}^2 = 1$$

Thus U fields are defined on S^3.

If U_0 and U_1 are two such fields, then we say that they are *homotopic,* $\left(U_0 \sim U_1\right)$ if we can continuous deformation of U_0 into U_1. In other words:

We say that $\left(U_0 \sim U_1\right)$, if there is set of mappings $S^3 \rightarrow S^3$, denoted by $U_{(\tau)}$ such that $U_{(\tau)}\left(\boldsymbol{x}\right)$ is continuous with respect of both $\boldsymbol{x}$ and τ and $U_{(0)} = U_0$ *and* $U_{(1)} = U_1$. In this case the problem is in enumeration of homotopic ally inequivalent mappings.

Non-trivial homotopic sector – homotopically inequivalent mappings

Let first consider the trivial field $U^{(0)}$, which transforms all $\boldsymbol{x}$ into one and the same point of $SU(2)$, which is the unit matrix, I. We can connect to $U^{(0)}$ all mappings U, which are homotopic to $U^{(0)}$. All such mappings form a trivial sector, Q_0. The usual pion physics was studied in this sector.

Let us now construct a non-trivial sector, say $U^{(1)} \in Q_1$. In order to guess how the corresponding field looks like, note that the correspondence between a 3-dimensional space with identified infinity and 3-dimensional sphere may be established by relations

$$N^0 = \cos\theta\left(r\right), \qquad N^i = n^i \sin\theta\left(r\right) \tag{253}$$

where $N^{\alpha}\,(\alpha = 0,1,2,3)$ is a unit radius-vector in 4-dimensional space, which parametrizes points of S^3. $n^i\left(i = 1,2,3\right)$ is an unit radius-vector in our 3-dimensional space, and $\theta\left(r\right)$ is arbitrary monotonic function of r with boundary conditions

$$\theta\left(0\right) = \pi \qquad \theta\left(\infty\right) = 0, \qquad \frac{d\theta\left(r\right)}{dr} = \theta'\left(r\right) < 0 \tag{254}$$

This last requirement guarantees that $\theta\left(r\right)$ decreases monotonically from π to zero, when r grows from 0 to ∞ and at the same time receives all the values between π and zero exactly once. Therefore the polar coordinates of the points on S^3 are

$$\left(\cos\theta\left(r\right), \hat{\boldsymbol{x}}\sin\theta\left(r\right)\right) \tag{255}$$

When $\boldsymbol{x}$ run all its values, this point reaches on S^3 to all values once and only once. In other words: by this $U^{(1)}$ map $SU(2)$ will cover exactly once. This is an example of such map, for which the so-called "winding number" equals to one. Like this mapping's homotopic mappings compose the class of equivalent mappings or the Q_1 sector with winding number 1.

In case of charge conjugation resulting map has a winding number $\left(-1\right)$, so the total sector, Q_{-1}.

A typical mapping with winding number n $\left(n = 0, \pm1, \pm2, \ldots\right)$ should be the nth degree of Q_1, $\left(U^{(1)}\right)^n$.

Physical meaning of homotopic classification

Let us consider the initial conditions for equation of motion at $t=0$ moment

$$t=0: \qquad \left(U^{(n)},\dot{U}^{(n)}\right), \qquad\qquad U^{(n)}\in Q_n$$

After some time T these values will change:

$$U^{(n)}\to\tilde{U}^{(n)}, \qquad \dot{U}^{(n)}\to\dot{\tilde{U}}^{(n)}$$

If $U\left(\boldsymbol{x},t\right)$ is a solution of equation of motion with a given initial conditions, it means that

$$U\left(\boldsymbol{x},0\right)=U^{(n)}\left(\boldsymbol{x}\right), \qquad\qquad U\left(\boldsymbol{x},T\right)=\tilde{U}^{(n)}\left(\boldsymbol{x}\right) \tag{256}$$

Because the time evolution is a continuous operation, it follows that $U^{(n)}$ and $\tilde{U}^{(n)}$ are homotopic to each other's. The homotopy may be realized with the aid of function

$$U\left(\boldsymbol{x},\tau T\right), \qquad \tau\in\left[0,1\right], \tag{257}$$

therefore $\tilde{U}^{(n)}\in Q_n$ and, so, the connected to U field the integer number, n, is a constant of motion, i.e. the characteristic index of homotopic classes is a constant of motion.

Lecture 17

The topological charge of non-linear sigma model

We have mentioned that in non-linear sigma model there is mapping

$$U\left(x\right): \quad S^3\to S^3 \tag{258}$$

This mapping is non-trivial. Homotopic sectors Q_n are characterized by integer number – called a winding number, which is conserved topologically. There corresponds a topological current, which has a form

$$J_\mu=\frac{1}{24\pi^2}\varepsilon_{\mu\nu\lambda\rho}Tr\left(L_\nu L_\lambda L_\rho\right) \tag{259}$$

Let us calculate the divergence

$$\partial^\mu J_\mu=\frac{1}{24\pi^2}\varepsilon_{\mu\nu\lambda\rho}Tr\left\{\partial^\mu L_\nu L_\lambda L_\rho+L_\nu\partial^\mu L_\lambda L_\rho+L_\nu L_\lambda\partial^\mu L_\rho\right\}$$

It follows from the definition that L_ν satisfies the Mauro-Cartan identity

$$\partial_\mu L_\nu-\partial_\nu L_\mu+\left[L_\mu,L_\nu\right]=0 \tag{260}$$

Using this identity a typical term in divergence equation may be rewritten as

$$\varepsilon_{\mu\nu\lambda\rho}\partial^{\mu}L_{\nu} = \frac{1}{2}\varepsilon_{\mu\nu\lambda\rho}\partial^{\mu}L_{\nu} + \varepsilon_{\mu\nu\lambda\rho}\partial^{\nu}L_{\mu} =$$
$$= \frac{1}{2}\varepsilon_{\mu\nu\lambda\rho}\left(\partial^{\mu}L_{\nu} - \partial^{\nu}L_{\mu}\right) = -\frac{1}{2}\varepsilon_{\mu\nu\lambda\rho}\left[L_{\mu}, L_{\nu}\right]$$

Therefore in divergence equation the typical term looks like

$$\varepsilon_{\mu\nu\lambda\rho}TrL_{\mu}L_{\nu}L_{\lambda}L_{\rho} \Rightarrow \varepsilon_{\nu\lambda\rho\mu}TrL_{\nu}L_{\lambda}L_{\rho}L_{\mu} = -\varepsilon_{\mu\nu\lambda\rho}TrL_{\mu}L_{\nu}L_{\lambda}L_{\rho} = 0 \qquad (261)$$

Thus it is proved that the vanishing divergence equation takes place geometrically, i.e. independent of equation of motion. It follows that the connected with this current charge is integral of motion,

$$B = \frac{1}{24\pi^2}\varepsilon_{ijk}\int d^3xTr\left(L_iL_jL_k\right) \qquad (262)$$

It is clear that this expression may be written in covariant form as well:

$$B = \frac{1}{24\pi^2}\varepsilon^{0\nu\alpha\beta}\int d^3xTr\left(L_{\nu}L_{\alpha}L_{\beta}\right) \qquad (262a)$$

Starting from the definition of L_{μ} one can show that $Tr\left(L_iL_jL_k\right)$ is a chiral invariant. Therefore its calculation is possible for arbitrary values of pion field. In particular, when this field is small, $L_{\mu} \to i\boldsymbol{\tau}\cdot\partial_{\mu}\boldsymbol{\phi}$. Then

$$J_0 = \frac{1}{24\pi^2}\varepsilon_{0\nu\lambda\rho}Tr\left(i\boldsymbol{\tau}\cdot\partial_{\nu}\boldsymbol{\phi}\right)\left(i\boldsymbol{\tau}\cdot\partial_{\lambda}\boldsymbol{\phi}\right)\left(i\boldsymbol{\tau}\cdot\partial_{\rho}\boldsymbol{\phi}\right) =$$
$$= \frac{1}{24\pi^2}\varepsilon_{0\nu\lambda\rho}\varepsilon^{mnk}\partial_{\nu}\phi^m\partial_{\lambda}\phi^n\partial_{\rho}\phi^k$$

After d^3x integration, we obtain

$$B = \frac{1}{24\pi^2}\int d^3x\varepsilon^{mnk}\varepsilon_{0\nu\lambda\rho}\partial_{\nu}\phi\partial_{\lambda}\phi\partial_{\rho}\phi \qquad (263)$$

We see that there appear the Jacobian of $R^3 \to S^3$ transformation. Therefore B is a winding number. The normalization factor is chosen so, that

$$\frac{1}{24\pi^2} = \frac{1}{3!2\pi^2}$$

Where $2\pi^2$ is the area of S^3 surface in R^4.

The size of soliton and the Skyrme term

In the previous considerations we made certain that the non-linear sigma model has a non-trivial topological structure. The field configurations are divided into the homotopically non-trivial sectors, each of them characterized by definite winding numbers (winding number often is named as the Pontryagin or Chern-Simons indice).

Configurations with finite energy have a localized energy densities in the finite area of R^3 space. Therefore, when we find a stable configurations with non-zero winding number satisfying above given boundary condition, then they correspond objects having extending particle properties – they are localized in space and their conserved current satisfied to continuous equation.

Unfortunately there are not stable configuration, which ensure minimization of the potential part of sigma model Lagrangian. Indeed, above we had a scale arguments about it, and now we can repeat it for this model .The energy is

$$E = \frac{f_\pi^2}{4}\int d^3x Tr\left(\partial^i U^+ \partial^i U\right) \tag{264}$$

Introduce a new field,

$$U_\lambda(\boldsymbol{x}) = U(\lambda \boldsymbol{x}), \qquad \lambda > 0$$

and calculate the energy for it

$$E_\lambda = \frac{f_\pi^2}{4}\int d^3x Tr\left(\partial^i U_\lambda^+ \partial^i U_\lambda\right) = \frac{1}{\lambda}E \tag{265}$$

It follows

$$\frac{dE_\lambda}{d\lambda} = -\frac{1}{\lambda^2}E \tag{266}$$

It is evident that the minimum energy corresponds to $\lambda = \infty$, i.e. energetically preferable state for finite size object is zero energy state. Or for finiteness of energy, particle size must be tend to zero. This means that in the non-linear sigma model solutions corresponding to a finite energy are not stable under the scale transformation (as we know from the Derick theorem)

The physical reason of this result is clear. The only parameter with the dimension of energy is a pion decay constant, f_π . If soliton solution has a characteristic size, say R , then its energy will be have order of $f_\pi^2 R$ .Therefore the ground state corresponds the limit $R \to 0$. In other words only the sigma model Lagrangian is unable to ensure a stability of soliton with finite size and finite energy: any such configuration suffers dissipation of energy because of pion radiation and are shrink to the point particle with zero energy – it is a particle with zero mass.

These consideration dictates the further strategy, if we want to have a soliton. It is necessary inclusion new terms to the Lagrangian. As the chiral invariants must be constructed by L_μ vectors, the even numbers of them is needed, i.e. the nearest additional term must contain a product of four space derivatives.

It must been noted that the effective Lagrangians are not defined uniquely. They are not bounded by the requirement for renormalizability, therefore they can contain U fields and their derivatives in arbitrary degrees. When we are interested by low energy phenomena, we can imagine that the Lagrangian is expanded in degrees of these derivatives and one can single out the leading terms. Exactly such leading term will be of the fourth order in our case. If one repeat consideration after inclusion of such terms, one obtains

$$E_\lambda = \frac{1}{\lambda} E^{(2)} + \lambda E^{(4)} \tag{267}$$

where $E^{(2),(4)}$ are corresponding energies of quadratic and fourth order terms. Then, the extremum condition looks like

$$\left.\frac{dE}{d\lambda}\right|_{\lambda=1} = \left(-\frac{1}{\lambda^2} E^{(2)} + E^{(4)} \right)_{\lambda=1} = 0$$

which has a solution

$$\lambda^2 = \frac{E^{(2)}}{E^{(4)}}, \qquad or \qquad E^{(4)} = E^{(2)} \tag{268}$$

It is expected that in this case we can have a finite size soliton.

The Skyrme Lagrangian

If Lagrangian must contain only by first derivative terms there remain only possible two fourth order terms, which at the same time satisfy the requirement of chiral symmetry, namely

$$\begin{aligned} L_1^{(4)} &= \left(\partial_\mu \boldsymbol{\pi} \cdot \partial^\mu \boldsymbol{\pi} \right)^2 \\ L_1^{(4)} &= \left(\partial_\mu \boldsymbol{\pi} \cdot \partial_\nu \boldsymbol{\pi} \right) \left(\partial^\mu \boldsymbol{\pi} \cdot \partial^\nu \boldsymbol{\pi} \right) \end{aligned} \tag{269}$$

There is no argument to prefer the first or second forms. But it is easy to convince that their difference

$$L_1^{(4)} - L_2^4$$

contains the time derivative only two times, exactly so, as the sigma Lagrangian. It is suitable for quantization procedure.

This is exactly the combination which was suggested by Skyrme (1961), for stabilization his Hedgehog like solution.

This tern has the following form by the means of L_μ fields

$$L^{(4)} = \frac{e^2}{4} Tr\left[L_\mu, L_\nu \right]^2 \tag{270}$$

The operatorial dimension of L_μ is $\left[m \right]^1$, therefore e is dimensionless parameter.

Thus the Skyrme model Lagrangian has the form

$$L_{Sk} = -\frac{f_\pi^2}{4} Tr\left(L_\mu L^\mu \right) + \frac{e^2}{4} Tr\left[L_\mu, L_\nu \right]^2, \qquad e > 0 \tag{271}$$

The stabilization of the soliton may be understand physically as follows: If R is a soliton size, then the Skyrme term (270) gives following contribution $\sim \frac{e^2}{R}$, therefore the total energy will be

$$c_1 f_\pi^2 R + c_2 \frac{e^2}{R}, \qquad c_{1,2} > 0 \tag{272}$$

This expression has a minimum for nonzero R.

Limitation of Energy by topological charge

Let us show that the energy of soliton is bounded from below by topological charge. Indeed, the energy of static configuration is equal to

$$E = \int d^3x Tr\left\{ -\frac{f_\pi^2}{4} L_i^2 - \frac{e^2}{4}\left[L_i, L_j\right]^2 \right\} \tag{273}$$

Let transform the second term in this way

$$\left[L_i, L_j\right]^2 = \left(\sqrt{2}\varepsilon_{ijk} L_j L_k\right)^2$$

Then

$$E = -\frac{f_\pi^2}{4}\int d^3x Tr\left\{ L_i^2 + \frac{e^2}{f_\pi^2}\left(\sqrt{2}\varepsilon_{ijk} L_j L_k\right)^2 \right\} \tag{274}$$

It is known from algebra that the arbitrary antihermitian matrix A obeys to inequality

$$TrA^2 \le 0 \tag{275}$$

Our matrices L_i, $\varepsilon_{ijk} L_j L_k$ both are antihermitian. Let compose the antihermitian combination

$$A_i = \frac{f_\pi}{2}\left(L_i - \sqrt{2}\frac{e}{f_\pi}\varepsilon_{ijk} L_j L_k \right) \tag{276}$$

Then

$$TrA_i^2 = \frac{f_\pi^2}{4} Tr\left[L_i^2 - 2\sqrt{2}\frac{e}{f_\pi}\varepsilon_{ijk} L_i L_j L_k + \frac{e^2}{f_\pi^2}\left(\sqrt{2}\varepsilon_{ijk} L_j L_k\right)^2 \right] \le 0$$

Therefore Eq.(274) transforms like

$$E \ge \frac{f_\pi^2}{4}\int d^3x \left| Tr\left(\frac{2\sqrt{2}e}{f_\pi}\varepsilon_{ijk} L_i L_j L_k \right)\right| =$$

$$= \frac{f_\pi \sqrt{2}e}{2}\left| \int d^3x \varepsilon_{ijk} L_i L_j L_k \right| = 12\sqrt{2}\pi^2 e f_\pi \left|B\right|$$

Therefore we have derived the bound on energy

$$E \ge 12\sqrt{2}\pi^2 e f_\pi \left|B\right| \tag{277}$$

where B is the topological charge (262a)

This inequality is known as Bogomolny bound, sometimes as Bogomolny-Prasad-Sommerfeld (BPS) bound. According to this inequality the lowest value must correspond to equality sign in (277). if this happens, the solution should be a self-dual:

$$L_i = \frac{2\sqrt{2}}{f_\pi} \varepsilon_{ijk} L_j L_k \tag{278}$$

But it is easily seen, that the self-duality contradicts to the Mauro-Cartan equation (260). Therefore the energy exceeds Bogomolny bound (Bogomolny bound is not saturated)

$$E > 12\sqrt{2}\pi^2 e f_\pi |B| \tag{279}$$

Here it is important that the soliton energy (mass) is bounded from below by the topological charge.

Lecture 18

Skyrmion

Skyrme's initial idea was to connect above defined topological current to the baryonic current and the conserved topological charge to the baryon number. This idea means the following: $U(\boldsymbol{x},t)$ field configurations, which in the vicinity of $U = I$ describes interacting pions, acquires new features when the topological charge differs from zero.

We have seen already that the Skyrme model consists spontaneously broken $SU(2)_L \times SU(2)_R$ chiral non-linear sigma model in the leading order, which satisfies requirements of algebra. But the stability of corresponding non-trivial configurations provides the Skyrme term. These configurations are called Skyrme solitons, or skyrmions. In general, as we know, soliton is classical, static, stable configurations with finite energy in weakly interacting non-linear field theories of bosons only, which are characterized by degenerate vacuum state. Solitons are heavy objects with exactly conserved topological charges. Soliton-soliton interaction is strong, but soliton-boson interaction is weak. After quantization solitons manifest a rich spectra. We'll see, that the skyrmions has many of features, listed here.

In his original papers Skyrme was convince that the field configurations with unit winding number $(B = 1)$ must be fermions in this model. According to him abovementioned topological current must be identified with baryonic current, which means, that skyrmions – are classical baryons. This suggestion was confirmed only 20-30 years after, in the frame of QCD, which will be elucidated in the forthcoming sections.

Classical equations of motion can be deduced from the Skyrme Lagrangian, constraining matrices by unitarity condition, $U^{+}U = 1$ or from the action functional in the first order with respect of fluctuations. We obtain the following equation:

$$\partial^{\mu} L_{\mu} - 2\frac{e^2}{f_{\pi}^2}\partial^{\mu}\left[L_{\nu},\left[L_{\mu},L_{\nu}\right]\right] = 0 \qquad (280)$$

The equation for R_{μ} looks analogously.

Usually Bogomolny constraint or the self-duality is used for simplification of similar equations. We have underlined above that in the Skyrme model self-duality contradicts to the Mauro-Cartan identity. Therefore we are not able in using this method here and as a result we are deal with very complicated equation, which can be studied only by numerical analysis.

Investigation of this equation is possible by symmetry considerations. As we know using the angular parametrization U matrix can be written as

$$U(\boldsymbol{x}) = \cos\phi + i\boldsymbol{\tau}\cdot\hat{\boldsymbol{\phi}}\sin\phi$$

If in course of variation of $\boldsymbol{x}$ the matrix $U(\boldsymbol{x})$ covers the S^3 sphere in $\boldsymbol{x}$ space completely then the unit vector $\hat{\boldsymbol{\phi}}$ must cover unit sphere S^2 in the isotopic space for arbitrary values of ϕ. In other words, the unit isovector $\hat{\boldsymbol{\phi}}$ as a function of $\boldsymbol{x}$ must cover total solid angle 4π in isospace, when $\boldsymbol{x}$ runs all such values in 3-space, for which ϕ has a constant value. The simplest way to reach this is a choice

$$\hat{\boldsymbol{\phi}}(\boldsymbol{x}) = \hat{\boldsymbol{x}}$$

This result can be understand also as follows: When $B \neq 0$ the matrix $U(\boldsymbol{x})$ cannot be translationary invariant, because in this case the field should be constant, but only $B = 0$ corresponds to constant field.

At the same time, when $B \neq 0$, the matrix $U(\boldsymbol{x})$ cannot be invariant under rotations, because such field would depended on distance r only, and in this case

$$U^{+}\partial_i U = \hat{x}_i U^{+}\partial_r U$$

and it follows for current density that

$$\varepsilon_{ijk} U^{+}\partial_i U U^{+}\partial_j U U^{+}\partial_k U = 0$$

There is one extra possibility – construct fields, which are invariant under generalized rotations

$$Diag\left[SU(2)_L \times SU(2)_R\right] \sim Diag\left[SO(3)_I \times SO(3)_J\right]$$

where $SO(3)_{I,J}$ denotes rotations in isotopic and ordinary spaces, respectively: The effect of a spatial rotation can be compensated by an isospin transformation.

Such a field is invariant under combine rotations in both spaces:

$$-i[\boldsymbol{x}\times\nabla]_i U(\boldsymbol{x})+\left[\frac{\boldsymbol{\tau}_i}{2},U(\boldsymbol{x})\right]=0$$

It is easy to verify that the general solution of this equation is

$$U_C(x)=\cos F(r)+i\tau\cdot\hat{x}\sin F(r)=\exp\{i\tau\cdot\hat{x}F(r)\} \tag{281}$$

The generator of this combined rotation is

$$\boldsymbol{K}=\boldsymbol{J}+\boldsymbol{I}=(\boldsymbol{L}+\boldsymbol{S})+\boldsymbol{I}$$

One can easily verify that

$$[\boldsymbol{K},U_C(\boldsymbol{x})]=0 \tag{282}$$

Resulting solution forms a Skyrme ansatz, which has a figurative name – "hedgehog", according its geometric picture: In all space points $\boldsymbol{x}$ corresponding isovector $\hat{\boldsymbol{\phi}}(\boldsymbol{x})$ is directed radially with respect to the origin $\boldsymbol{x}=0$, where the center of considered object is located.

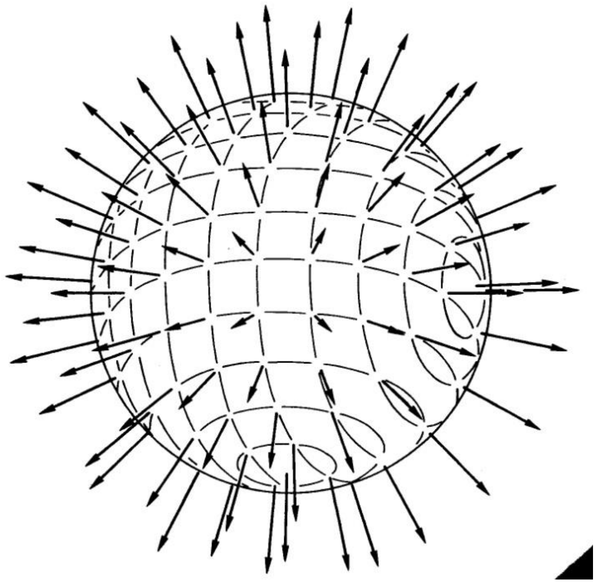


Fig.* The Hedgehog configuration: Arrows indicate the directions of the isovector field $\boldsymbol{\phi}$ at different points in coordinate space

Boundary conditions for the hedgehog solution

As we know, one of the fundamental requirement is tending to unit matrix at infinity $U(\boldsymbol{x}) \to I \quad at \quad |\boldsymbol{x}| \to \infty$. Therefore, we must have

$$F(\infty) = 2\pi \times \text{integer number},$$

which can be chosen as zero. Otherwise, one can change the definition $F(r) \to F(r) - F(\infty)$.

Therefore, we take

$$F(\infty) = 0 \tag{283}$$

We must also require that $F(r)$ is well-defined at origin. If we take $\sin F(0) \neq 0$, then U will tend to $\hat{\boldsymbol{x}}$ depending limit, which is not well-defined at $\boldsymbol{x = 0}$: the space origin $r = 0$ has to reflected on S^3 sphere in one point, therefore we must require that $\sin F(0) = 0$. Hence

$$F(0) = n\pi, \qquad n = 0, \pm 1, \pm 2, ... \tag{284}$$

One can calculate the winding number:

$$B = \frac{1}{24\pi^2} \varepsilon_{ijk} \int d^3x Tr\left\{\left(U^+\partial_i U\right)\left(U^+\partial_j U\right)\left(U^+\partial_k U\right)\right\} =$$

$$\frac{1}{2\pi^2} \int d^3x \frac{\sin^2 F}{r^2} \frac{dF}{dr} = \frac{2}{\pi} \int\limits_{F(0)}^{F(\infty)} \sin^2 F dF$$

Therefore

$$B = \frac{1}{\pi}\left[F(0) - F(\infty)\right] + \frac{1}{2\pi}\left[\sin 2F(\infty) - \sin 2F(0)\right] = n \tag{285}$$

In conclusion, we have shown that the fields with generalized spherical symmetry give for winding number arbitrary integer values.

Because the Skyrme ansatz commutes with the generator $\boldsymbol{K}$ of diagonal group $SO(3)_J \times SO(3)_I$, the configuration "hedgehog" is a scalar in the K-space $(K = 0)$.

Remember that the parity on $\boldsymbol{\phi}$ field is defined as

$$P\boldsymbol{\phi}(\boldsymbol{x}, t) P^{-1} = -\boldsymbol{\phi}(-\boldsymbol{x}, t)$$

therefore

$$PU(\boldsymbol{x}, t) P^{-1} = U^+(-\boldsymbol{x}, t) \tag{286}$$

This means that the Skyrme ansatz is invariant under parity, therefore the quantum numbers of skyrmion are $K^P = 0^+$ and it may be considered as a mixed state of positive parity of $J = I$ states.

Radial equation for the Skyrme profile function

If we substitute the Skyrme ansatz into the equation of motion (280) after some manipulations we obtain the following radial equation

$$F'' + \frac{2}{r}F' - \frac{\sin 2F}{r^2} - 8\frac{e^2}{f_\pi^2}\left[\frac{\sin 2F \sin^2 F}{r^4} - \frac{F'^2 \sin 2F}{r^2} - \frac{2F'' \sin^2 F}{r^2}\right] = 0 \qquad (287)$$

with boundary conditions

$$F(0) = \pi, \qquad F(\infty) = 0 \qquad (288)$$

There are numerical solutions of this radial equation. We can investigate here analytically some general properties of it.

Character of asymptotic: At the small distances $F(r)$ manifests a linear dependence on r

$$F(r) \simeq n\pi - \alpha r, \qquad r \to 0 \qquad \textbf{(289)}$$

Exercise: Show this.

At the large distances because of boundary condition (288) terms in parenthesis decay more rapidly and if we restrict ourselves by the expansion of sine up to the first order, there remains the equation

$$F'' + \frac{2}{r}F' - \frac{2F}{r^2} = 0$$

Decreasing solution of this equation is,

$$F \sim 1/r^2$$

Such a behavior is expected for the source of massless particles.

Let us now bring an expression for skyrmion energy, which can be easily derived by substitution of the Skyrme ansatz

$$M_{Sk} = 4\pi\int_0^\infty r^2 dr\left\{\frac{f_\pi^2}{2}\left[F'^2 + \frac{2\sin^2 F}{r^2}\right] + 4e^2\frac{\sin^2 F}{r^2}\left[2F'^2 + \frac{\sin^2 F}{r^2}\right]\right\} \qquad (290)$$

If we minimize this expression with respect to chiral angle $F(r)$ evidently we obtain the equation of motion. Therefore for calculation of minimum of mass, we can use here the equation of motion, which after using the virial theorem means:

$$E^{(4)} = E^{(2)}$$

Therefore

$$M_{Sk} = 4\pi f_\pi^2\int_0^\infty r^2 dr\left[F'^2 + \frac{2\sin^2 F}{r^2}\right] \qquad (291)$$

introducing dimensionless variable as follows

$$x = \frac{\sqrt{2} f_\pi}{e} r \tag{292}$$

the mass can be rewritten in the form

$$M_{Sk} = 4\pi\sqrt{2} e f_\pi \int_0^\infty x^2 dx \left[\left(\frac{dF}{dx} \right)^2 + \frac{2\sin^2 F}{x^2} \right] \tag{293}$$

It is interesting to note that after using this new dimensionless variable the equation of motion becomes

$$\left(x^2 + 2\sin^2 F\right) F'' + \frac{1}{2} x F' + F'^2 \sin 2F - \frac{1}{4}\sin 2F - \frac{\sin^2 F \sin 2F}{x^2} = 0 \tag{294}$$

This is a one dimensional ordinary, but non-linear equation. It does not contain any free parameters, so it can be solved by numerical methods. Numerical solution of this equation is shown in Fig. (25) , and is adopted from the article of G.S.Adkins et al. Nucl. Physics B228 (1983)552-556.

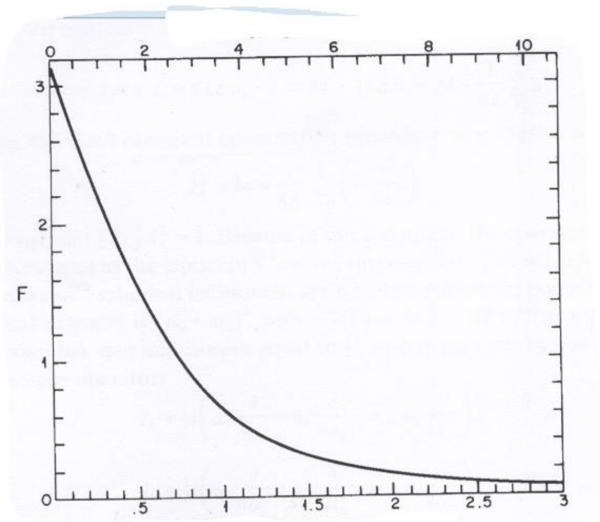


Fig. 25 **The numerical solution for F, Eq. (294). The radial distance is measured in fm in dimensionless variable** $\tilde{r} = e f_\pi r / 2$

Some Phenomenology of the Skyrme model

The first phenomenological calculations was performed by S.Adkins, C. Nappi and E. Witten (Nucl. Physics B228 (1983) 552-566. They found the soliton's profile function, depicted above and by its using they calculated many static properties of nucleon and delta, such as isoscalar charge and magnetic radii of nucleon, magnetic moments of proton and neutron and $N\Delta$ transition, axial constant g_A, also $g_{\pi NN}$ and checked that the Goldberger-Traiman relation is valid theoretically. In calculations they used nucleon and delta masses as inputs, while f_π and the Skyrme constant e as drive parameters.

Below we reproduce their calculations briefly.

So, if $U_0 \equiv \exp\{iF(r)\boldsymbol{\tau}\cdot\hat{\boldsymbol{x}}\}$ is the soliton solution, then $U = AU_0A^{-1}$, where A is an arbitrary constant $SU(2)$ matrix, is a finite energy solution as well. A solution with any given A is not an eigenstate of spin and isospin. We need to treat A as a quantum mechanical variable, as a collective coordinate. The simplest way to do this is to write the Lagrangian and all physical variables in terms of a time dependent $SU(2)$ matrix A. We substitute $U = A(t)U_0A^{-1}(t)$ in the Lagrangian. This procedure will allow us to write a Hamiltonian which we diagonalize. The Eigenstates with the proper spin and isospin will correspond to the nucleon and delta.

Substituting $U = A(t)U_0A^{-1}(t)$, after a lengthy calculation, we get for the Lagrangian

$$L = -M + \lambda Tr\{\partial_0 A \partial_0 A^{-1}\} \tag{295}$$

where M was defined above as skyrmion mass and $\lambda = \frac{2}{3}\pi\left(1/e^3F_\pi\right)\Lambda$ with

$$\Lambda = \int_0^\infty \tilde{r}^2 \sin^2 F\left[1+4\left(F'^2+\frac{\sin^2 F}{\tilde{r}^2}\right)\right]d\tilde{r}$$

Numerically, $\Lambda = 50.9$. The $SU(2)$ matrix A can be written as $A = a_0 + i\boldsymbol{a}\cdot\boldsymbol{\tau}$, $a_0^2 + \boldsymbol{a}^2 = 1$
In terms of these matrices

$$L = -M + 2\lambda\sum_{i=0}^{3}(a_i)^2 \tag{296}$$

Substituting the conjugate momenta $\pi_i = \frac{\partial L}{\partial \dot{a}_i} = 4\lambda \dot{a}_i$, we can now write the Hamiltonian

$$H = \pi_i \dot{a}_i - L = 4\lambda \dot{a}_i \dot{a}_i - L = M + \frac{1}{8\lambda}\sum_{i=0}^{3}\pi_i^2 \tag{297}$$

Performing the usual canonical quantization procedure $\pi_i = -\frac{\partial}{\partial \dot{a}_i}$, we get

$$H = M + \frac{1}{8\lambda}\sum_{i=0}^{3}\left(-\frac{\partial^2}{\partial \dot{a}_i^2}\right) \tag{298}$$

with the constraint $\sum_{i=0}^{3} a_i^2 = 1$, because of which the operator $\sum_{i=0}^{3}\partial^2/\partial a_i^2$ is to be interpreted as the Laplacian ∇^2 on the 3-sphere. The wave functions are traceless symmetric polynomials in the a_i. A typical example is $(a_0 + ia_1)^l$, with $-\nabla^2(a_0+ia_1)^l = l(l+2)(a_0+ia_1)^l$. Such a wave function has spin and isospin equal to $\frac{1}{2}l$, as one may see by considering the spin and isospin operators

$$
\begin{aligned}
I_k &= \frac{1}{2} i \left(a_0 \frac{\partial}{\partial a_k} - a_k \frac{\partial}{\partial a_0} - \varepsilon_{klm} a_l \frac{\partial}{\partial a_m} \right) \\
J_k &= \frac{1}{2} i \left(a_k \frac{\partial}{\partial a_0} - a_0 \frac{\partial}{\partial a_k} - \varepsilon_{klm} a_l \frac{\partial}{\partial a_m} \right)
\end{aligned}
\qquad (299)
$$

An important physical point must be addressed here. Since the non-linear sigma model field is $U = AU_0A^{-1}$ both A and $-A$ correspond to the same U. Naively, one might expect to insist that for wave function $\psi(A)$ there are two consistent ways to quantize the soliton as a boson or as a fermion depending on the sign of $\psi(A) = \pm\psi(-A)$. The choice of minus sign corresponds to quantizing it as fermion. In this case our wave function will be polynomials of odd degree in the a_i. So the nucleons of $I = J = 1/2$ corresponds to wave function linear in a_i. While the deltas of $I = J = 3/2$ correspond to cubic functions. Wave functions of fifth and higher orders correspond to highly excited states. the properly normalized wave functions for proton and neutron states of spin up or spin down along the z-axis, and same of the Δ wave functions are:

$$
\begin{aligned}
&\left| p \uparrow \right\rangle = \frac{1}{\pi}(a_1 + ia_2), \qquad \left| p \downarrow \right\rangle = \frac{1}{\pi}(a_1 - ia_2) \\
&\left| n \uparrow \right\rangle = \frac{1}{\pi}(a_0 + ia_3), \qquad \left| n \downarrow \right\rangle = (a_0 - ia_3) \\
&\left| \Delta^{++}, s_z = 3/2 \right\rangle = \frac{\sqrt{2}}{\pi}(a_1 + ia_2)^2, \\
&\left| \Delta^{+}, s_z = 1/2 \right\rangle = -\frac{\sqrt{2}}{\pi}(a_1 + ia_2)\left(1 - 3\left(a_0^2 + a_3^2\right)\right)
\end{aligned}
\qquad (30)
$$

Returning to the Hamiltonian, eigenvalues are $E = M + \frac{1}{8\lambda} l(l+2), \qquad l = 2J$ .So

$$
M_N = M + \frac{1}{2\lambda}\frac{3}{4} \qquad\qquad M_\Delta = M + \frac{1}{2\lambda}\frac{15}{4} \qquad (301)
$$

where $M = M_{Sk} = 36.5 f_\pi / 2e$ is evaluating numerically, moreover $\lambda = \frac{1}{3}\pi e^3 f_\pi 50.9$, as already was said. It was found that the best procedure in dealing with this model is to adjust e and f_π to fit the nucleon and delta masses. The results are $e = 5.45 \quad and \quad f_\pi = 64{,}5 MeV$.

Currents, charge radii and magnetic moments

In order to compute weak and electromagnetic couplings of baryons, it is needed first to evaluate the currents in terms of collective coordinates. The Noether current associated with the V-A transformation $\delta U = iQU$ is

$$J^{\mu}_{V-A}=\frac{1}{8}i\frac{f_{\pi}^{2}}{4}Tr\left[\partial^{\mu}UU^{+}Q\right]+\frac{i}{8e^{2}}Tr\left\{\left[\left(\partial_{\nu}U\right)U^{+},Q\right]\left[\left(\partial^{\mu}U\right)U^{+},\left(\partial^{\nu}U\right)U^{+}\right]\right\} \quad (302)$$

Anomalous baryon current is

$$B^{\mu}=\frac{\varepsilon^{\mu\nu\alpha\beta}}{24\pi^{2}}Tr\left[\left(U^{+}\partial_{\nu}U\right)\left(U^{+}\partial_{\alpha}U\right)\left(U^{+}\partial_{\beta}U\right)\right] \quad (303)$$

If one substitutes $U=A(t)U_{0}A^{-1}(t)$ we get rather complicated expressions for the vector and axial currents. The following angular integrals are adequate for these purposes:

$$\begin{aligned}&\int d\Omega V^{a,0}=\frac{1}{3}i4\pi\Lambda' Tr\left[\left(\partial_{0}A\right)A^{-1}\tau^{a}\right]\\&\int d\Omega \boldsymbol{q}\cdot\boldsymbol{x}V^{a,i}=\frac{1}{3}i\pi\Lambda' Tr\left(\boldsymbol{\tau}\cdot\boldsymbol{q}\tau^{i}A^{-1}\tau^{a}A\right)\\&\int d\Omega A^{a,i}=\frac{1}{3}\pi D' Tr\left(\tau^{i}A^{-1}\tau^{a}A\right)\end{aligned} \quad (304)$$

where

$$\Lambda'=\sin^{2}F\left[f_{\pi}^{2}/4+\frac{4}{e^{2}}\left(F'^{2}+\frac{\sin^{2}F}{r^{2}}\right)\right] \quad (305)$$

$$D'=\frac{f_{\pi}^{2}}{4}\left(F'+\frac{\sin 2F}{r}\right)+\frac{4}{e^{2}}\left(\frac{\sin 2F}{r}F'^{2}+2\frac{\sin^{2}F}{r^{2}}F'+\frac{\sin^{2}F\sin 2F}{r^{3}}\right)$$

In the computation of the above formulas terms quadratic in time derivatives are neglected, because they are of higher order in the semiclassical approximation.

It follows for the baryon current and charge density that

$$\begin{aligned}B^{0}&=-\frac{1}{2\pi^{2}}\frac{\sin^{2}F}{r^{2}}F'\\B^{i}&=i\frac{\varepsilon^{ijk}}{2\pi^{2}}\frac{\sin^{2}F}{r}F'\hat{x}_{k}Tr\left[\left(\partial_{0}A^{-1}\right)A\tau_{i}\right]\end{aligned} \quad (306)$$

The baryon charge per unit r is therefore

$$\rho_{B}(r)=4\pi r^{2}B^{0}(r)=-\frac{2}{\pi}\sin^{2}FF' \quad (307)$$

The isoscalar mean square radius is given by

$$\left\langle r^{2}\right\rangle_{I=0}=\int_{0}^{\infty}r^{2}\rho_{B}(r)\,dr=\frac{4.47\times 4}{e^{2}f_{\pi}^{2}}=4.47(0.28)^{2}\ fm^{2} \quad (308)$$

And we get $\left\langle r^{2}\right\rangle_{I=0}^{1/2}=0.59\,fm$, while the corresponding experimental value is 0.72fm.

Moreover, for the isovector charge density per unit r we obtain

$$\rho_{I=1}(r)=\frac{r^2\sin^2 F\left[\frac{f_\pi^2}{4}+\frac{4}{e^2}\left(F'^2+\frac{\sin^2 F}{r^2}\right)\right]}{\int_0^\infty r^2\sin^2 F\left(f_\pi^2/4+4/e^2\right)\left(F'^2+\frac{\sin^2 F}{r^2}\right)dr} \tag{309}$$

Now one can derive the proton and neutron charge distributions. They are plotted in Figure below

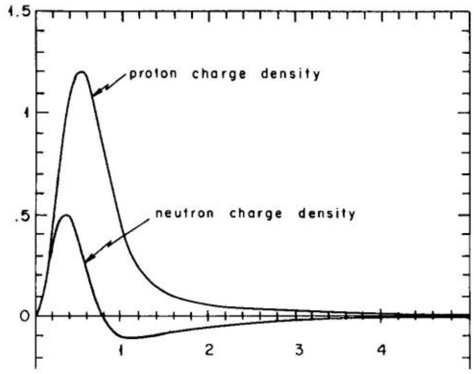


Fig.26 Charge densities are given as functions of the radial distance r and include a factor $4\pi r^2$

By analogy other static characteristics may be calculated. Results of calculations are summarized in the Table below

Table. Static quantities of nucleons in the Skyrme model

Quantity	Prediction	Experiment
M_N	input	939 MeV
M_Δ	input	1232 MeV
$2f_\pi$	129MeV	186 MeV
$\langle r^2\rangle_{I=0}^{1/2}$	0.59fm	0.72fm
$\langle r^2\rangle_{M,I=0}^{1/2}$	0.92fm	0.81 fm
μ_P	1.87	2.79
μ_n	-1.31	-1.91
$\lvert\mu_P/\mu_n\rvert$	1.43	1.46
g_A	0.61	1.23
$g_{\pi NN}$	8.9	13.5
$g_{\pi N\Delta}$	13.2	20.3
$\mu_{N\Delta}$	2.3	3.3

Agreement is satisfactory within the 30%.

In these calculation the pion mass was zero. In the following attempts were undertook to include into account the pion mass term as well, which breaks the chiral invariance. This term has a form

$$L_m = -\langle \bar{\psi}\psi \rangle Tr\left(mU + m^{+}U^{+}\right) \tag{310}$$

Where $\langle \bar{\psi}\psi \rangle$ is a quark condensate, and m - quarks' mass matrix. In this case pions acquire masses and the profile function asymptotic is changed $1/r^2 \to \exp(-m_\pi r)$. In addition to quantities from the Table above nucleon isotriplet charge radius $\langle r \rangle_{I=1} = \left(\langle r^2 \rangle_p - \langle r^2 \rangle_n\right)^{1/2}$ was calculated, which diverges logarithmically in the chiral limit. Moreover the so-called σ-term was obtained,

$$\sigma_{\pi NN} = \langle N | m\bar{\psi}\psi | N \rangle = \frac{f_\pi^2 m_\pi^2}{8} \int d^3x \left(2 - TrU\right) \tag{311}$$

The results look like

$\langle r \rangle_{m,I=1}$	1.04 fm	0.80 fm
$\sigma_{\pi NN}$	49 MeV	36 ± 20 MeV

Comparison shows that the numerical values of various quantities do not change essentially in result of accounting the pion mass term.

The significance of these results is in demonstration that the simplest quantitative realization of the idea "baryon as soliton" gives reasonable numbers.

Lecture 19

X. The Wess-Zumino term

The chiral anomaly of QCD and effective chiral model

We have already mention that in massless quarks limit QCD is characterized by extra $U(1)_A$ symmetry, which is explicitly broken by Adler-Bell-Jackiw anomaly. In the $SU(2)$ model this anomaly does not appear. Therefore consider the case of three massless quarks, $m_u = m_d = m_S = 0$ when the global symmetry is $U(3)_L \times U(3)_R$. We mentioned that this symmetry is broken spontaneously by the following way

$$U(3)_L \times U(3)_R / U(1)_A \equiv SU(3)_L \times SU(3)_R \times U(1)_V \to SU(3)_V \times U(1)_V \tag{312}$$

And therefore we have an octet of massless pseudo scalar particles. For non-linear realization of which we must take U matrix, defined on the $SU(3)$ manifold:

$$U(x,t) = \exp\left\{ i\lambda^a \frac{\pi^a(x,t)}{f_\pi} \right\}, \qquad a = 1,2,\dots,8 \tag{313}$$

Here λ^a are Gell-Mann $SU(3)$ matrices, normalized as $Tr\left(\lambda^a \lambda^b\right) = 2\delta^{ab}$. The explicit form of pseudo scalar octet is well known:

$$\lambda^a \pi^a(\boldsymbol{x},t) \equiv \sqrt{2} \begin{pmatrix} \frac{\pi^0}{\sqrt{2}} + \frac{\eta}{\sqrt{6}} & \pi^+ & K^+ \\ \pi^- & -\frac{\pi^0}{\sqrt{2}} + \frac{\eta}{\sqrt{6}} & K^0 \\ K^- & \tilde{K}^0 & -\frac{2}{\sqrt{6}}\eta \end{pmatrix} \tag{314}$$

Under Parity operation these fields transform as usual

$$P\pi^a(\boldsymbol{x},t)P^{-1} = -\pi^a(-\boldsymbol{x},t) \tag{315}$$

E. Witten mentioned that the sigma model Lagrangian, which is invariant under $U(3)_L \times U(3)_R$ global transformations and Parity transformation – as in QCD

$$PU(\boldsymbol{x},t)P^{-1} = U^+(-\boldsymbol{x},t) \tag{316}$$

manifests additional symmetry with respect of QCD, the symmetry under separate transformations

$$\begin{aligned} &1. \quad U(\boldsymbol{x},t) \to U(-\boldsymbol{x},t) \\ &2. \quad U(\boldsymbol{x},t) \to U^+(\boldsymbol{x},t) \end{aligned} \tag{317}$$

These invariances forbid the following processes

$$\begin{aligned} &K^+K^- \to \pi^+\pi^0\pi^- \\ &\pi^+\pi^- \to \pi^+\pi^-\eta \end{aligned}, \tag{318}$$

In other words –processes, in which the even number of pseudo scalar particles transform into to odd numbers and vice versa, are forbidden, the like processes are acceptable in QCD because we do not have here separate symmetries 1. and 2., but only their combined symmetry – parity. It is remarkable that in QCD these processes are allowed by anomaly and in the sigma model they are suppressed by kinematics, as we do not have an anomalous Ward identities.

E.Witten changed classical sigma model Lagrangian adding $U(3)_L \times U(3)_R$ invariant term, which brakes the additional symmetry, but maintains their combination. This program was realized as follows:

In order to break the (317-1) symmetry without explicit breaking of Lorentz-invariance we must introduce in the equation of motion totally antisymmetric tensor

$$\frac{f_\pi^2}{2}\partial^\mu L_\mu + \lambda \varepsilon^{\mu\nu\alpha\beta} L_\mu L_\nu L_\alpha L_\beta = 0 \tag{319}$$

This additional term contains the time derivative only linearly. Under $\boldsymbol{x} \to -\boldsymbol{x}$, Levy-Chivita tensor changes sign $\varepsilon^{\mu\nu\alpha\beta} \to -\varepsilon_{\mu\nu\alpha\beta}$, $\partial^{\mu} \to \partial_{\mu}$, $L_{\mu} \to L_{\mu}$, and we derive

$$\frac{f_{\pi}^{2}}{2}\partial_{\mu}L^{\mu} - \lambda\varepsilon_{\mu\nu\alpha\beta}L^{\mu}L^{\nu}L^{\alpha}L^{\beta} = 0 \tag{320}$$

When we perform the second transformation,

$$\begin{aligned} &\pi^{a} \to -\pi^{a}, \quad U(x) \to U^{+}(x), \quad L_{\mu} \to R_{\mu} = -UL_{\mu}U^{+} \\ &\partial^{\mu}L_{\mu} \to -\partial^{\mu}\left(UL_{\mu}U^{+}\right) = -U\partial^{\mu}L_{\mu}U^{+} \end{aligned} \tag{321}$$

the equation becomes

$$-\frac{f_{\pi}^{2}}{2}\partial^{\mu}L_{\mu} + \lambda\varepsilon^{\mu\nu\alpha\beta}L_{\mu}L_{\nu}L_{\alpha}L_{\beta} = 0 \tag{322}$$

Hence, the new term breaks both symmetries, but combined symmetry leaves inviolable.
Now we need the action functional for deriving equation of motion. This is far non-trivial task, because the explicit candidate

$$\varepsilon^{\mu\nu\alpha\beta}Tr\left(L_{\mu}L_{\nu}L_{\alpha}L_{\beta}\right)$$

is identically zero in (3+1)-dimension owing to cyclic property of trace.

Fortunately, this problem has a well-known solution. The analogical problem arises in monopole case. Consider a particle of mass m constrained to move on an ordinary two-dimensional sphere of radius one. The Lagrangian is

$$\mathcal{L} = \frac{1}{2}m\dot{\boldsymbol{r}}^{2} \tag{323}$$

and the equation of motion is

$$m\ddot{x}_{i} + mx_{i}\sum_{k}x_{k}^{2} = 0 \tag{324}$$

with the constraint $\sum_{k} x_{i}^{2} = 1$

This system respects the symmetries $t \leftrightarrow -t$ and separately $x_{i} \leftrightarrow -x_{i}$. If we want an equation that is only invariant under the combined operation $t \leftrightarrow -t$, $x_{i} \leftrightarrow -x_{i}$, the simplest choice is

$$m\ddot{x}_{i} + mx_{i}\sum_{k}x_{k}^{2} = \alpha\varepsilon_{ijk}x_{i}\dot{x}_{k}\ , \tag{325}$$

where α is a constant. To derive this equation from a Lagrangian is again troublesome. There is no obvious term whose variation equals the right-hand side (since $\varepsilon_{ijk}x_{i}x_{j}\dot{x}_{k} = 0$) . it seems that the situation is analogous to our problem.

But the solution of this problem exists: The right-hand side of the last equation can be understood as the Lorentz force for an electric charge interacting with a magnetic monopole located at the center of sphere. Introducing a vector potential $\boldsymbol{A}$ such that $\boldsymbol{B} = g\nabla \times \boldsymbol{A} = g\boldsymbol{x}/|x|^3$, the action of our problem will be

$$I = \int \left[\frac{1}{2} m\dot{x}_i^2 + \alpha A_i x_i \right] dt \tag{326}$$

where in case of monopole $\alpha = eg$, i.e. the product of electric and monopole charges. When $\alpha = 0$ or when monopole is absent, the action is invariant under separate transformations: $t \leftrightarrow -t$ and $x_i \leftrightarrow -x_i$. But when monopole is included, only combined transformation conserves.

However, this Lagrangian is problematical because A_i contains a Dirac string and certainly does not respect the symmetries of our problem. Indeed, at the one hand $\boldsymbol{B} \sim \boldsymbol{x}$,therefore $\nabla \cdot \boldsymbol{B} \neq 0$, but at the same time $\boldsymbol{B} = \nabla \times \boldsymbol{A}$ and $\nabla \cdot \boldsymbol{B} = 0$. So, $\boldsymbol{A}$ must be singular in R^3. Moreover, this part of action is not gauge invariant:
Under gauge transformations

$$\boldsymbol{A} \to \boldsymbol{A}^{\phi} = \boldsymbol{A} - \nabla \phi$$

the action changes as

$$I \to I^{\phi} - eg \int d\boldsymbol{r} \cdot \nabla \phi = I - eg \int d\phi \tag{327}$$

We are not able to restrict ϕ at the endpoints, because ϕ is arbitrary.

Noninvariance of the action has no importance in classical physics, because here the invariance of equation of motion is interesting for us, which is satisfied really. But in quantum mechanics the action takes place in transition matrix elements. It is obvious from the Feynman form of generating functional

$$Z[T] = Tr\left[e^{-iHT}\right] = \int_{\boldsymbol{r}(T)=\boldsymbol{r}(0)} d\left[\boldsymbol{r}(t)\right] e^{\frac{i}{\hbar} I[\boldsymbol{r}(t)]} \tag{328}$$

In quantum mechanics one can maintain the gauge invariance, If I would change be multiple of $\hbar$.

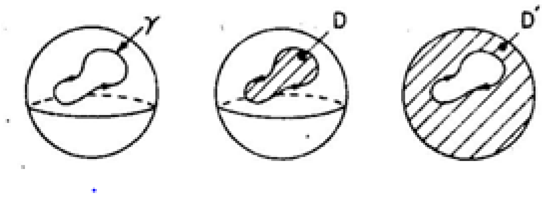
In I the troublesome term is

$$\exp\left(\frac{i}{\hbar} \alpha \int_{\gamma} A_i dx_i \right),$$

where the integration goes over the particle orbit γ , a closed orbit, if we discuss the simplest object $Tre^{-\beta H}$. By Gauss's law we can eliminate the vector potential from above integrand in favor of the magnetic field. In fact, the closed orbit γ in S^2 of Fig.27 (a) is the boundary of a disc D , and by Gauss's law we can write exponent in terms of the magnetic flux through D , if the integrand is not singular. But the monopole field $\boldsymbol{A}$ is necessarily singular owing to a Dirac string, the position of which depends on gauge choice. If we integrate in surface D , string becomes safe, if it threads the second surface D' . Therefore

$$\exp\left(\frac{i}{\hbar}\alpha\int_{\gamma} A_i dx^i\right) = \exp\left(\frac{i}{\hbar}\alpha\int_{D} F_{ij} d\Sigma^{ij}\right) \qquad (329)$$

The circle γ in S^2 is the boundary of a disc D (or more exactly, a mapping γ of a circle into S^2 can be extended to a mapping of a disc into S^2).



(a) (b) (c)

Fig. 27 A particle orbit γ on the two-sphere: (part (a)); bounds the disc D (part (b)) and D' (part (c))

The right-hand side of previous equality is manifestly well defined, unlike the left-hand side, which suffers from a Dirac string. We could try to use the right-hand side in a Feynman path integral. There is only one problem: D is not unique. The curve γ also bounds the disc D'. (Fig.27c). There is no consistent way to decide whether to choose D or D' (the curve γ could continuously be looped around the sphere or turned inside out). Working with D' we would get

$$\exp\left(\frac{i}{\hbar}\alpha\int_{\gamma} A_i dx^i\right) = -\exp\left(-\frac{i}{\hbar}\alpha\int_{D'} F_{ij} d\Sigma^{ij}\right), \qquad (330)$$

where a crucial minus sign on the right-hand side appears because γ bounds D in a right-hand sense, but bounds D' in a left-hand sense. If we are to introduce the right-hand side of previous forms in a Feynman integral we must require that they be equal. This is equivalent to

$$1 = \exp\left(\frac{i}{\hbar}\alpha\int_{D\cup D'} F_{ij}\Sigma^{ij}\right) \qquad (331)$$

Since $D + D'$ is the whole two sphere S^2, and $\int_{S^2} F_{ij} d\Sigma^{ij} = \int_{S^2} d\Omega = 4\pi$ the previous relation is obeyed if and only if α is an integer or half-integer.

$$\frac{4\pi eg}{\hbar} = 2\pi n, \qquad n \in Z \qquad \Rightarrow \qquad eg = \frac{n}{2}\hbar \qquad (332)$$

Evidently, we could derive uniqueness also in case when eg would be integer number. This is Dirac's quantization condition for the product of electric and magnetic charges. If we choose even n, we obtain the Schwinger monopole.

Interesting enough that this quantization has a *topological meaning*. Indeed, We saw that by gauge transformation the action changes as

$$\Delta I = eg\int d\phi \tag{333}$$

For closed contours $\frac{1}{2\pi}\int d\phi$ is a U (1) winding number, which is topological charge. So

$$\Delta I = 2\pi egB,$$

and the Feynman amplitude changes on value

$$\exp\left(\frac{i}{\hbar}2\pi egB\right) \tag{334}$$

And if it equals to 1, it means the requirement of uniqueness. B is quantized topologically

$$\left[\pi_1\left(U(1)\right)\sim Z\right]$$

The uniqueness of transition amplitude gives again the Dirac quantization rule.

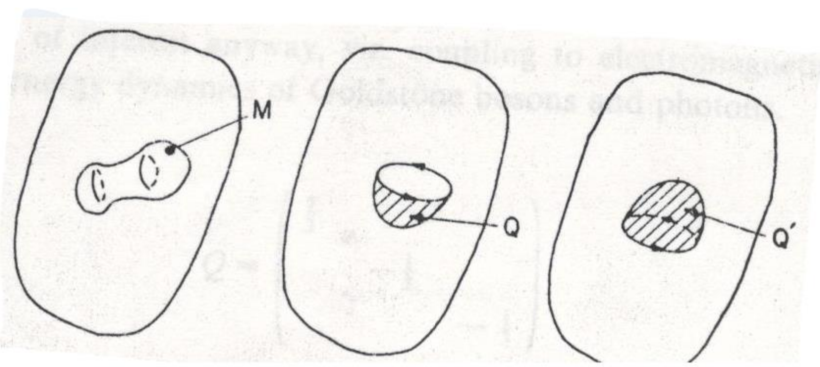
Therefore, *the uniqueness is provided by quantization of topological charge.*

Now let us return to our original problem. We imagine space-time to be a very large four-dimensional sphere M. A given field U is a mapping of M into the $SU(3)$ manifold (Fig.28a) Since $\pi_4\left(SU(3)\right)=0$, the four-sphere in $SU(3)$ defined by $U\left(\boldsymbol{x}\right)$ is the boundary of a five-dimensional disc Q.

By analogy with the previous problem, let us try to find some object that can be integrated over Q to define an action functional. On the $SU(3)$ manifold there is a unique fifth rank antisymmetric tensor ω_{ijklm} that is invariant under $SU(3)_L\times SU(3)_R$. Analogously of above consideration, we define

$$\Gamma = \int_Q \omega_{ijklm}d\Sigma^{ijklm}$$

As before, we hope to include $\exp\left(i\Gamma\right)$ in a Feynman path integral. Again, the problem is that Q is not unique. Our four-sphere M is also boundary of another five-disc Q' (Fig.28 a,b,c)



(a) (b) (c)

Fig.28 **Space-time, a four sphere, is mapped into the $SU(3)$ manifold. In part (a), space-time is symbolically**

denoted as a two sphere. In parts (b) and (c), space-time is reduced to a circle that bounds the discs Q and Q'. the SU(3) manifold is symbolized in these sketches by the interior of the oblong

If we let

$$\Gamma' = -\int_{Q'} \omega_{ijklm} d\Sigma^{ijklm}$$

(with again a minus sign because M bounds Q' with opposite orientation) then we must require $\exp(i\Gamma) = \exp(i\Gamma')$ or equivalently $\int_{Q+Q'} \omega_{ijklm} d\Sigma^{ijklm} = 2\pi \cdot \text{integer}$. Since $Q+Q'$ is closed five-dimensional sphere, our requirement is

$$\int_S \omega_{ijklm} d\Sigma^{ijklm} = 2\pi \cdot \text{ integer} \tag{335}$$

for any five-sphere S in the $SU(3)$ manifold.

We thus need the topological classification of mappings of the five-sphere into $SU(3)$. If we have such topology, there is a theorem that every sphere in $SU(3)$ is topologically a multiple of the basic five sphere S_0^5. The normalization may be chosen so that

$$\int_{S_0^5} \omega_{ijklm} d\Sigma^{ijklm} = 2\pi$$

and then we may work with the action

$$I = \frac{1}{4} f_\pi^2 \int d^4x Tr \partial_\mu U \partial^\mu U^{-1} + n\Gamma_0 \tag{336}$$

where n is an arbitrary integer and Γ_0 is, in fact, the Wess-Zumino Lagrangian. This Lagrangian was derived by Wess and Zumino restricting anomalies in the chiral $SU(3)_L \times SU(3)_R$ model.
As regards of Γ_0, it may be written in 5-dimensional sphere as follows: Let us introduce variables y^i, $(i = 1, 2, ..., 5)$, being coordinates for the disc Q. Then on Q

$$d\Sigma^{ijklm} \omega_{ijklm} = -\frac{i}{240\pi^2} d\Sigma^{ijklm} Tr\left[U^{-1} \frac{\partial U}{\partial y^i} U^{-1} \frac{\partial U}{\partial y^j} U^{-1} \frac{\partial U}{\partial y^k} U^{-1} \frac{\partial U}{\partial y^l} U^{-1} \frac{\partial U}{\partial y^m} \right]$$

Therefore, with the aid of this the modified action of sigma model is to be written as

$$S_\pm = -\frac{f_\pi^2}{4} \int d^4x Tr\left(L_\mu L^\mu\right) \pm \frac{(-i)\xi}{240\pi^2} \int_{Q,Q'} d^5x \varepsilon^{\mu\nu\alpha\beta\gamma} Tr\left[L_\mu L_\nu L_\alpha L_\beta L_\gamma\right]$$

Here it is implied that $U(y) = U(x, s)$ is a continuation of $U(x)$ to fifth dimension.

Variation of Wess-Zumino Lagrangian

Let us calculate the variation of Γ_0, which as was mentioned above, coincides with the Wess-Zumino Lagrangian. Consider first the right-handed transformations

$$\begin{aligned} &U \to U(1+ir), \qquad or \qquad \delta U = iUr \\ &U^+ \to (1-ir)U^+ \qquad\qquad \delta U^+ = -irU^+ \end{aligned} \tag{337}$$

As usual, r is thought as Hermitian matrix. From the definition of $L_\mu = U^+\partial_\mu U$ it follows that

$$\begin{aligned} \delta L_\mu &= \delta U^+\partial_\mu U + U^+\partial_\mu \delta U = -irU^+\partial_\mu U + U^+\partial_\mu (iUr) = \\ &= -irL_\mu + iL_\mu r + i\partial_\mu r \end{aligned}$$

The typical term derived by this variation is

$$Tr\left(-irL_\mu L_\nu L_\alpha L_\beta L_\gamma + iL_\mu r L_\nu L_\alpha L_\beta L_\gamma\right)\varepsilon^{\mu\nu\alpha\beta\gamma} \tag{338}$$

which inside of trace is zero after some rearrangements. Therefore the variation of the Wess-Zumino Lagrangian takes the final form as follows

$$\begin{aligned} \delta_R\Gamma &= -\frac{5i\xi}{240\pi^2}\int_Q d^5x\varepsilon^{\mu\nu\alpha\beta\gamma} Tr\left[i\partial_\mu r L_\nu L_\alpha L_\beta L_\gamma\right] = \\ &= -\frac{5i\xi}{240\pi^2}\int_Q d^5x\varepsilon^{\mu\nu\alpha\beta\gamma} Tr\left[i\partial_\mu r \partial_\nu U^+\partial_\alpha U\partial_\beta U^+\partial_\gamma U\right] \end{aligned} \tag{339}$$

Therefore transition of ∂_μ from $\partial_\mu r$ to other terms vanishes all getting terms. It means that

$$\delta_R\Gamma = \frac{\xi}{48\pi^2}\int d^5x\varepsilon^{\mu\nu\alpha\beta\gamma}\partial_\mu Tr\left[rL_\nu L_\alpha L_\beta L_\gamma\right]$$

which is a total divergence. therefore one can use the Stake’s theorem and carry integration to boundary of D_5 disk, which is a physical space-time, i.e.

$$\delta_R\Gamma = \frac{\xi}{48\pi^2}\int d^4x\varepsilon^{\nu\alpha\beta\gamma} Tr\left[rL_\nu L_\alpha L_\beta L_\gamma\right] \tag{340}$$

This leads us to the Witten’s equation of motion.

Interesting enough that that this equality may be rewritten as

$$\delta_R\Gamma = \frac{\xi}{48\pi^2}\int d^4x\varepsilon^{\nu\alpha\beta\gamma} Tr\left[r\partial_\nu U^+\partial_\alpha U\partial_\beta U^+\partial_\gamma U\right] \tag{341}$$

It follows, that

$$\begin{aligned} &\varepsilon^{\nu\alpha\beta\gamma}\partial_\nu U^+\partial_\alpha U\partial_\beta U^+\partial_\gamma U = -\partial_\mu\left[\varepsilon^{\mu\alpha\beta\gamma}\left(U^+\partial_\alpha U\right)\left(U^+\partial_\beta U\right)\left(U^+\partial_\gamma U\right)\right] = \\ &\qquad \partial_\mu\left[\varepsilon^{\mu\alpha\beta\gamma}\left(U^+\partial_\alpha U\right)\partial_\beta U^+\partial_\gamma U\right] \end{aligned}$$

It means that if we assume a sufficiently fast falling of fields at infinity and same behavior of $r(x,s)$ then the variation becomes

$$\delta_R\Gamma = \frac{\xi}{48\pi^2}\int d^4x\varepsilon^{\mu\alpha\beta\gamma} Tr\left[\partial_\mu r L_\alpha L_\beta L_\gamma\right]$$

In other words, the right currents of Noether corresponding to this part of variation, according to Gell-Mann-Levy equations take the form

$$J^a_{R\mu} = \frac{\xi}{48\pi^2}\varepsilon^{\mu\alpha\beta\gamma}Tr\left[\lambda^a L_\alpha L_\beta L_\gamma\right] \tag{341}$$

Analogously one can consider the left-hand transformations and get the left-hand currents

$$J^a_{L\mu} = -\frac{\xi}{48\pi^2}\varepsilon^{\mu\alpha\beta\gamma}Tr\left(\lambda^a R_\alpha R_\beta R_\gamma\right) \tag{342}$$

Therefore, we have shown that if we take the action in the form of Eq. (336), it follows the Witten equation of motion. But it remains to make clear, what is here the parameter ξ . It is easy exercise to show from uniqueness that

$$\xi = n$$

So, all ingredients are known in Witten's term.

Lecture 20

The physical consequences of this can be made more transparent as follows.

Using

$$U^{-1}\partial_i U = \frac{4i}{f_\pi}\partial_i A + O\left(A^2\right), \qquad A = \sum_a \lambda^a \pi^a \tag{343}$$

we derive

$$\omega_{ijklm}d\Sigma^{ijklm} = \frac{2\cdot 2^5}{15\pi^2 f_\pi^5}d\Sigma^{ijklm}Tr\left(\partial_i A\partial_j A\partial_k A\partial_l A\partial_m A\right) + O\left(A^6\right) =$$
$$= \frac{2\cdot 2^5}{15\pi^2 f_\pi^5}d\Sigma^{ijklm}\partial_i\left(TrA\partial_j A\partial_k A\partial_l A\partial_m A\right) + O\left(A^6\right)$$

So $\int_Q \omega_{ijklm}d\Sigma^{ijklm}$ is (to order A^5 and in fact also in higher orders) the integral of a total divergence which can be expressed by Stokes' theorem as an integral over the boundary of Q. By construction, this boundary is precisely space-time. We have then,

$$n\Gamma = n\frac{2\times 2^5}{15\pi^2 f_\pi^5}\int d^4x\varepsilon^{\mu\nu\alpha\beta}TrA\partial_\mu A\partial_\nu A\partial_\alpha A\partial_\beta A \text{ + higher order terms} \tag{344}$$

In a hypothetical world of massless kaons and pions, this effective Lagrangian rigorously describes the low-energy limit of $K^+K^- \to \pi^+\pi^0\pi^-$. We reach the remarkable conclusion that in any theory with $SU(3)\times SU(3)$ broken to diagonal $SU(3)$, the low –energy limit of the amplitude for this reaction must be an integer (in units, used above).

The magnitude of integer n in QCD.

Witten considered the coupling of Goldstone bosons with fermions, in order to get their dynamics at low energies. Let us take the Quarks' electric charge matrix

$$Q = \begin{pmatrix} 2/3 & & \\ & -1/3 & \\ & & -1/3 \end{pmatrix}$$

and consider it as a generator of $U_V(1)$ group. Wess-Zumino action Γ is invariant under the global rotations by charge operator

$$U \to U + i\varepsilon\left[Q,U\right] \qquad (345)$$

where ε is a constant. We wish to promote this to a local symmetry,

$$U \to U + i\varepsilon(x)\left[Q,U\right]. \qquad (346)$$

with $\varepsilon(x)$ an arbitrary function of x .It is necessary to introduce the photon field A_μ which transforms as $A_\mu \to A_\mu - (1/e)\partial_\mu \varepsilon$; e is the charge of proton.

Usually a global symmetry can straightforwardly be gauged by replacing derivatives by covariant ones, $\partial_\mu \to D_\mu = \partial_\mu + ieA_\mu$. In the case at hand, Γ is not given as the integral of a manifestly $SU(3)_L \times SU(3)_R$ invariant expression, so the standard road to gauging global symmetry is not available. One can still resort to the trial and error Noether's method, widely used in supergravity. Under a local charge rotation one finds

$$\Gamma \to \Gamma - \int d^4x \partial_\mu \varepsilon J^\mu . \qquad (347)$$

where

$$J^\mu = \frac{1}{48\pi^2}\varepsilon^{\mu\nu\alpha\beta} Tr\left[Q\left(U^{-1}\partial_\nu U\right)\left(U^{-1}\partial_\alpha U\right)\left(U^{-1}\partial_\beta U\right) + Q\left(\partial_\nu U U^{-1}\right)\left(\partial_\alpha U U^{-1}\right)\left(\partial_\beta U U^{-1}\right)\right]$$

(348)

is the extra term in the electromagnetic current required due to addition of Γ to the Lagrangian. The first step in the construction is to add the Noether coupling, $\Gamma \to \Gamma' = \Gamma - e\int d^4x A_\mu J^\mu(x)$. This expression is still not gauge invariant, because J^μ is not, but by trial and error one finds that by adding an extra term one can form a gauge invariant functional

$$\tilde{\Gamma}(U, A_\mu) = \Gamma(U) - e\int d^4x A_\mu J^\mu + \frac{ie^2}{24\pi^2}\int d^4x \varepsilon^{\mu\nu\alpha\beta}\left(\partial_\mu A_\nu\right)A_\alpha \times$$
$$\times Tr\left[Q^2\left(\partial_\beta U\right)U^{-1} + Q^2 U^{-1}\left(\partial_\beta U\right) + QUQU^{-1}\left(\partial_\beta U\right)U^{-1}\right]$$

The gauge invariant Lagrangian will then be

$$\mathcal{L} = \frac{1}{64} f_\pi^2 \int d^4x Tr D_\mu U D_\mu U^{-1} + n\tilde{\Gamma} \qquad (349)$$

What value of the integer n will reproduce QCD results?

Here we find a surprise. The last term in (348) has a piece that describes $\pi^0 \to \gamma\gamma$. Expanding U and integrating by parts, (348) has a piece

$$A = \frac{ne^2}{24\pi^2 f_\pi}\pi^0 \varepsilon^{\mu\nu\alpha\beta} F_{\mu\nu} F_{\alpha\beta} \tag{(350)}$$

This agrees with the result from QCD triangle diagrams if $n = N_C$, the number of colors. The Noether coupling $-eA_\mu J^\mu$ describes, among other things, a $\gamma\pi^+\pi^0\pi^-$ vertex

$$B = -\frac{2}{3} ie \frac{n}{\pi^2 F_\pi^2} \varepsilon^{\mu\nu\alpha\beta} A_\mu \partial_\nu \pi^+ \partial_\alpha \pi^- \partial_\beta \pi^0 \tag{351}$$

again agrees with calculations based on the VAAA anomaly of QCD if $n = N_C$. The effective action $N_C \tilde{\Gamma}$ (Wess-Zumino action) precisely describes all effects of QCD anomalies in low-energy processes with photons and Goldstone bosons.

Calculation of the Wess-Zumino term for $SU(2)$ group

One of the important property of the Wess-Zumino term has a linear dependence from the time derivative through $L_0 = U^+ \partial_0 U$ and contains integration by time, without loss of generality we may assume that t varies from 0 to 2π. One has to single out time explicitly. For that consider an $SU(3)$ -skyrmion in 3-th dimensional space-time with topology $S^3 \times S^1$. To leading order in $\hbar$ the vacuum to vacuum amplitude of a skyrmion at rest is given by

$$\left\langle S(T) \middle| S(0) \right\rangle \approx e^{-iTM/\hbar} \left(1 + O(\hbar)\right) \tag{352}$$

where M is the skyrmion mass. Now, rotate the skyrmion over 2π along S^1, infinitely slowly.

According to quantum mechanics the skyrmion wave function acquires a phase factor $\exp\left(-i2\pi J/\hbar\right)$, where J is the skyrmion spin. In other words,

$$\left\langle S(T) \middle| S(0) \right\rangle_{2\pi} \simeq e^{-iTM/\hbar} e^{i2\pi J/\hbar} \left(1 + O(\hbar)\right) \tag{353}$$

which determines J up to an integer. The phase factors here are given by the classical action of an adiabatically rotated skyrmion. To determine the latter, consider a $SU(2)$ hedgehog $U_H(\boldsymbol{x})$ embedded in $SU(3)$, i.e.

$$U(\boldsymbol{x}) = \begin{pmatrix} U_H(\boldsymbol{x}), & 0 \\ 0, & 1 \end{pmatrix}$$

Because of the hedgehog character of $U_H(\boldsymbol{x})$, rotations are equivalent to isorotations

$$U(\boldsymbol{x}, t) = e^{i\lambda_3 t/2} U(\boldsymbol{x}) e^{-i\lambda_3 t/2} = U\left(R_3(t)\boldsymbol{x}\right)$$

Explicitly $U(\boldsymbol{x}, t)$ reads

$$U(x,t)=\begin{pmatrix} e^{it/2} & 0 & 0 \\ 0 & e^{-it/2} & 0 \\ 0 & 0 & 1 \end{pmatrix} U(\boldsymbol{x}) \begin{pmatrix} e^{-it/2} & 0 & 0 \\ 0 & e^{it/2} & 0 \\ 0 & 0 & 1 \end{pmatrix} =$$

Multiplying both sides on unity $1=e^{-t/2}e^{t/2}$, we get

$$U(x,t)=\begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-it} & 0 \\ 0 & 0 & e^{-it/2} \end{pmatrix} U(\boldsymbol{x}) \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{it} & 0 \\ 0 & 0 & e^{it/2} \end{pmatrix} =$$

Multiplying again on unit matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{3it/2} \end{pmatrix} \times \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{3it/2} \end{pmatrix}^{-1} = 1$$

and using commutativity of appearing here matrix with $U(x)$, owing to its structure, we obtain

$$U(x,t)=\begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-it} & 0 \\ 0 & 0 & e^{it} \end{pmatrix} U(\boldsymbol{x}) \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{it} & 0 \\ 0 & 0 & e^{-it} \end{pmatrix}$$

Therefore the obtained matrix has the form

$$U(x,t)=\begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-it} & 0 \\ 0 & 0 & e^{it} \end{pmatrix} \begin{pmatrix} U_H & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{it} & 0 \\ 0 & 0 & e^{-it} \end{pmatrix} \equiv A(t)U(\boldsymbol{x})A^{-1}(t) \qquad (354)$$

A new $A(t)$ matrix is periodic and may be extended from a circle

$$\rho=1, \qquad 0\le t\le 2\pi$$

to a disk

$$S^1; \quad 0\le\rho\le 1$$

For example, very often the following extension is used

$$A(t,\rho)=\begin{pmatrix} 1 & 0 & 0 \\ 0 & \rho e^{it} & \sqrt{1-\rho^2} \\ 0 & -\sqrt{1-\rho^2} & \rho e^{-it} \end{pmatrix} \qquad (355)$$

One may substitute thus well-defined extension

$$U(\boldsymbol{x},t,\rho)=A(t,\rho)U(\boldsymbol{x})A^{-1}(t,\rho)$$

into the definition of Wess-Zumino action

$$S_{WZ} = -\frac{iN_C}{240\pi^2}\int d^5xTr\left[\varepsilon^{\mu\alpha\beta\gamma\sigma}\left(U^+\partial_\mu U\right)\left(U^+\partial_\alpha U\right)\left(U^+\partial_\beta U\right)\left(U^+\partial_\gamma U\right)\left(U^+\partial_\sigma U\right)\right]$$

First of all calculate the variation under the following transformation

$$U(x,t,\rho)\to U(x,t,\rho)\left(1+ir(x,t,\rho)\right)$$

and make use the steps, applying above, one easily derives

$$\delta_r S_{WZ} = \frac{N_C}{48\pi^2}\int d^4xTr\left[\partial_\mu r\varepsilon^{\mu\alpha\beta\gamma}\left(U^+\partial_\alpha U\right)\left(U^+\partial_\beta U\right)\left(U^+\partial_\gamma U\right)\right] \tag{356}$$

Remember now, that $A(t)$ is a unitary matrix from $SU(3)$, rotating spherically symmetric chiral soliton. From previous consideration we know that the corresponding currents have the form (341-342). Therefore the piece, following from Wess-Zumino term , could be

$$\Delta Q_R^a = \frac{N_C}{48\pi^2}\int d^3x\varepsilon_{ijk}Tr\left\{\lambda^a\left(U^+\partial_i U\right)\left(U^+\partial_j U\right)\left(U^+\partial_k U\right)\right\}$$

$$\Delta Q_L^a = -\frac{N_C}{48\pi^2}\int d^3x\varepsilon_{ijk}Tr\left\{\lambda^a\left(U^+\partial_i U\right)\left(U^+\partial_j U\right)\left(U^+\partial_k U\right)\right\}$$

Substituting a “rotated” ansatz and remembering the definition of baryonic charge, after simple manipulations we get

$$\Delta Q_R^a = \Delta Q_L^a = \frac{N_C B}{4}\int d^3xTr\left\{A^+\lambda^a A\begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}\right\} = \frac{N_C B}{4\sqrt{3}}Tr\left(A^+\lambda^a A\lambda^8\right) \tag{357}$$

$$\lambda^8 = \frac{1}{\sqrt{3}}\begin{pmatrix}1&0&0\\0&1&0\\0&0&-2\end{pmatrix}$$

It follows that the contribution to the axial charge $\left(\Delta Q_L - \Delta Q_R\right)$ from the Wess-Zumino term equal zero, but to the vector charge is

$$\Delta Q^V = \Delta Q_R^a + \Delta Q_L^a = \frac{N_C B}{2\sqrt{3}}\int d^3xTr\left[A^+(t)\lambda^a A(t)\lambda^8\right] \tag{358}$$

Now it is an easy task to find the Wess-Zumino term itself in case of considered rotated ansatz. Indeed, the vector charge according to the Noether theorem (or Gell-Mann-Levy equation) is a coefficient in front of $\dot{\alpha}^a(t)$ in the variation of the sought-for functional $S_{WZ}\left[A(t)\right]$ in case of vector transformation, depended only on time

$$A(t)\to\left[1-i\alpha^a(t)\lambda^a\right]A(t),\quad A^+(t)\to A^+(t)\left[1+i\alpha^a(t)\lambda^a\right],$$

The functional, possessing this property, has the form

$$S_{WZ} = \frac{N_C B}{2\sqrt{3}}\int dtTr\left(iA^+\dot{A}\lambda^8\right) \tag{359}$$

and if we substitute the explicit form of $A(t)$ into this general form, we find the needed result

$$S_{WZ} = N_C B\pi \tag{360}$$

Discussion: Soliton and QCD

Before going in advance let us summarize the principal results obtained earlier and look for left problems in the framework of QCD.

Till now we have suggested to develop an idea that baryons are nonlinear waves (solitons) as a phase of chiral condensate, which appears in QCD. Now we want transfer attention to the most principal problems, which are related to the quantum numbers of the chiral soliton. First of all let us remember some of the considered results.

Because pseudo scalar mesons are fluctuations (or chiral condensate phases) of unit matrices of $SU(2)\times SU(2)$ *or* $SU(3)\times SU(3)$, the most suitable parameterization for them was exponential one, introduced by unitary matrix

$$U(x) = \exp\left\{i\lambda^a\pi^a(x)/F_\pi\right\}, \qquad \lambda^a \text{ - Gell-Mann matrices}$$

It is well-known that in the chiral limit (massless quarks) QCD is invariant under chiral transformations of the quark fields

$$\psi_{Li} \to A_{ij}\psi_{Lj}. \qquad \psi_{Ri} \to B_{ij}\psi_{Rj} \tag{361}$$

where A *and* B are arbitrary 3×3 unitary matrices, $A^+A = B^+B = I$. At the same time $U(x)$ transforms as a composed meson field $U_{ij} \sim \psi_{Li}\bar{\psi}_{Rj}$ or under the above chiral transformations

$$U(x) \to AU(x)B^+, \qquad U^+ \to BU^+A^+$$

Comment it is easy to guess that A *and* B are above considered matrices L *and* R. Invariance of QCD with respect of these transformations requires that the action written in terms of $U(x)$ fields must be invariant under these global transformations.

The first term of effective chiral action (the non-linear sigma model) has the form

$$S_{ch} = \frac{F_\pi^2}{4}\int d^4x Tr\left(\partial_\mu U\partial^\mu U^+\right) =$$

$$= \int d^4x\left\{\frac{1}{2}\left(\partial_\mu\pi^a\right)^2 + \frac{1}{3F_\pi^2} f^{abc}f^{cde}\partial_\mu\pi^a\partial^\mu\pi^e\pi^b\pi^d + O\left(\pi^6\right)\right\}$$

Here f^{abc} are the structure constants of $SU(3)$ group.

The second row is derived after expansion of $U(x)$ matrix:

$$U(x) = 1 + i\frac{\pi^a\lambda^a}{F_\pi} - \frac{1}{2F_\pi^2}\pi^a\pi^b\lambda^a\lambda^b + \ldots$$

The following important step was the introduction of the Skyrme term, in order the topological soliton becomes dynamically stable.

As regards of skyrmion's quantum numbers, here very important is the Wess-Zumino-Witten term, considered above. Its presence follows in QCD taking into account the Axial anomaly: Accounting a local gauge (x-depended) chiral transformations and corresponding gauge transformations for external compensating fields L_μ and R_μ, which interact with the left and right currents of quarks, then the $U_A(1)$ symmetry breaks and the Adler anomalies appear. If we agree that the chiral soliton theory must be compatible with QCD, then exactly the same anomaly must appear in its effective Lagrangian. In this way we have recovered the Wess-Zumino term above.

We wrote the WZ term as an integral over 5-dimensional space

$$S_{WZ} = -\frac{iN_C}{240\pi^2}\int d^5x\varepsilon^{\alpha\beta\gamma\delta\sigma}Tr\left\{\left(U^+\partial_\alpha U\right)\left(U^+\partial_\beta U\right)\left(U^+\partial_\gamma U\right)\left(U^+\partial_\delta U\right)\left(U^+\partial_\sigma U\right)\right\}$$

To achieve the dynamical stability the consideration of higher order derivative terms is needed. There are many possibilities and unfortunately, there is no way to pick out from them. But if we confine ourselves to quadratic term in time derivative (note, that the WZ term is linear in time derivatives), we must confined ourselves by Skyrme term

$$S_{Sk} = N_C const\int d^4xTr\left[U^+\partial_\mu U, U^+\partial_\nu U\right]^2$$

Here a const is a dimensionless number.

There were models, in which the chiral soliton stability is provided by inclusion of vector fields, but these models are not well-interpret theoretically.

Let us consider results following from this action.

Currents and charges

The expressions of currents and charges in terms of $U(x)$ we have derived earlier. Let us now connect them to QCD.

On the quark language left and right currents are expressed as a half-sum or half- difference of vector and axial currents

$$J^a_{L\mu} = \bar{q}\gamma_\mu\frac{1+\gamma_5}{2}\lambda^a q; \qquad J^a_{R\mu} = \bar{q}\gamma_\mu\frac{1-\gamma_5}{2}\lambda^a q \tag{362}$$

These currents may be derived by the Noether's theorem. For this aim we consider the transformation of quark fields

$$q_L \to \exp\left\{-i\alpha^a(x)\lambda^a\right\}q_L, \qquad q_R \to \exp\left\{-i\beta^a(x)\lambda^a\right\}q_R$$

and make the parameters infinitesimal. Currents are derived as coefficients of derivatives of parameters (Gell-Mann-Levy) in action variation

$$\delta S = \int d^4x \left(\partial_\mu \alpha^a J^a_{L\mu} + \partial_\mu \beta J^a_{R\mu} \right) \tag{363}$$

Application of this method in chiral Lagrangian requires performing of infinitesimal transformation of $U(x)$ field. We derive:

$$J^a_{L\mu} = -i\frac{F_\pi}{2} Tr\lambda^a U\partial_\mu U^+ - \frac{N_C}{48\pi^2}\varepsilon^{\mu\alpha\beta\gamma} Tr\left[\lambda^a \left(U\partial_\alpha U^+\right)\left(U\partial_\beta U^+\right)\left(U\partial_\gamma U^+\right) + \ldots\right] \tag{364}$$

$$J^a_{R\mu} = -i\frac{F_\pi}{2} Tr\lambda^a U^+\partial_\mu U + \frac{N_C}{48\pi^2}\varepsilon^{\mu\alpha\beta\gamma} Tr\left[\lambda^a \left(U^+\partial_\alpha U\right)\left(U^+\partial_\beta U\right)\left(U^+\partial_\gamma U\right) + \ldots\right] \tag{365}$$

Here the first term follows from kinetic term, while the second one from S_{WZ} , ellipsis denote contributions from the other possible terms. Note that if we expand the first term in small π^a fields we derive $J^a_{L\mu} \approx -F_\pi \partial_\mu \pi^a$. This current interacting with a weak W boson causes the decay $\pi \to \mu\nu$. Really, the normalization of numerical value of $F_\pi \approx 96 MeV$ is derived from this decay.

Now the vector and axial vector currents are

$$J^a_{V\mu} = J^a_{L\mu} + J^a_{R\mu}, \qquad J^a_{A\mu} = J^a_{L\mu} - J^a_{R\mu}$$

Baryon charge

In Chromodynamics with N_C numbers of quarks the baryon is composed by N_C quarks and baryonic current in the quark language will be

$$J^B_\mu = J^B_{L\mu} + J^B_{R\mu} = \frac{1}{N_C}\bar{q}\gamma_\mu \frac{1+\gamma_5}{2} q + \frac{1}{N_C}\bar{q}\gamma_\mu \frac{1-\gamma_5}{2} q = \frac{1}{N_C}\bar{q}\gamma_\mu q$$

Therefore, quarks baryonic charge is $1/N_C$

Left and right baryonic currents can be formally derived according to the Noether's theorem, if we consider transformations;

$$q_L \to \exp\left(-i\frac{\alpha}{N_C}\right) q_{L,} \qquad q_R \to \exp\left(-i\frac{\beta}{N_C}\right) q_R$$

and we may derive the left- and right- handed baryonic currents in chiral theory from (364-365). Herewith the first term does not contribute, which means that baryonic current follows only from S_{WZ} .

$$J^\mu_B = \frac{1}{24\pi^2}\varepsilon^{\mu\alpha\beta\gamma} Tr\left(U^+\partial_\alpha U\right)\left(U^+\partial_\beta U\right)\left(U^+\partial_\gamma U\right)$$

and the baryonic charge is

$$B = \int d^4x J^0_B(x) = \frac{1}{24\pi^2}\int d^3x \varepsilon^{ijk} Tr\left(U^+\partial_i U\right)\left(U^+\partial_j U\right)\left(U^+\partial_k U\right)$$

This is, as we remember, normalizable expression of the topological mapping of homotopic group $\pi_3\left(SU(3)\right)=Z$.

Let us underline once again how the baryon charge appears by pseudo scalar mesons. This last expression is expected to be zero. Indeed, if we take the meson field to be small and expand $U(x)$ we derive

$$B=\frac{i}{24\pi^2}\int d^3x\varepsilon^{ijk}\partial_i\pi^a\partial_j\pi^b\partial_k\pi^c Tr\left(\lambda^a\lambda^b\lambda^c\right)=$$
$$=-\frac{1}{12\pi^2}\int d^3x\partial_i\left(\varepsilon^{ijk}f^{abc}\pi^a\partial_j\pi^b\partial_k\pi^c\right)$$

It is integral from total divergence (this could be shown even without expansion). If π fields do not have singularities and decrease rather quickly at infinity, integral is zero – it is in accordance with intuition – meson fields do not carry baryonic charges.

But this exercise says, in what cases can be derived non-zero baryonic charge - π fields must have singularity at some point. In this case it is not available to take π field small and we have to use the exact expression. Hence, appearing of baryonic charge must be a topological effect.

To be convince in that one can consider the spherical ansatz

$$\pi^a\left(x\right)=\begin{cases}\frac{1}{F_\pi}n^a f\left(r\right), & a=1,2,3; \quad n^a=\frac{x^a}{r}\\ 0, & a=4,5,6,7,8\end{cases} \tag{366}$$

Consequently,

$$U\left(x\right)=\begin{pmatrix}\cos f\left(r\right)+i\tau^a n^a\sin f\left(r\right), & 0\\ 0 & 1\end{pmatrix} \tag{367}$$

and the baryonic charge for such matrix will be

$$B=\frac{2}{\pi}\int_0^\infty dr\sin^2 f\left(r\right)\frac{df}{dr}=\frac{2}{\pi}\int_0^\infty dr\frac{d}{dr}\frac{1}{2}\left(f-\frac{\sin 2f}{2}\right)=\frac{1}{\pi}\left(f-\frac{\sin 2f}{2}\right)_0^\infty \tag{368}$$

Therefore, the baryonic charge is determined by the boundary conditions for profile function $f\left(r\right)$.

Let take vanishing condition at infinity, $f\left(\infty\right)=0$. As regards of origin, $f\left(0\right)$ it cannot be arbitrary, but $n\pi$, where n is an integer, which follows from the finiteness of soliton energy. Therefore, according to (368) the baryonic charge on “hedgehog” ansatz is an integer and is determined by boundary behavior at $r\to 0$ (!).

In particular, the value B=1 is achieved when $f\left(0\right)=-\pi$, antibaryon corresponds to $f\left(0\right)=+\pi$.

Why is a chiral soliton fermion?

The most astonishing is that the characteristics of the chiral soliton rather easily can be established in case of three quarks, when the matrix of pseudo scalar meson octet is a 3×3 unitary matrix of the $SU(3)$ group.

Let us follow to Witten's point of view. Consider $\exp(-iHT)$ evolution operator's matrix element between soliton states. It may be written in form of Feynman's functional integral

$$\langle soliton|e^{-iHT}|soliton\rangle = \int_{t=0:U(x)}^{t=T:U(x,T)} \mathscr{D}U(x,t)\,e^{i(S_{ch}+S_{WZ}+...)} \tag{369}$$

When $T\to\infty$ only lowest mass state remains on the left-hand side – baryon with mass M . Therefore in this limit the left-hand side is,

$$\langle soliton|baryon\rangle\langle baryon|e^{-iHT}|baryon\rangle\langle baryon|soliton\rangle$$

Now let us consider the similar matrix element between $U(x)$ soliton and rotated $A(t)U(x)A^{+}(t)$ soliton, which during the $T\to\infty$ time is rotated on 2π by $A(t)$ matrix around some axis. According to previous calculation we can write:

$$\langle rot.soliton|e^{-iHT}|soliton\rangle = \int_{t=0:U(x)}^{t=T:AU(x)A^{+}} \mathscr{D}U(x,t)\,e^{i(S_{ch}+S_{WZ}+...)} \approx \simeq e^{-iMT+iN_C B\pi} \tag{370}$$

Witten's argument is: S_{ch} is quadratic in time derivative, therefore in case of adiabatic rotation S_{ch} practically does not differ from the initial result, but S_{WZ} consists time derivative linearly, therefore this term makes difference between solitons in rest and rotated one.

In summary, we have obtained, that rotated soliton acquires the phase

$$e^{iN_C BT} = (-1)^{N_C B} \tag{371}$$

This means that when $N_C B$ is odd number, the chiral soliton is fermion, because its wave function changed the sign after rotation on 2π .

This result means also that for odd N_C every soliton is a fermion, while for even N_C - they are bosons.

Lecture 21

Justification of the Skyrme model in QCD

$1/N$ *expansion in quantum mechanics*

Above we saw that the solitons of the non-linear sigma model have precisely the quantum numbers of QCD baryons provided one includes the effects of the Wess-Zumino-Witten coupling. The basic fact that makes QCD a difficult theory to understand is that in QCD, as in atomic physics, the coupling constant can be scaled out of the problem. This fact is probably one of the most subtle discoveries in particle physics. It was properly appreciated only after the discovery of "asymptotic freedom" (the weakness of the QCD interaction at very high energies) and it played a great role in pinpointing quantum chromodynamics as the correct theory. In a QCD the probability amplitude for a quark to emit a gluon is proportional to the "color charge" g of the quark. This quantity is known as the QCD coupling constant. The renormalization group can be used to show that this constant does not have a characteristic value, rather, its value depends on the energy scale of processes one considers – or on the units in which one measures energy.

But the variable nature of the QCD coupling constant its numerical value can be absorbed by properly defining the overall scale of energies – Nothing depends on the coupling constant except this overall scale and therefore perturbation theory cannot answer such problems as explaining confinement or predicting the mass spectrum.

To solve these problems, we must somehow circumvent the apparent absence in QCD of a relevant expansion parameter. The $1/N$ expansion of QCD, originally suggested by t'Hooft , is an attempt to do this.

Because the reasoning behind the $1/N$ expansion is a little bit abstract, let us describe the $1/N$ expansion in some simple situation in atomic physics.

Let us consider the familiar Hamiltonian of the hydrogen atom:

$$H = \frac{p^2}{2m} - \frac{e^2}{r} \tag{372}$$

One might think that for small e^2 one could understand the hydrogen atom by treating the potential energy as a perturbation. This does not work because e^2 is not dimensionless and it does not make sense to say that e^2 is "large" or "small" – the value of e^2 just depends on the choice of units. After a rescaling

$r \to rt, \quad p \to p/t, \quad with \quad t = 1/me^2$, the Hamiltonian becomes

$$H = \left(me^4\right)\left(\frac{p^2}{2} - \frac{1}{r}\right) \tag{373}$$

and one sees that the "coupling constant" e^2 appears only in an overall factor me^4 multiplying the whole Hamiltonian, which merely helps set the overall scale of energies. Therefore, the hydrogen atom is a simple example of a problem without a free parameter, because it can be described by the reduced Hamiltonian

$$H = \frac{p^2}{2} - \frac{1}{r} \tag{374}$$

in which there is no free parameter. Likewise, other atoms and molecules can be described by the reduced Hamiltonian with e^2 scaled out.

Without a free parameter there is no perturbation expansion. What can one do?

To make progress, we must make an expansion of some kind. We must find a quantity one usually regards as given and fixed that we may treat as a free, variable parameter.

Instead of studding atomic physics in three dimensions, where it possesses a $SO(3)$ rotation symmetry, let us consider atomic physics in N dimensions, so that the symmetry is $O(N)$. We will see that atomic physics simplifies as $N \to \infty$ and that it can be solved for large N by expansion in $1/N$. For simplicity let us consider the s-states only. For these states wave function ψ is a function of r only, and the Schrodinger equation can be written as

$$\left[-\frac{1}{2m}\left(\frac{d^2}{dr^2}+\frac{N-1}{r}\frac{d}{dr}\right)-\frac{e^2}{r}\right]\psi = E\psi \qquad (375)$$

To eliminate terms with first derivative from the Hamiltonian , we make the transformation

$$\psi \to r^{(1-N)/2}\psi$$

and then defining (rescaling) $r = (N-1)^2 R$, then in terms of R the Hamiltonian becomes

$$H = \frac{1}{(N-1)^2}\left(-\frac{1}{2m(N-1)^2}\frac{d^2}{dR^2}+\frac{N-3}{8m(N-1)R^2}-\frac{e^2}{R}\right) \qquad (376)$$

Apart from the overall factor $1/(N-1)^2$, which only determines the overall scale of energy or time, the only N in this Hamiltonian is the N^2 that appears with the mass in the kinetic energy term. This is a Hamiltonian for a particle with an effective mass $M_{eff} = m(N-1)^2$, moving in an effective potential

$$V_{eff}(R) = \frac{\gamma}{8mR^2}-\frac{e^2}{R}; \qquad \gamma = \frac{N-3}{N-1} \qquad (377)$$

For large N the effective mass is very large, so that the particle simply sits in the bottom of the effective potential well – the quantum fluctuations are negligible. The ground state energy is simply the absolute minimum of V_{eff}, when $\gamma = 1$. In this case we find

$$E_0 = -2me^4(N-1)^{-2} \qquad (378)$$

Which in case $N = 3$ gives $E_0 = -me^4/2$, i.e. exact value.

To calculate the excitation spectrum, one may, for large N simply make a quadratic approximation to the effective potential near its minimum, because large effective mass ensures that the particle stays very close to the minimum of V_{eff}. The inharmonic terms in the expansion of V_{eff} around its minimum can be included as perturbations; this leads to an expansion in powers of $1/N$.

It was demonstrated in many papers that the quantitative accuracy can be obtained from $1/N$ expansion at $N = 3$.

$1/N$ *method in QCD*

To describe the $1/N$ expansion, it is necessary to describe QCD in somewhat more detail. In QCD there is actually not just one type of quark, but three types or “colors” of quarks. Here we will label the quark colors by number: q^i is the quark of type (color) i, where i may be equal 1, 2, or 3.
Apart from color, quarks can be distinguished by another property known as “flavor” (up, down, strange, charm, etc.). The quark flavor is very important in weak and electromagnetic interactions, but unimportant for strong interactions, so we will simply think of quarks as coming in three quark colors.

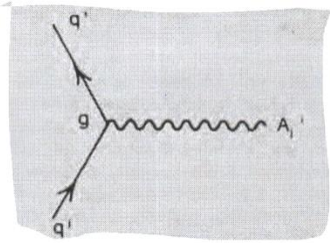


Fig.29 qqg-vertex

Each color of quarks participates equally in strong interactions. This is expressed mathematically by saying that there is a symmetry group, denoted by $SU(3)$, relating the three kinds (color) of quark; mathematically, the group has properties similar to the rotation group.

One of the basic processes in QCD is the process in which a quark emits a gluon (see Fig.29)
quark $\rightarrow$ quark +gluon).

The initial and final quarks have three color states each; the gluon field is a 3×3 matrix A_j^i in color space. Thus, the most general allowed process is that a quark of type q^i emits a gluon of type A_j^i and becomes a quark of type q^j. Because the 3×3 matrix for the gluon field is required to be a traceless matrix, it has not 9 but 8 independent components. This fact plays no role in the large - N expansion, and we may simply think of gluon field as a 3×3 matrix.

t’Hooft in 1974 suggested that one generalize from three quark colors to N colors. We still label quarks as q^i, but now i runs from 1 to N. The symmetry group becomes $SU(N)$ rather than $SU(3)$. The gluon field is now a $N\times N$ rather than 3×3 matrix.

The step is similar to the method in atomic physics, of generalizing from 3 to N dimensions and from $O(3)$ to $O(N)$ rotation symmetry.

t’Hooft showed that QCD also simplifies for large N. The basic reason that QCD simplifies for large number of colors is very simple. For large N the gluon field A_j^i has N^2 (actually N^2-1, but the difference is unimportant) components. For large N the Feynman diagrams contain large combinatory factors, arising from the large number of possible intermediate states. Only the diagrams with the largest possible combinatory factors need to be included when N is large. So only a subclass of diagrams is relevant, and the theory simplifies.

To see how this works in more detail, let us consider the lowest order contribution to the gluon vacuum polarization (Fig. 30). This is the lowest order quantum correction to the gluon propagator.

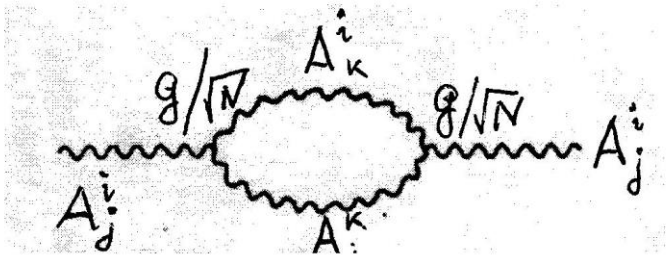


Fig. 30. Gluon propagator at one loop

It is not hard to see that for any choice of initial and final states, there are N possibilities for the intermediate state in the diagram. If the initial state gluon is of type A_j^i it can split into a pair of gluons, one of type A_k^i and one of type A_j^k, where k is arbitrary, and runs N possible values. Therefore, there are N possibilities for the intermediate state. We then must sum over all possible intermediate states. Therefore, the contribution of this diagram is proportional to a combinatorial factor of N from a sum over N different intermediate states. If QCD is to have a smooth limit for large N, this factor of N must somehow be cancelled. If the correction to the propagator of the gluon were to diverge for large N in proportion to N, all the other calculations would also give divergent results, and we could not construct a useful QCD for large numbers of colors.

There is only one way to cancel the combinatorial factor of N. We must remember that in our calculations for each of two vertices there is a factor of coupling constant. If we choose the coupling constant to be $g/\sqrt{N}$, where g is to be held fixed as $N \to \infty$, then the factors of N cancel out in this diagram, because $N\left(g/\sqrt{N}\right)^2 = g^2$, independent of N. So the choice of the coupling as $g/\sqrt{N}$ gives a smooth limit to the one-loop diagram (Fig. 27).

Moreover, this is the only choice of coupling constant that gives a smooth limit to this one-loop diagram. With any other choice the coupling constant factor will not cancel the combinatorial factor, and the large - N limit of QCD will not exist. But choosing the coupling constant as $g/\sqrt{N}$ is a fateful choice. Complicated diagrams will have factors of $g/\sqrt{N}$ at each vertex and so will vanish for large N unless, like the simple one-loop diagram, they have combinatory factors large enough to cancel the factors at the vertex.

It turns out that a certain class of diagrams, the so-called "planar" diagrams, have combinatorial factors large enough to just cancel the vertex factors. All other diagrams have smaller combinatorial factors and vanish for large N. The large N limit is therefore given by the sum of the planar diagrams.

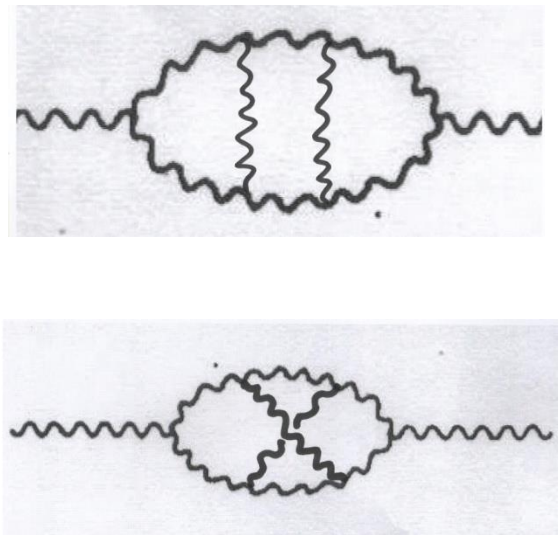


Fig. 31. **Planar and nonplanar diagrams at three loops**

For example, the three-loop diagrams in Fig.31 have factors of $\left(g/\sqrt{N}\right)^6$ from six vertices. The first turns out to have a combinatorial factor of N^3 from summing over the various intermediate states. Since $N^3\left(g/\sqrt{N}\right)^6 = g^6$, which is independent of N, the first diagram survives and has a smooth limit for large N. However, the second diagram in this figure has only a combinatorial factor of N, and vanishes for large N, as $1/N^2$.

The general class of diagrams that survives for large N was originally determined by t'Hooft. The diagrams that survive are the "planar" diagrams – i.e. which can be drown in the plane with no two lines crossing. The second diagram in figure is not a planar diagram, since two gluon lines cross at the center of the diagram, and it vanishes for large N.

The planar diagrams are a vast class of diagrams. Summing the planar diagrams is clearly very ambitious task. Since 1974, when t'Hooft first proposed the $1/N$ expansion, this problem has been the subject of some fairly intensive study. As we were convinced above on the basis of detail analysis of symmetry in QCD Witten and others concluded that the effective action of the sigma model necessarily consists the kinetic term and the Wess-Zumino term, caused by axial anomaly, as we have saw above

$$S = -\frac{F_\pi^2}{16}\int d^4x Tr\left(L_\mu L^\mu\right) - \frac{iN_C}{240\pi^2}\int_{D_5} d^5x\varepsilon^{\mu\nu\alpha\beta\gamma} Tr\left(L_\mu L_\nu L_\alpha L_\beta L_\gamma\right) \qquad (379)$$

The main difficulty in deriving terms of higher orders is in calculation of general path integral according to quark degrees of freedom. Using the special transformations for quark fields Adrianov and others succeeded to calculate the forth order terms in L_μ. They have a form

$$\begin{aligned} L_{eff}^{(4)}\left(U\right) = \frac{N_C}{384\pi^2} Tr\Big\{\left[L_\mu, L_\nu\right]\left[L^\mu, L^\nu\right]\Big\} - \\ -2\left(L_\mu L^\mu\right)^2 + 4\partial_\mu\partial^\mu U\partial_\mu\partial^\mu U^{-1} \end{aligned} \qquad (380)$$

We see that together with the Skyrme term (the first term here) there are also other forth order terms.

As we mentioned above, only anti-symmetric Skyrme term is of second order by time derivatives, which is important for quantization.

Lecture 22

Stabilization problem of classical soliton solutions in generalized Skyrme-like models

Very often for improving of phenomenological needs various generalizations of the original Skyrme model are considered. The main modifications consist in inclusion of unforeseen terms of forth or higher orders. In some papers such terms are derived from the model Lagrangians of Nambu-Johna-Lasinio or QCD-like in low energy region. But Hamiltonians obtained in this way are not always positive-definite ones, there arise the question of stability of soliton solutions of corresponding equations of motion. For completeness of exposition below we bring some confirmations from the theory of variational calculus.

Let us consider the functional

$$I[y]=\int_{x_1}^{x_2} dxf\left(x,y(x),y'(x)\right) \qquad (381)$$

Then the following theorem takes place. **Theorem**:

Suppose, that $f\in C_2(Q)$ *and the functional* $I[y]$ *reaches its weak local minimum on functions* $y\in D_1$. *Then the function* $y(x)$ *satisfies to the Legendre condition*

$$\tau(x)=f_{y'y'}\left(x,y(x),y'(x)\right)\geq 0, \qquad x\in[x_1,x_2] \qquad (382)$$

Here $C_2(Q)$ is a class of functions, having continuous second order derivatives in region Q, points of which are (x,y,y'), and D_1 is a class of functions having derivatives from the spline (or piecewise functions). It can be shown that the inverse is not valid always, so this theorem establishes only necessary, but not sufficient condition of weak local minimum.

There is also theorem about the sufficient condition:

Suppose, that the function $y(x)$ *fulfils the following conditions: a)* $f\in C_3(Q)$; b) $y\in C_2[x_1,x_2]$ is a stationary function of the functional I; c) $\tau(x)>0,\ x\in[x_1.x_2]$; d) $\Lambda_y(x)\notin 0,\ \ x\in[x_1,x_2]$. Then the functional $I[y]$ reaches a weak local minimum on functions $y(x)$. Here $\Lambda_y(x)$ denotes a solution of Jacobi differential equation

$$\left(p(x)-q(x)\right)\eta(x)-\tau'(x)\eta'(x)-\tau(x)\eta''(x)=0 \qquad (383)$$

where $p(x)=f_{yy}$, $q(x)=f_{yy'}$, $\tau(x)=f_{y'y'}$ with boundary conditions $\Lambda_y(x_1)=0,\quad \Lambda'_y(x_1)=1$.

Above exhibited statements are generalized to cases, when the functional $I[y]$ depends on higher derivatives. For example, the Legendre condition says:

If $f \in C_2(Q)$ *and the function* $y \in D_1$ causes a weak local minimum to functional

$$I[y] = \int_{x_1}^{x_2} dx f\left(x, y(x), y'(x), ..., y^{(n)}(x)\right), \qquad (384)$$

then the following inequality is valid

$$f_{y^{(n)}y^{(n)}}\left(x, y(x), y'(x), ..., y^{(n)}(x)\right) \geq 0, \qquad x \in [x_1, x_2] \qquad (385)$$

It is important in this statement that the inequality takes place for the second derivative of integral-ground expression with respect to the senior derivative.

Arming by these theorems let us return to our problem.

Structure of extra terms in modified Skyrme models

In traditional Skyrme model the Lagrangian of nonlinear sigma model

$$L_2 = \frac{f_\pi^2}{4} Tr\left[\partial_\mu U \partial^\mu U^+\right] \qquad (386)$$

is supplemented by the forth order term – square of antisymmetrized expression

$$L_4^{(1)} = \frac{1}{32e^2} Tr\left[\partial_\mu UU^+, \partial^\mu UU^+\right]^2 \qquad (387)$$

When $e^2 > 0$, the Skyrme term $L_4^{(1)}$ stabilizes static soliton, which has a form

$$U = U(\boldsymbol{r}) = \exp\left\{i\boldsymbol{\tau} \cdot \hat{\boldsymbol{r}} F(r)\right\} \qquad (388)$$

The profile function $F(r)$ satisfies to the Euler-Lagrange equation

$$F\left(\frac{\tilde{r}^2}{4} + 2\sin^2 F\right) + \frac{1}{2}\tilde{r}F' - \frac{1}{4}\sin 2F - F'^2 \sin 2F - \frac{1}{\tilde{r}^2}\sin 2F \sin^2 F = 0 \qquad (389)$$

Here $F' = \dfrac{dF}{d\tilde{r}}$ and $\tilde{r} = 2ef_\pi r$ is a dimensionless variable. In the vicinity of origin the Chiral angle $F(r)$ behaves as

$$F(r \to 0) \simeq n\pi - \alpha r\,. \qquad (390)$$

Moreover, it decreases at infinity as

$$F(r \to \infty) \sim r^{-2}. \qquad (390')$$

These boundary conditions lead to integer values for topological charge, which is defined as integral from zero component of conserved topological current

$$B_\mu = \frac{1}{24\pi^2}\varepsilon_{\mu\alpha\beta\delta} Tr\left[U^+\partial_\alpha U U^+\partial_\beta U U^+\partial_\delta U\right] \qquad (391)$$

Usually modification of the Skyrme model is made in two directions: By adding of 6th order term like

$$L_6 = -\frac{e_6^{\;2}}{4} B_\mu B^\mu \qquad (392)$$

and reflecting the sign of e^2, or including the symmetric forth order term

$$L_4^{(2)} = \frac{\gamma}{8e^2}\left[Tr\left(\partial_\mu U \partial^\mu U^+\right)\right]^2 \qquad (393)$$

Sometimes term with second derivative is also included

$$L_4^{(3)} = -\xi Tr\left(\partial^2 U \partial^2 U^+\right) \qquad (394)$$

It is easy to check that the inclusion of these terms do not have an influence on leading asymptotes (390-390') and hence, on boundary conditions in case of Skyrme ansatz (388).

The explicit forms of these terms look like:

$$L_2 = -\frac{1}{2f_\pi^2}\left(F'^2 + \frac{2}{r^2}\sin^2 F\right); \qquad L_4^{(1)} = -\left(\frac{1}{2e^2r^2}\right)\left(\frac{1}{r^2}\sin^2 F + 2F'^2\right)$$

$$L_4^{(1)} = -L_4^{(2)} = \frac{\gamma}{2e^2}\left(F'^2 + \frac{2}{r^2}\sin^2 F\right) \qquad (395)$$

$$L_4^{(3)} = -2\xi\left(F''^2 + F'^4 + \frac{4}{r}F''F'\right) + \frac{4}{r^2}F'^2 - \frac{2}{r^2}F'^2\sin 2F -$$

$$-\frac{4}{r^2}F'^2\sin^2 F - \frac{4}{r^3}\mathrm{F}'\sin 2F + \frac{4}{r^4}\sin^2 F$$

$$L_6 = -\frac{e_6^2}{16\pi^2 r^4}F'^2\sin^4 F$$

In these expressions $F' = dF/dr$.

Application of variational principles

(i) Let us consider first the model with symmetric term $L = L_2 + L_4^{(1)} + L_4^{(2)}$. Taking into account the explicit expressions, given above, mass functional can be written as

$$M = \left(\frac{\pi f_\pi}{e}\right)\int_0^\infty \tilde{r}^2 d\tilde{r}\left\{\left(F'^2 + \frac{2\sin^2 F}{\tilde{r}^2}2\sin^2 F\right) + \frac{4\sin^2 F}{\tilde{r}^2}\left(\frac{\sin^2 F}{\tilde{r}^2} + 2F'^2\right)\right\}$$

$$-\left(\frac{\pi f_\pi}{e}\right)\int_0^\infty \tilde{r}^2 d\tilde{r}4\gamma\left(F'^2+\frac{2\sin^2 F}{\tilde{r}^2}\right)^2 \qquad - \tag{396}$$

The solution of Euler-Lagrange equation $F(\tilde{r})$ must provide a local minimum to this expression, because the symmetric term gives a negative contribution. Moreover, local minimum must arise for γ below some critical value, γ_C.

It is evident from this expression (396) that

$$\tau(\tilde{r}) = f_{F'F'} = 2+\frac{16\sin^2 F}{\tilde{r}^2} - 16\gamma\left(3F'^2+\frac{2\sin^2 F}{\tilde{r}^2}\right) \tag{397}$$

Because the function $F(\tilde{r})$ is monotonically decreasing (soliton solution), then the principal contribution from γ term is expected at $r \to 0$, i.e. the sign of the last expression is determined by the behavior at the origin. Therefore considering this expression at $r \to 0$ and using boundary condition, we derive

$$f_{F'F'} = 2+16\tilde{\alpha}^2 - 80\gamma\tilde{\alpha}^2, \qquad \tilde{\alpha} = \alpha / 2ef_\pi \tag{398}$$

Requiring the Legendre condition, we obtain a restriction $\gamma \le 1/2+1/40\tilde{\alpha}^2$. Minimum of $\gamma \equiv \gamma_C$ is reached in case of equality.

We see that the value of γ_C depends on the slope of $F(r)$ at $r \to 0$, and is monotonically decreasing function of $\tilde{\alpha}^2$ (see Fig. 32).

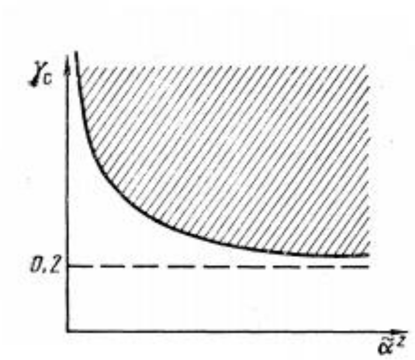


Fig.32 **Allowed region for parameters**

The range above the curve is forbidden. The lowest value of γ_C is $\gamma_C(\min) = 0.2$. Let us now investigate the Jacobi equation near $r \to 0$. After calculating of derivatives $p = f_{FF}$, $q = f_{F'F}$ in the limit $r \to 0$ and taking into account the strong Legendre condition $f_{F'F'} > 0$, the Jacobi equation reduces to

$$\eta'' + 2\eta'/\tilde{r} - 2\eta/\tilde{r}^2 = 0 \tag{399}$$

The general solution of which is $\eta(\tilde{r}) = c_1\tilde{r} + c_2\tilde{r}^{-2}$. In order to fulfill Jacobi conditions, we must take $c_1 = 1, \quad c_2 = 0$, in such case $\eta(0) = 0, \;\; \eta'(0) = 1$. Therefore, if $\gamma < \gamma_C(\min)$, then the mass functional should have a local minimum on soliton solutions. Interesting enough that phenomenological calculations performed in diversity of papers do not contradict to soliton stabilization in a such model with parameter $\gamma = 0.12$, which is lower than $\gamma_C(\min) = 0.2$

(ii) As regards of another model with the Lagrangian $L = L_2 + L_4^{(1)} + L_6$, mass functional takes the form

$$M = \left(\frac{\pi f_\pi}{e}\right)\int_0^\infty \tilde{r}^2 d\tilde{r}\left\{F'^2 + 2\frac{\sin^2 F}{\tilde{r}^2} - 4\sin^2 F\right\}\times \\ \times\left(\frac{\sin^2 F/\tilde{r}^2 + 2F'^2}{\tilde{r}^2} + \lambda F'^2 \sin^4 F/\tilde{r}^2\right), \quad (400)$$

where $\lambda = 2\pi^{-4}e^4 e_6^2 f_\pi^2$.

Now $\tau(\tilde{r}) = f_{F'F'} = 2 - 16\tilde{r}^{-2}\sin^2 F + 2\lambda\tilde{r}^{-4}\sin^4 F$ (401)

which for $r \to 0$ reduces to the expression $\tau(\tilde{r}) = 2 - 16\tilde{\alpha}^2 + 2\lambda\tilde{\alpha}^4$. According to Legendre condition

$$\lambda \ge \tilde{\alpha}^{-4}\left(8\tilde{\alpha}^2 - 1\right)$$

The critical value $\lambda_C\left(\tilde{\alpha}^2\right) = \tilde{\alpha}^{-4}\left(8\tilde{\alpha}^2 - 1\right)$ is a function of $\tilde{\alpha}^2$, approximate course of which is exhibited on the picture below

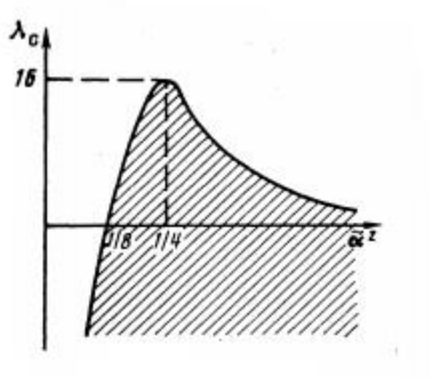


Fig. 33 **Forbidden range for shown parameters**

The range below the curve is forbidden. When $\lambda \ge \lambda_C(\max) = 16$, the Legendre condition is satisfied for arbitrary $\tilde{\alpha}^2$. Then it follows the restriction $e_6^2 \ge 8\pi^2 f_\pi^{-2} e^{-4}$.

In case of strict inequality above the Jacobi equation takes the same form, and consequently has a needed solution to guarantee a weak local minimum.

(iii) In case of inclusion the second derivative terms, we must be careful not to break the needed inequality for derivatives of functional under second derivative. When the sign of ξ is chosen correctly, then the corresponding model with all forth order terms is renormalizable at one loop level and the term with the second derivatives play the role of Pauli-Villars regulators.

Concluding remarks on the Skyrme model

As we have seen above the Skyrme model, as a non-linear chiral theory of pions, provides an approximate description of hadron physics in the low-energy limit. In this theory the nucleon emerges as a non-perturbative solution of the field equations, or more precisely as a topological soliton. This model is also seen as a prototype which might be applicable in various physical contexts where one could expect soliton solutions to occur (e.g. condensed matter physics (baby skyrmions), wrapped branes, ...). more recently, this model was applied for explanation for the newly discovered hadronic states.

We know that the original Skyrme Lagrangian is a naive extension of the non-linear sigma model consisting of a fourth-order field derivative term. This is nonetheless sufficient to stabilize the soliton against scale transformations and to reach at least a 30% accuracy with respect to physical observables. In order to incorporate effects due to higher-spin mesons and improve the fit on most observables a number of alternate Skyrme-like models which preserved the form of original Lagrangian while extending it to higher orders has been proposed and analyzed.

In the absence of exact analytical solutions, the only alternative to numerical treatment is the use of apply chosen analytical forms which provide sometimes a reasonable approximation but which may not reproduce the correct behavior in the limits $r \to 0, \infty$. For example, one can analyze the quantum behavior of the Skyrme model soliton based on a family of trial functions, taking into account breathing motion and spin-isospin rotations.

Conceptually different attempts of stabilizing the nonlinear soliton are those avoiding the Derick theorem by dropping the condition of stationarity (which is necessary condition of the theorem) and taking into account the quantum fluctuations of rotational and vibrational degrees of freedom.

We saw that the most attractive features in the Skyrme model are provided mainly by topological structure on the non-linear chiral sigma model, which unfortunately don’t gives stability of classical solitonic solutions. Besides, introducing new terms in this model brings more free parameters into the theory which is also undesirable.

Let us remember some ingredients of the nonlinear sigma model. Let write the Lagrangian in the form

$$L = -\frac{f_\pi^2}{4} Tr\left[\partial_\mu U^+ \partial^\mu U\right] \qquad (402)$$

where $f_\pi = 93 MeV$ is the pion decay constant. We can look for static solutions using Skyrme’s “hedgehog” ansatz

$$U = U_0 = \exp\left[i\boldsymbol{\tau}\cdot\boldsymbol{n}F\left(r\right)\right] \qquad \boldsymbol{n} \equiv \boldsymbol{r}/r \qquad (403)$$

The topological charge equals to

$$Q = \frac{i}{24\pi^2}\varepsilon^{ijk}\int_0^\infty d^3x Tr\left[U_0^+\partial_i U_0\right]\left[U_0^+\partial_j U_0\right]\left[U_0^+\partial_k U_0\right] =$$

$$=\left[F(0)-F(\infty)\right]/\pi \qquad (404)$$

So, if the profile function $F(r)$ satisfies the boundary condition $F(0)=n\pi$ (n being integer) and $F(\infty)=0$, then $Q=n$. Mass of hedgehog configuration is given by

$$M_{cl}=2\pi F_\pi^2\int_0^\infty drr^2\left[\left(\frac{dF}{dr}\right)^2+\frac{2}{r^2}\sin^2 F(r)\right] \qquad (405)$$

Corresponding Euler-Lagrange equation is the following:

$$r^2\frac{d^2F}{dr^2}+2r\frac{dF}{dr}=\sin\left(2F(r)\right) \qquad (406)$$

The Euler-Lagrange equation results from the extremum condition for mass functional (405). Usually, when the equation can not be solved analytically one tries numerical methods or look for minimum of mass functional using trial profile functions (and paying no attention to equation of motion). Obviously satisfactory description of static properties of baryons is always possible by choosing appropriate trial function. But as long as these trial profile functions have nothing to do with the solutions of Euler-Lagrange equation with relevant boundary conditions, the result can not be reliable.

Therefore, we'll say that given profile function describes soliton solution if it minimizes the mass and at the same time satisfies the Euler-Lagrange equation with relevant boundary conditions.

After these general remarks let us get back to our problem. Substituting scaled profile function $F(r/R)$ instead of $F(r)$ in the mass functional one finds

$$M_{cl}\left[F(r/R)\right]=RM_{cl}\left[F(r)\right] \qquad (407)$$

Clearly while reducing the characteristic scale of soliton, (R) the mass is reduced too – the soliton is collapsing. This is a consequence od the Derick theorem. To prevent the soliton from shrinking one may add to Lagrangian new terms with different behavior under scalling transformations.

It is interesting to find out how the instability of soliton manifests itself in the equation of motion. The Euler-Lagrange equation (406) is invariant under the change $F(r)\to F(r/R)$ and $F(r/R)$ will be solution of (406) provided $F(r)$ is. It means that R can be identified as one of the two constants of integration and the general solution of (406) must have the form

$$F=F(r;C_1,C_2)=F(r/C_2;C_1)=F(r/R;C_1) \qquad (408)$$

Exploiting the boundary condition $F(0)=\pi$ in Eq.(406) one finds the asymptotic behavior near origin

$$F(r\to 0)=\pi-r/R$$

and the equation can not determine the value of R (it could be fixed had we imposed the boundary condition upon the first derivative of $F(r)$. Choosing $F(0)=\pi$ we fix only one of the constants of integration, namely C_1, because the change of the other one, $C_2\left(or \quad R\right)$ does not affect the initial value $F\left(r=0\right)$:

$$F\left(r=0\right)=F\,(r\,/\,R)\big|_{r=0}$$

But this change can not affect the asymptotic value of $F\left(r\right)$ at the spatial infinity either:

$$F\left(r\right)_{r\to\infty}=F\left(r\,/\,R\right)_{r\to\infty}$$

Therefore the whole single-parameter family of solutions of Eq. (406) with the boundary condition $F\left(0\right)=\pi$ have the same asymptotic value at the spatial infinity. As long as Eq.(406) is invariant under shifting by π $\left(F\left(r\right)\to F\left(r\right)+\pi\right)$ the same may by said about solutions starting from the points $F\left(0\right)=n\pi$.

The actual value of $F\left(0\right)-F\left(\infty\right)$ was proved to be equal to $\pm\pi/2$. So the numerical solutions exhibiting in Figure below starting from the point $F\left(0\right)=\pi$ reach the same asymptotic value $\pi/2$ for different values of negative slope $\left(dF\,/\,dr\right)_{r=0}$ or R .

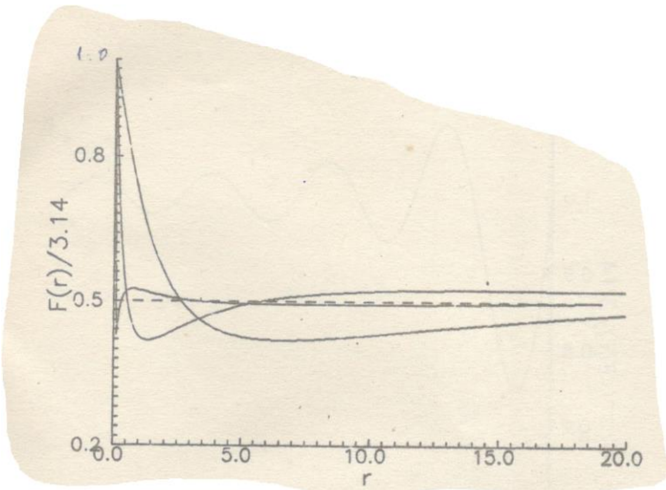


Fig.34. **Solutions of Eq.(406) with** $F\left(0\right)=\pi$ **for different slopes** $\left(dF\,/\,dr\right)_{r=0}$

Thus instability of soliton in NLM shows up in absence with proper boundary conditions.

Consider now the chiral symmetry breaking via standard pion mass term

$$L_{m_\pi}=\left(m_\pi^2F_\pi^2\,/\,4\right)Tr\left(U+U^{+}-2\right) \tag{409}$$

Corresponding expressions for soliton mass and equation of motion for profile function have the following form:

$$M = 2\pi F_\pi^2 \int_0^\infty drr^2 \left[\left(\frac{dF}{dr} \right)^2 + \frac{2}{r^2} \sin^2 F(r) \right] + 8\pi F_\pi^2 m_\pi^2 \int_0^\infty drr^2 \sin^2 (F/2) \qquad (410)$$

and

$$r^2 \frac{d^2F}{dr^2} + 2r \frac{dF}{dr} = \sin(2F) + m_\pi^2 r^2 \sin F \qquad (411)$$

First of all notice that the Eq.(411) excludes solutions with $F(r \to \infty) \to \pi/2$. Again the Derick theorem forbids the existence of stable soliton solutions because under scaling transformation the first and the second terms are multiplied by $R \quad and \quad R^3$, respectively and hence soliton collapses.

If we impose $F(0) = \pi$ then

$$F(r)_{r \to 0} = \pi - r/R$$

and

$$F(r)_{r\to\infty} = \begin{cases} \pi(\bmod 2\pi) + C_1 \sin(m_\pi r) + C_2 \cos(m_\pi r)/r \\ 0(\bmod 2\pi) + C_3 \exp(-m_\pi r)/r \end{cases} \qquad (412)$$

Now let us multiply Eq. (411) by dF/dr and integrate from 0 to some $r = r_0$ and perform partial integrations in the first and the last integrals. Using $F(0) = \pi$, we get

$$\frac{r_0^2}{2} \left[\frac{dF}{dr} \right]^2_{r=r_0} + \int_0^{r_0} \left[\frac{dF}{dr} \right]^2 rdr + 4m_\pi^2 \int_0^{r_0} drr \sin^2(F/2) =$$

$$= 2m_\pi^2 r_0^2 \sin^2 (F(r_0)/2) + \sin^2 (F(r_0)) \qquad (413)$$

The question we want to answer is whether the function $F(r)$ starting from $F(0) = \pi$ can approach zero at spatial infinity. If we make use of the second line of (411) then it can be easily shown that when $r_0 \to \infty$ the right-hand side of (413) tends to zero while the left-hand side is always positive. The only chance to make it zero too is to set $dF/dr = 0$ for all $r > 0$. So $F(r)$ has to have the form:

$$F(r) = \begin{cases} n, & r = 0 \\ 0, & 0 < r < \infty \end{cases} \qquad (414)$$

But for any continuous function with non-zero difference of boundary conditions the left-hand side of (413) is non-zero and positive.

Therefore no continuous solution of (411) with $F(0) = \pi$ can approach 0 (or 2π). The same is valid for any finite value of r_0 - it is impossible to satisfy (413) at the same requirements. Therefore, the only allowed asymptotic behavior of any continuous solution of (413) starting from $F(0) = \pi$ is

$$F(r)_{r\to\infty} = \pi + C_1 \sin(m_\pi r)/r + C_2 \cos(m_\pi r)/r$$

Topological charge of this configuration will equal to zero and it will have an infinite mass because of oscillations at infinity.

In the same manner one can prove that solutions with $F(0)=0$ behave at large distances like

$$F(r)_{r\to\infty} = \pm\pi + C_1 \sin(m_\pi r)/r + C_2 \cos(m_\pi r)/r$$

Again the solution is oscillating (resulting an infinite mass) but it has integer topological charge $Q=\pm1$.

So while smooth solutions of NLM equation of motion without chiral symmetry breaking have an infinite mass (caused by incorrect asymptotic value $F(\infty)=\pi/2$ and half-integer topological charge, those solutions with standard pion mass term have, again, infinite mass (caused by large-distance oscillations) but integer topological charge $Q=0,\pm1$. The numerical solutions are displayed in Figs.(31a,b)

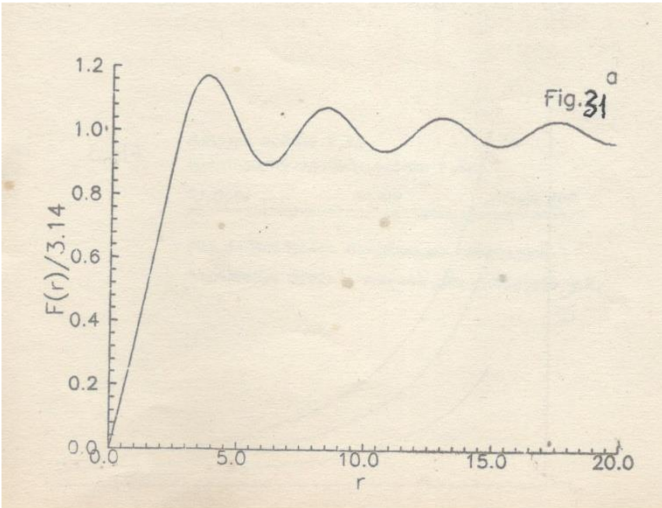


Fig.35a. **Solution with** $F(0)=\pi$

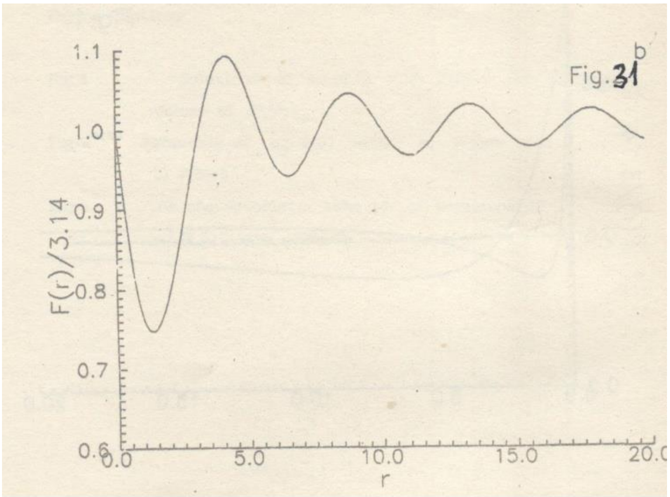


Fig. 35b**. Solution with** $F(0)=0$

Notice that the mass functional is positive and its minimal value $M=0$ is produced by profile function (414). But this function can not be obtained by solving the equation of motion. The reason is that solutions

must possess smooth derivatives up to the second order while the extremal function of variational problem may not belong to that class of functions.
One can consider another mass terms for breaking the chiral symmetry. The analogous study shows that there is no mass terms producing solutions with finite energy and unit topological charge.

Therefore, if smooth solutions of non-linear sigma model lead to infinite masses (because of incorrect asymptotics) and to the half-integer topological charges, moreover inclusion of pion mass terms made the charge integer, but the mass is nevertheless divergent (because of oscillations at infinity).

Quantization of various modes

1. Take into account the rotational degrees of freedom. After performing the standard semiclassical quantization one derives for mass the following expression

$$M_T\left[F(r)\right]=M_0\left[F(r)\right]+T(T+1)/2I\left[F(r)\right] \qquad (415)$$

Here T is the isospin and $I\left[F(r)\right]$ is the moment of inertia:

$$I\left[F(r)\right]=\frac{8\pi}{3}F_\pi^2\int_0^\infty drr^2\sin^2 F \qquad (416)$$

The Derick theorem is no longer valid because new ansatz is not static. The two terms in the right-hand side of (415) behave under transformation $F(r)\to F(r/R)$ like R and R^{-1} respectively. At the first sight the scaling behavior of mass functional ensures the existence of soliton sector with mass spectrum bounded from below. The Euler-Lagrange equation is

$$r^2d^2F/dr^2+2rdF/dr=\left(1-\mathrm{P}r^2\right)\sin(2F) \qquad (417)$$

where

$$P\equiv P\left[F(r)\right]=T(T+1)/3I^2\left[F(r)\right] \qquad (418)$$

These equations are no longer invariant under the change $F(r)\to F(r/R)$. A lot of work has been done in order to derive the masses and other static properties of baryons using (415) and some trial profile functions. It has been stressed that the criterion for choosing profile function shall not be that of satisfactory description of experimental data but first of all the stability of soliton solutions satisfying the equation of motion.

Let us look for the solutions with asymptotic value $F(\infty)=0\ (\mathrm{mod}\,2\pi)$, We obtain immediately (for constant $P\left[F\right]$)

$$F(r)_{r\to\infty}=\frac{C_1}{r}\cos\left(\sqrt{2\mathrm{P}}r\right)+\frac{C_2}{r}\sin\left(\sqrt{2\mathrm{P}}r\right) \qquad (419)$$

and in so far as $P>0$ the solution is oscillating. The oscillations are damped only by a factor $1/r$ and so it is impossible to find self-consistent solutions - it is clear that the moment of inertia diverges and consequently $P\left[F\right]$ goes to zero. But for $P\left[F\right]=0$ Eq.(417) reduces to that of non-linear sigma model with solutions asymptotically approaching $\pi/2$. They have an infinite mass, half-integer topological charge, infinite moment of inertia and it is meaningless to speak about their stability.

Addition of the pion mass term transforms the equation of motion into:

$$r^2 d^2F/dr^2 + 2rdF/dr = \left(1 - Pr^2\right)\sin\left(2F\right) + m_\pi^2 r^2 \sin\left(F\right) \qquad (420)$$

The possible large-distance behavior for self-consistent solutions all are oscillating except followings

$$F\left(r\right)_{r\to\infty} = \begin{cases} osc. \ \text{solutions} \\ 0\left(\text{mod}\, 2\pi\right) + C_1 \exp\left(-r\sqrt{m_\pi^2 - 2P}\right)/r \\ 0\left(\text{mod}\, 2\pi\right) + C_2/r^2 \end{cases} \qquad (421)$$

Arguments using in the previous case give us that there are no solutions with $P \neq 0$ and finite $F\left(\infty\right)$.

It can be shown that in spite of absence of relevant solutions of Eq. (420) the functional of mass has non-trivial minimum only after including the pion mass term. In particular, it is obtained that in this case there appears the inequality

$$M_T \geq m_\pi \sqrt{3T\left(T+1\right)/2} \qquad (422)$$

If we take the profile function like

$$F\left(r\right)_{r\to\infty} = \frac{C}{Cr + r^2}$$

then by choosing sufficiently large C and small R to get as near to

$$M = M_{\min} = m_\pi \sqrt{3T\left(T+1\right)/2}$$

as we like.

Therefore, incorporating the pion mass term stabilizes rotating soliton in the sense of ensuring the nonzero value of mass functional but the configuration minimizing mass is pathological in the sense that it does not obey the equation of motion and doesn't suit for description baryons (for instance, the average square radius of nucleon is zero)

Quantization of vibrational (breather) mode

Semiclassical quantization of vibrational mode is carried out in the same manner as that of rotational one. But in this case there is no symmetry associated with radial scaling transformations – there is no zero mode.

The idea comes from analogy with the particle in potential well with the minimum $V\left(x=0\right)=0$. While the classical particle will have zero energy and zero coordinate $x=0$, the quantization will give rise to nonzero expectation value $\left\langle x^2 \right\rangle \neq 0$ and nonzero ground state energy. Hence if we consider the size of the soliton as the dynamic variable (i.e. introduce time-dependence of R in $F\left(r/R\right)$ and substitute $F = F\left(r/R\left(t\right)\right)$ into non-linear lagrangian then the standard quantization enables us to derive the Schrodinger equation for the energy spectrum of soliton.

As a result of variety of investigations one can underline that the quantization of breathing mode in chiral invariant model can not lead to stable soliton solutions.

Various modifications were also considered. One of the attractive idea was to cut off the short distances (in nonrenormalizable theory like considered one it doesn't seem very unnatural). The mass functional now depends on the cutoff parameter ε :

$$M_0 = 2\pi F_\pi^2 \int_\varepsilon^\infty drr^2 \left[\left(dF / dr \right)^2 + \left(2 / r^2 \right) \sin^2 F\left(r\right) \right] \qquad (423)$$

and $F\left(r\right)$ is subject to boundary conditions: $F\left(\varepsilon\right) = \pi, \quad F\left(\infty\right) = 0$. It was shown that now the model has stable soliton solutions. We can remember previous analysis – any solution starting from $F\left(0\right) = n\pi$ approaches asymptotically $\pm\pi / 2 + n\pi$ (see, Fig. 36). But if we start from the other end – try to find where the solution with $F\left(\infty\right) = 0$ leads to – we'll find that it goes to infinity. In Fig. (36) we have shown the one parameter family of solutions with $F\left(\infty\right) = 0$.

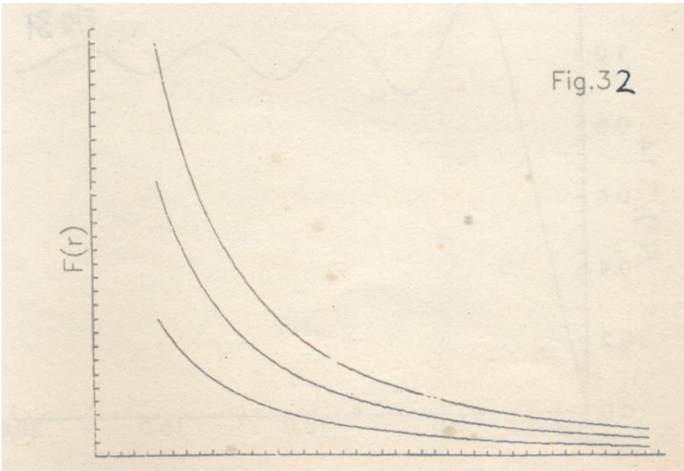


Fig. 36 **Characteristic behavior of solutions of Eq. (411) with** $F\left(\infty\right) = 0$

All of these solutions are connected by familiar transformation $F\left(r\right) \to F\left(r / R\right)$. Note that this transformation changes the boundary value at nonzero and finite r : $F\left(r = \varepsilon\right) \neq F\left(r / R = \varepsilon / R\right)$. So choosing R we can satisfy both the equation and boundary conditions simultaneously. The corresponding soliton solution will have minimal but nonzero energy. Thus introducing a cutoff parameter stabilizes soliton but brings in undesirable arbitrariness which can be avoided by quantization of ε. however now the soliton becomes not stable again.

Look now how can standard pion mass term affect the spectrum of Schrodinger equation. In this case we are faced to the following equation

$$\left[-\frac{d^2}{dZ^2} + \left(\frac{b^3}{4a} \right)^{1/3} Z^{2/3} + \left(3/4 + 2T\left(T+1\right) \frac{a}{I} \right) \frac{1}{Z^2} + \frac{Cm_\pi^2}{4F_\pi^2} \frac{Z}{a} \right] \phi = \frac{E}{F_\pi} \phi \qquad (424)$$

where

$$a = \frac{8\pi}{9}\int_0^\infty dyy^4\left(\frac{dF}{dy}\right)^2$$

$$b = 2\pi\int_0^\infty dyy^2\left[\left(\frac{dF}{dy}\right)^2 + \frac{2}{y^2}\sin^2 F(y)\right]$$

$$I = \frac{8\pi}{3}\int_0^\infty dyy^2 \sin^2 F(y)$$

$$c = 8\pi\int_0^\infty dyy^2 \sin^2\left(F(y)/2\right)$$

Here $y = F_\pi r$ - dimensionless variable, $Z = X\sqrt{4a}$, $X = R^{2/3}$ and the wave function $\phi = \psi X^{3/2}$.

We have a one-dimensional problem with effective potential

$$W_{eff} = \alpha Z^{2/3} + \beta / Z^2 + \gamma Z^2 \tag{424}$$

Here

$$\alpha \equiv \left(\frac{b^3}{4a}\right)^{1/3}; \qquad \beta \equiv \frac{3}{4} + 2T(T+1)a/I, \quad \gamma \equiv m_\pi c/\left(4aF_\pi^2\right)$$

It is evident that the effective potential (424) will have nontrivial minimum if at least one of the two parameters, α or γ is nonzero. The scaling property of c can be deduced from the definition. It follows

$$c\left[F(r/R)\right] = R^3 c\left[F(r)\right]$$

and recalling the scaling property of a, we can conclude that the new γ term is invariant too. Now the profile function with large distance behavior which produced the vanishing of γ -term can do no harm, c like a turns out to behave like ε^{-1} and $\gamma \sim c/a$ is independent of ε. So there is a hope that the new term ensures stability.

In order to convince in that let us suppose that for some profile function coefficient α becomes zero. Then remaining Schrodinger equation is solved explicitly and gives

$$\begin{aligned} E_n &= \frac{m_\pi}{2}\left(\frac{c}{a}\right)^{1/2}\left(4n+2+\sqrt{4+\frac{8T(T+1)a}{I}}\right) = \\ &= \frac{m_\pi}{2}\left((4n+2)\left(\frac{c}{a}\right)^{1/2} + \sqrt{\frac{4c}{a}+\frac{8T(T+1)c}{I}}\right); \qquad n = 0,1,2,\ldots \end{aligned} \tag{425}$$

When $\varepsilon = 0$ the coefficient α is zero and we have an analytic solution

$$E_n = (m_\pi/2)(c/a)^{1/2}\left[4n+2+\sqrt{4+8T(T+1)a/I}\right]$$

When $F(r)_{r\to\infty} = const.\times r^{-3/2}$, then $c/a = 1 \quad and \quad a/I = 3/4$ and

$$E_n = (m_\pi/2)\left[4n+2+\sqrt{4+6T(T+1)}\right] \tag{426}$$

The ground state energy for $T = 1/2$ equals to

$$E_{0,T=1/2} \sim 2.34 m_\pi$$

Conclusions.

The quantization of only rotational modes with pion mass term leads to soliton solutions with ground state energy bounded from below. But the profile functions in that case turn out to be pathological. Stability of the soliton can be achieved by quantizing vibrational mode with massive pions. Quantization of both modes together doesn't affect this statement. The only question that remains is whether the profile functions minimizing energy are smooth or pathological. Having no explicit expression for energy one can not obtain an equation for profile function and look for its self-consistent solutions.

The other side of the problem is that the numerical value of soliton mass (426) is much less then the nucleon mass. So it will not suit for description of static properties of baryons. But, *the fact of existence of energy spectrum bounded from below is important.* Besides the profile functions usually used give an estimate for nucleon mass exceeding the actual value. So as long as the model with chiral symmetry breaking produces much lower values of soliton mass it gives better opportunities to find profile functions corresponding to experimental value of nucleon mass.

Vanishing energy follows from (425) not only in case when $\frac{c}{a} = 0$, but also if $\frac{c}{I} = 0$. But looking on explicit expressions we can see that $\frac{c}{I} \geq \frac{3}{4}$. Indeed,

$$\frac{c}{I} = 3\frac{\int dy y^2 \sin^2 \frac{F(y)}{2}}{\int dy y^2 \sin^2 F(y)} = \frac{3}{4}\frac{\int dy y^2 \sin^2 \frac{F(y)}{2}}{\int dy y^2 \sin^2 \frac{F(y)}{2}\left(1 - \sin^2 \frac{F(y)}{2}\right)} \geq \frac{3}{4}$$

Therefore, in this case

$$E_0 \geq m_\pi \sqrt{\frac{3T(T+1)}{2}} \qquad (426)$$

We see that the minimal energy is provided owing to the rotating mode.

It is evident, however that for description of static properties of baryons, profile function, that minimizes of mass, is not suitable. On the other hand, we are able to choose the profile function in such a way that to get the better agreement with experimental data. Let us remember that the Skyrme model gives higher values for baryon masses and lower values of average square radii (if for input parameter pion decay constant is taken) In considered model these inconsistencies can be improved. Below the calculated values for baryonic masses are given by using the following decaying profile functions at spatial infinity

$$F_1(r) = \frac{2\pi}{1 + r + \exp(0.15r)},$$

$$F_2(r) = \frac{\pi}{1 + r + 0.25r^2}$$

Obtained results are summarized in the Table below

Mass (in GeV)	M_N	M'_N	M''_N	M_Δ	M'_δ
experiment	0.94	1.44	1.71	1.23	1.60
F_1	*input*	1.34	1.70	1.22	1.58
F_2	*input*	1.32	1.66	1.24	1.58

It seems that the different and simplest profile functions – short-range and long-range – describe spectra rather well. It is expected that by using of more intricate profile functions one may achieve good description of other parameters as well.

The analysis presented above shows that the essential role in soliton stabilization by quantum fluctuations plays the breaking of chiral symmetry, i.e. turning on the pion mass term. However, the positive results follow only by such profile functions, which are not solutions of Euler-Lagrange equation of motion.