

Dark Matter, Higgs Bosons, Supersymmetry and all that

Jack Gunion
U.C. Davis

March 10, 2011

References: For early material see Reviews of Particle Properties, Weinberg, and other standard texts. For Robertson-Walker metric and Riemann and Ricci tensors a good reference is Relativistic Astrophysics and Cosmology: A Primer by Peter Hoyle. For basic cosmology material and Boltzmann equation material related to dark matter see Kolb and Turner chapters 3 and 5, notes prepared by Bohdan Grzadkowski (http://www.fuw.edu.pl/~bohdang/wyklady/Cosmology/cosmo_09_10.html) and notes by A. Lewis (http://cosmologist.info/teaching/EU/notes_EU1_thermo.pdf). For Supersymmetry, I will follow to some extent the Supersymmetry Primer by S. Martin. Details regarding the Higgs sector of the MSSM and NMSSM will mainly follow the Higgs Hunters Guide (Gunion et al.) and related papers.

Evidence for Dark Matter

1. Luminous objects move faster than one would expect if they only felt the gravitational attraction of other visible objects.

Rotation velocity v of an object on a stable Keplerian orbit obeys $v(r) = \sqrt{GM(r)/r}$ ($M(r)$ = mass inside orbit).

If r lies outside the visible stuff and mass tracks visible then $v(r) \propto 1/\sqrt{r}$.

Instead, $v \sim \text{constant}$ as far as can be measured.

Thus, need dark matter halo with $\rho(r) \propto 1/r^2$ ($\Rightarrow M(r) \propto r$ and $v \sim \text{const.}$).

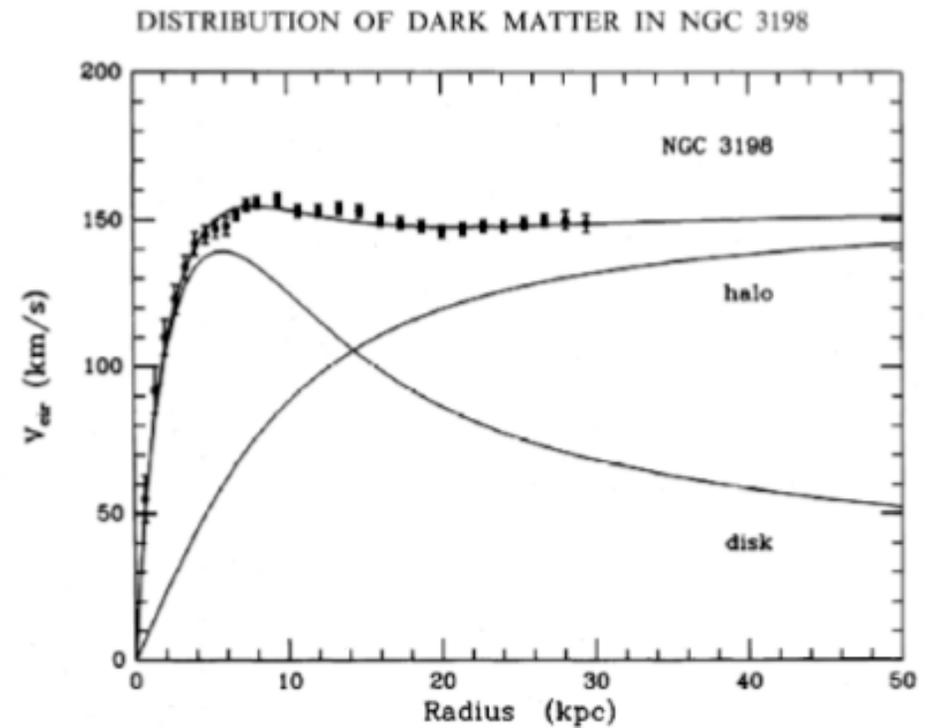
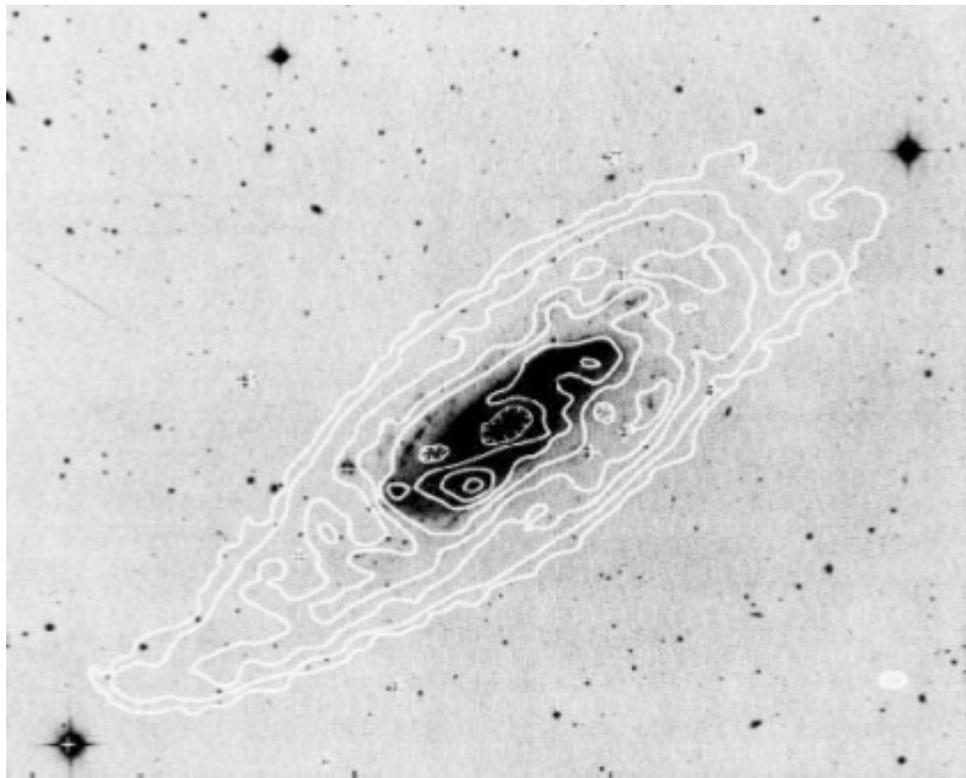
We would like to get the total mass of a given galaxy, which means we would like to be able to observe at least the start of $v \propto 1/\sqrt{r}$ and compute $M(r) = rv^2(r)/G$.

But, the rotation curve is hard to get once we run out of stars to look at.

The solution is to observe neutral Hydrogen at $\lambda = 21.1$ cm with a radio telescope.

- Most of the Hydrogen lines are in the optical or ultra-violet, but there is a very tiny magnetic energy difference between spin of the proton parallel to the spin of the electron and spin of the proton anti-parallel to the spin of the electron.
- This tiny difference in energy yields photons of $\lambda = 21.1$ cm (for Hydrogen at rest — for large z , λ is larger by factor of $1 + z$, where $1 + z \equiv R(t_0)/R(t_1)$, $R(t)$ being the radius of the universe as a function of time, see Chapt. 2 of Kolb and Turner “The Early Universe”).
- The Hydrogen has little total mass, but we can trace its orbit to measure the total mass.

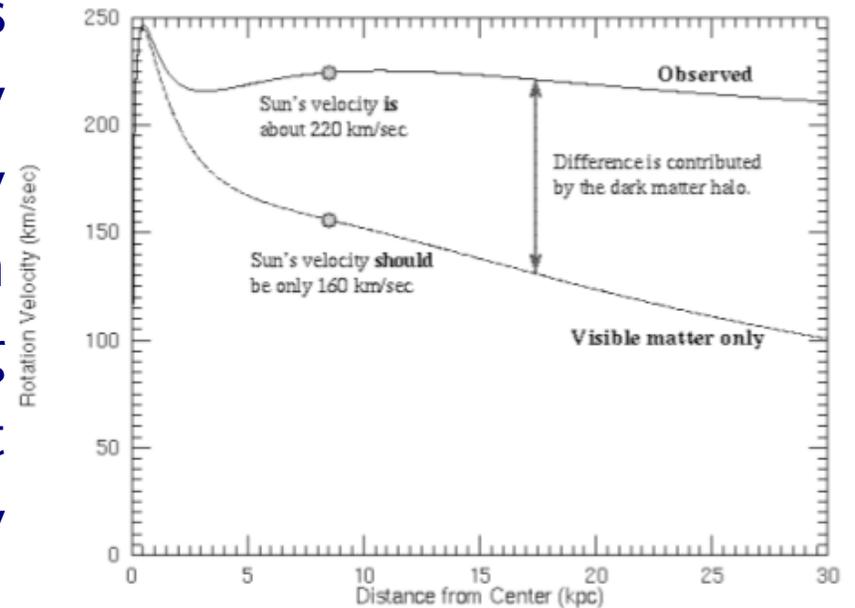
Below, on the left, is a $\lambda = 21.1$ cm radio map superimposed upon a negative optical image of galaxy NGC 3198. Note that the radio map goes way past the visible image. The rotation curve extracted from the radio image is given on the right.



- The result is that although the stars in this galaxy extend out to only 10 kpc, the rotation curve remains flat out to 30 kpc.

- The curve labeled “disk” indicates the expected rotation curve due to the visible stars in the galaxy.
- The curve labeled “halo” indicates the rotation curve due to the “dark matter halo” of the galaxy, the nature of which is not yet known.
- The exact amount of the mass associated with the stars isn’t known very well since massive stars produce most of the light but there could be many low mass stars that produce little light.
- Hence, there are other possible fits to the same $v(r)$ curve.
- Also, we do not yet see the $v(r) \propto 1/\sqrt{r}$ drop, so there is undoubtedly still more total mass beyond the Hydrogen we can detect in this way.

For our galaxy, a velocity plot is that given to the right. The gravity of the visible matter in the Galaxy is not enough to explain the high orbital speeds of the stars, including the sun which is moving about 60 km/s too fast. The discrepancy is ascribed to a dark matter halo.

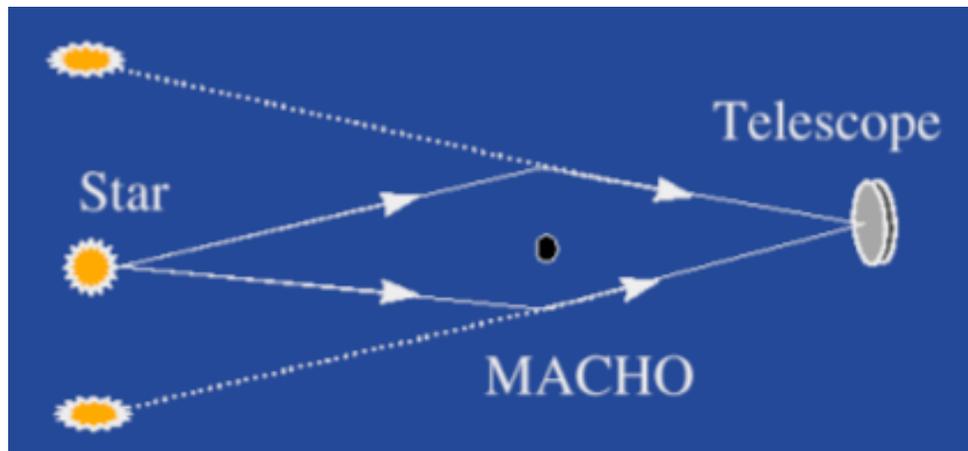


Putting a bunch of such observations together $\Rightarrow \Omega_{DM} > 0.1$ where $\Omega_X \equiv \rho_X / \rho_{crit}$, where ρ_{crit} is the critical mass density such that $\Omega_{tot} = 1$ corresponds to a flat universe (which is observationally verified to be approximately the case).

2. Observations of peculiar velocities of galaxies within clusters of galaxies, measurements of the X -ray temperature of the hot gas in the cluster (which correlates with the gravitational potential felt by the gas) and studies of (weak) gravitational lensing of background

galaxies all point to $\Omega_{DM} \sim 0.2$.

For example, in gravitational lensing, you look for multiple images of a single source as shown in the l.h. diagram. Typically one sees arc images as illustrated in the r.h. picture, which is the image of the cluster 0024+1654.



From the amount of lensing, one can determine the mass of the invisible (dark) matter between the cluster and the observer.

3. The famous bullet cluster that passed through another cluster

shows baryonic (visible) matter being decelerated and shocked, whereas the galaxies in the clusters proceeded on ballistic trajectories. Gravitational lensing shows that most of the total mass also moved ballistically, indicating that DM self-interactions are weak.

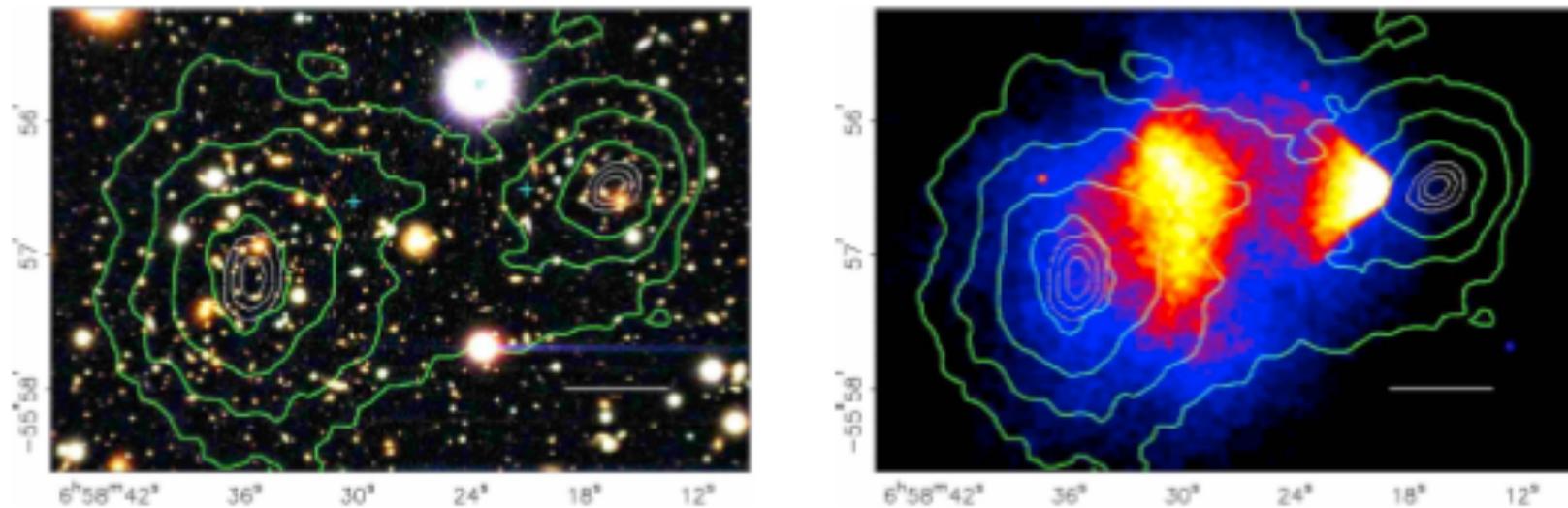
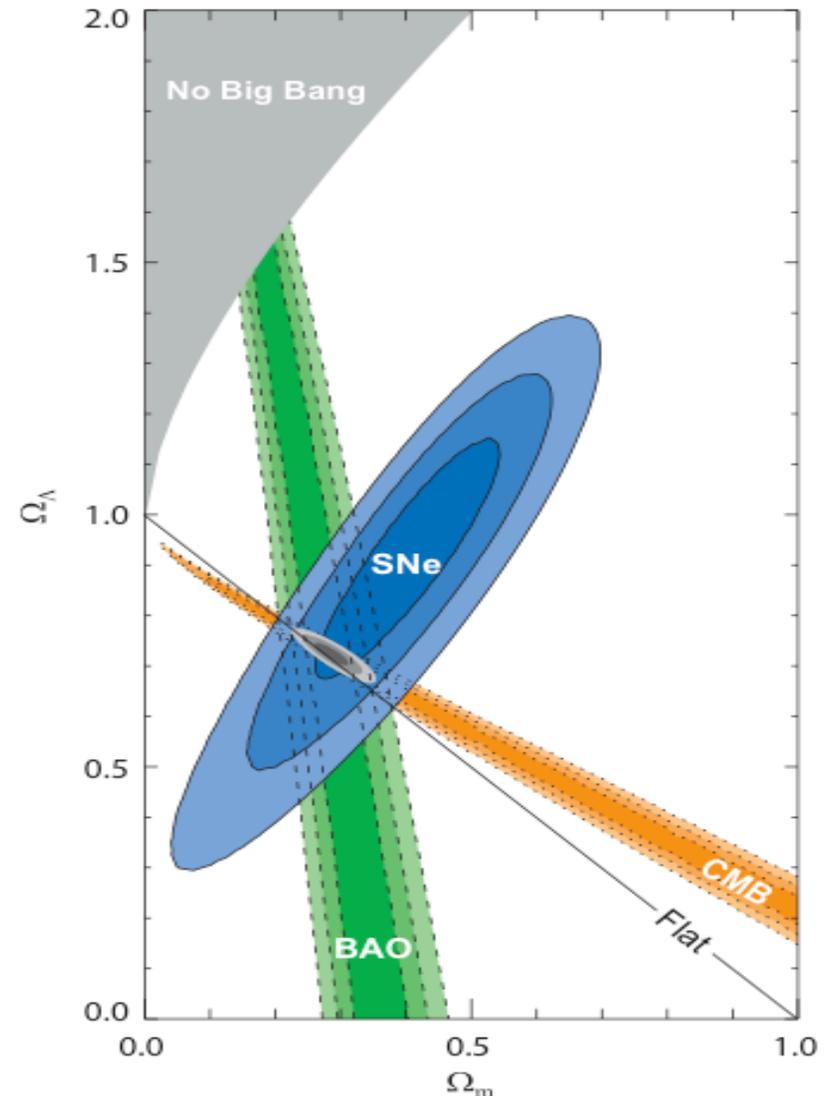
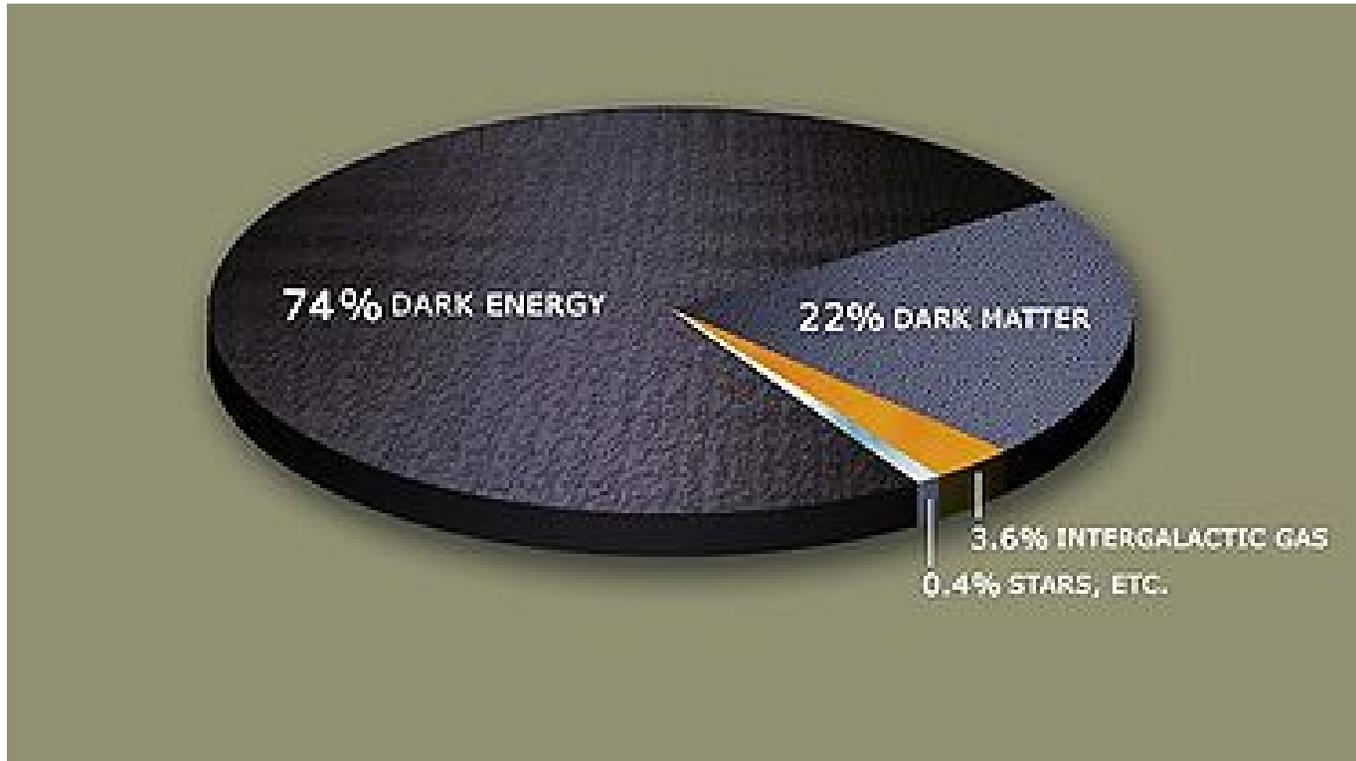


FIG. 1.— Shown above in the top panel is a color image from the Magellan images of the merging cluster 1E0657–558, with the white bar indicating 200 kpc at the distance of the cluster. In the bottom panel is a 500 ks Chandra image of the cluster. Shown in green contours in both panels are the weak lensing κ reconstruction with the outer contour level at $\kappa = 0.16$ and increasing in steps of 0.07. The white contours show the errors on the positions of the κ peaks and correspond to 68.3%, 95.5%, and 99.7% confidence levels. The blue +s show the location of the centers used to measure the masses of the plasma clouds in Table 2.

4. The most accurate determination of Ω_{DM} (albeit somewhat indirect) comes from a simultaneous fit to a variety of cosmological measurements.

The summary plot for the energy/matter content of the universe (normalized to $\Omega = 1$) is shown to the right. We are interested in the mass component, most of which is not visible, *i.e.* is **dark matter**. It comprises about 20% of the total. The observations are from Super Novae, Cosmological Microwave Background (WMAP5), and Baryonic Acoustic Oscillations (also WMAP5). WMAP8 (Gawiser colloquium) further reduces errors. Galaxy clustering also gives an ellipse that crosses the others at the common point.





5. In terms of Ωh^2 , where $h =$ Hubble constant in units of $100 \text{ km}/(s \cdot \text{Mpc})$ and $h \sim 0.7$ is the measured value,

$$\Omega_{nbm} h^2 = 0.11 \pm 0.006, \quad \text{vs.} \quad \Omega_b h^2 \sim 0.0227 \pm 0.006, \quad (1)$$

where Ω_{nbm} is the density of **cold, non-baryonic** dark matter and Ω_b is the density of **all** baryonic matter, whether visible or invisible (*e.g.*

MACHOs or cold molecular gas clouds).

Table 21.2: Parameter constraints reproduced from Dunkley *et al.* [2] and Komatsu *et al.* [3], with some additional rounding. All columns assume the Λ CDM cosmology with a power-law initial spectrum, no tensors, spatial flatness, and a cosmological constant as dark energy. Above the line are the six parameter combinations actually fit to the data; those below the line are derived from these. Two different data combinations are shown to highlight the extent to which this choice matters. The first column is WMAP5 alone, while the second column shows a combination of WMAP5 with BAO and SNe data as described in Ref. 3. The perturbation amplitude $\Delta_{\mathcal{R}}^2$ is specified at the scale 0.002 Mpc^{-1} . Uncertainties are shown at 68% confidence, and caution is needed in extrapolating them to higher significance levels due to non-Gaussian likelihoods and assumed priors.

	WMAP5 alone	WMAP5 + BAO + SN
$\Omega_b h^2$	0.0227 ± 0.0006	0.0227 ± 0.0006
$\Omega_{\text{cdm}} h^2$	0.110 ± 0.006	0.113 ± 0.003
Ω_Λ	0.74 ± 0.03	0.726 ± 0.015
n	$0.963^{+0.014}_{-0.015}$	0.960 ± 0.013
τ	0.087 ± 0.017	0.084 ± 0.016
$\Delta_{\mathcal{R}}^2 \times 10^9$	2.41 ± 0.11	2.44 ± 0.10
h	0.72 ± 0.03	0.705 ± 0.013
σ_8	0.80 ± 0.04	0.81 ± 0.03
$\Omega_m h^2$	0.133 ± 0.006	0.136 ± 0.004

6. The local DM density in the neighborhood of our solar system can be estimated using the motion of nearby stars transverse to the galactic plane and by other local observables. One finds

$$\rho_{DM}^{local} \simeq 0.3 - 0.5 \frac{\text{GeV}}{\text{cm}^3}, \quad (2)$$

which is not too different from that of luminous matter (stars, gas, dust). The most recent analyzes favor values towards the upper end of this range.

The above local density is far above the average DM density for the universe as a whole. This is, of course, expected since we reside in a dark matter halo. More precisely, the average dark matter content of the universe is about $0.22 \times \rho_{crit}$ where

$$\rho_{crit} = \frac{3H_0^2}{8\pi G_N} = 1.05368 \times 10^{-5} h^2 \frac{\text{GeV}}{\text{cm}^3} \stackrel{h \sim 0.7}{\sim} 0.5 \times 10^{-5} \frac{\text{GeV}}{\text{cm}^3}. \quad (3)$$

In any case, there is little doubt that there is a great deal of dark matter present in typical galaxies and galaxy clusters, but as of the moment we have no idea what it is.

Candidates for Dark Matter

The favorite possibility is that there is an invisible (we don't see it), weakly interacting (or we would already have seen its interactions), neutral (if charged, we would have seen tracks in emulsions, strange charged bound state particles,) particle which has significant density throughout the universe.

A DM candidate must be stable on cosmological time scales (otherwise no longer around), interact weakly with electromagnetic radiation (otherwise not dark) and must have interactions and thermal history such as to give the measured Ω_{DM} .

Possibilities include the following.

1. **Neutrinos** One early idea was that maybe the neutrinos comprised dark matter. However, neutrinos are now known to have such low masses that they would be rather relativistic (termed “warm”),

whereas cosmological observations show that the dark matter should be “cold” (*i.e.* have mass of order 1 keV or larger).

The limit on warm dark matter requires (WMAP5 + ...)

$$\Omega_\nu h^2 \leq 0.0067 \quad 95\% \text{ CL}. \quad (4)$$

This agrees well with direct upper bounds for light neutrinos.

2. **“Primordial” Black holes** They would need to be formed before the era of Big-Bang nucleosynthesis, since otherwise they would have been counted in Ω_b , which value comes from considering abundances of elements formed during BBN.

This is not absolutely impossible, but requires a very contrived cosmological model.

3. **Axions** The axion was proposed as a way to solve the strong CP problem of QCD; they also occur naturally in superstring theories.

They are pseudo Nambu-Goldstone bosons (PNGBs) associated with the (mostly) spontaneous breaking of a global “Peccei-Quinn” (PQ) $U(1)$ symmetry at high energy scale f_a .

Although very light, axions would constitute cold DM since they were produced non-thermally.

At temperatures T well above the QCD phase transition (at $T \sim \text{GeV}$), the axion is massless and the axion field can take any value, parametrized by the “misalignment angle”, θ_i .

At $T \leq \text{GeV}$, the axion develops a mass m_a due to instanton effects. Unless the axion field a happens to find itself at the minimum of its potential ($\theta_i = 0$), it will begin to oscillate once m_a becomes comparable to the Hubble parameter H . These coherent oscillations transform the energy originally stored in the axion field into physical axion quanta. The result is

$$\Omega_a h^2 = \kappa_a (f_a / 10^{12} \text{ GeV})^{1.175} \theta_i^2, \quad (5)$$

where $\kappa_a \sim 0.5 - \text{few}$. If $\theta_i \sim \mathcal{O}(1)$, this result saturates $\Omega_{n b m} h^2 \sim 0.11$ for $f_a \sim 10^{11}$ GeV, comfortably above laboratory and astrophysical constraints. This would correspond to $m_a \sim 0.1$ MeV.¹

However, if the post-inflationary reheat temperature $T_R > f_a$, cosmic strings will form during the PQ phase transition at $T \simeq f_a$, and their decay will give an additional contribution to Ω_a which is often bigger than the above result. This would require a smaller f_a and, therefore, larger m_a .

On the other hand, values of f_a near M_{P} become possible if θ_i is small for some reason.

4. **WIMPs** WIMP stands for “weakly-interacting massive particle”, conventionally denoted by χ . These are particles with mass roughly

¹The axion mass is given by $m_a f_a \simeq m_\pi f_\pi$ where $m_\pi = 135$ MeV and $f_\pi = 92$ MeV. In more detail one finds $m_a = z^{1/2} (1+z)^{-1} \frac{f_\pi m_\pi}{f_a} = \frac{0.6 \text{ MeV}}{(f_a/10^{10} \text{ GeV})}$, where $z = m_u/m_d$. Above, we have used the canonical value $z = 0.56$, although the range $z = 0.35 - 0.60$ is plausible.

between *few* GeV and *few* TeV, and with cross sections of approximately weak strength.

Within standard cosmology, their present relic density can be calculated reliably if the WIMPs were in thermal and chemical equilibrium with the hot soup of Standard Model (SM) particles after inflation. In this case, their density would become exponentially (Boltzmann) suppressed at $T < m_\chi$.

The WIMPs therefore drop out of thermal equilibrium (freeze out) once the rate of reactions that change SM particles into WIMPs or vice versa,

$$rate \propto n_{WIMP} \times \sigma_A \times v_{rel} \quad (6)$$

becomes smaller than the Hubble expansion rate of the Universe. Here, n_{WIMP} is the number density of WIMPs, σ_A is the cross section for WIMP-pair-annihilation to SM particles, and v_{rel} is the relative velocity of the annihilating WIMPs.

After freeze out, the co-moving WIMP density remains essentially constant; if the Universe evolved adiabatically after WIMP decoupling, this implies a constant WIMP number to entropy density ratio.

Their present relic density is then approximately given by (ignoring logarithmic corrections)

$$\Omega_\chi h^2 \simeq \text{const.} \times \frac{T_0^3}{M_{\text{P}}^3 \langle \sigma_A v \rangle} \simeq \frac{0.1 \text{ pb} \cdot c}{\langle \sigma_A v \rangle}. \quad (7)$$

Here T_0 is the current CMB temperature, M_{P} is the Planck mass, c is the speed of light, σ_A is the total annihilation cross section of a pair of WIMPs into SM particles, v is the relative velocity between the two WIMPs in their cms system, and $\langle \dots \rangle$ denotes thermal averaging.

Freeze out happens at temperature $T_F \simeq m_\chi/20$ almost independently of the properties of the WIMP. This means that WIMPs are already non-relativistic when they decouple from the thermal plasma; it also

implies that Eq. (7) is applicable if $T_R > T_F$.

Notice that the 0.1 pb in Eq. (7) contains factors of T_0 and M_P ; it is, therefore, quite intriguing that it happens to come out near the typical size of weak interaction cross sections. This is called the **WIMP Miracle**.

WIMP Candidates

Heavy neutrino

The seemingly most obvious WIMP candidate is a heavy neutrino. However, an $SU(2)$ doublet neutrino will have too large a cross section and, therefore, too small a relic density if its mass exceeds $m_Z/2$, as required by LEP data.

One can suppress the annihilation cross section, and hence increase the relic density, by postulating mixing between a heavy $SU(2)$

doublet and some sterile $SU(2) \times U(1)_Y$ singlet neutrino. However, one also has to require the neutrino to be stable; it is not obvious why a massive neutrino should not be allowed to decay.

LSP

The currently best motivated WIMP candidate is, therefore, the lightest superparticle (LSP) in supersymmetric models with exact R -parity (which guarantees the stability of the LSP).

Searches for exotic isotopes imply that a stable LSP has to be neutral.

This leaves basically two candidates among the superpartners of ordinary particles:

- (a) a sneutrino (supersymmetric partner of a neutrino),
- (b) and a neutralino (a mixture of the spin-1/2 supersymmetric partners of the γ , Z gauge bosons and the two neutral Higgs bosons of the minimal supersymmetric model plus, possibly, others).

Sneutrinos

Sneutrinos have quite large annihilation cross sections; their masses would have to exceed several hundred GeV for them to make good DM candidates.

This is uncomfortably heavy for the lightest sparticle, in view of naturalness arguments.

Moreover, the negative outcome of various WIMP searches (see below) rules out ordinary sneutrinos as a primary component of the DM halo of our galaxy. (In models with gauge-mediated SUSY breaking, the lightest messenger sneutrino could make a good WIMP.)

Neutralinos

The most widely studied WIMP is therefore the lightest neutralino. Detailed calculations (some of which we shall do) show that the

lightest neutralino will have the desired thermal relic density Eq. (7) in at least four distinct regions of parameter space:

- (a) χ could be (mostly) a bino or photino (the superpartner of the $U(1)_Y$ gauge boson and photon, respectively), if both χ and some sleptons have mass below ~ 150 GeV;
- (b) if m_χ is close to the mass of some sfermion (so that its relic density is reduced through co-annihilation with this sfermion);
- (c) if $2m_\chi$ is close to the mass of the CP-odd Higgs boson present in supersymmetric models;
- (d) if χ has a large higgsino or wino component.

Other WIMP Models

- (a) Many nonsupersymmetric extensions of the Standard Model also contain viable WIMP candidates. Examples are the lightest T -odd particle in Little Higgs models with conserved T -parity, or technibaryons in scenarios with an additional, strongly interacting

(technicolor or similar) gauge group.

- (b) Models where the DM particles, while interacting only weakly with ordinary matter, have quite strong interactions within an extended dark sector of the theory. These were spurred by measurements by the PAMELA, ATIC and Fermi satellites indicating excesses in the cosmic e^+ and/or e^- fluxes at high energies.

However, these excesses are relative to background estimates that are clearly too simplistic (e.g., neglecting primary sources of electrons and positrons, and modeling the galaxy as a homogeneous cylinder).

Moreover, the excesses, if real, are far too large to be due to usual WIMPs, but can be explained by astrophysical sources.

It therefore seems unlikely that they are due to Dark Matter.

- (c) Although thermally produced WIMPs are attractive DM candidates because their relic density naturally has at least the right order of magnitude, non-thermal production mechanisms have also been

suggested, e.g., LSP production from the decay of some moduli fields, from the decay of the inflaton, or from the decay of Q balls (non-topological solitons) formed in the wake of Affleck-Dine baryogenesis.

Although LSPs from these sources are typically highly relativistic when produced, they quickly achieve kinetic (but not chemical) equilibrium if T_R exceeds a few MeV (but stays below $m_\chi/20$). They therefore also contribute to cold DM.

DM Detection

- Primary black holes (as MACHOs), axions, and WIMPs are all (in principle) detectable with present or near-future technology. This was presumably what some of you learned about in the fall quarter.
- There are also particle physics DM candidates which currently seem almost impossible to detect, unless they decay; the present lower

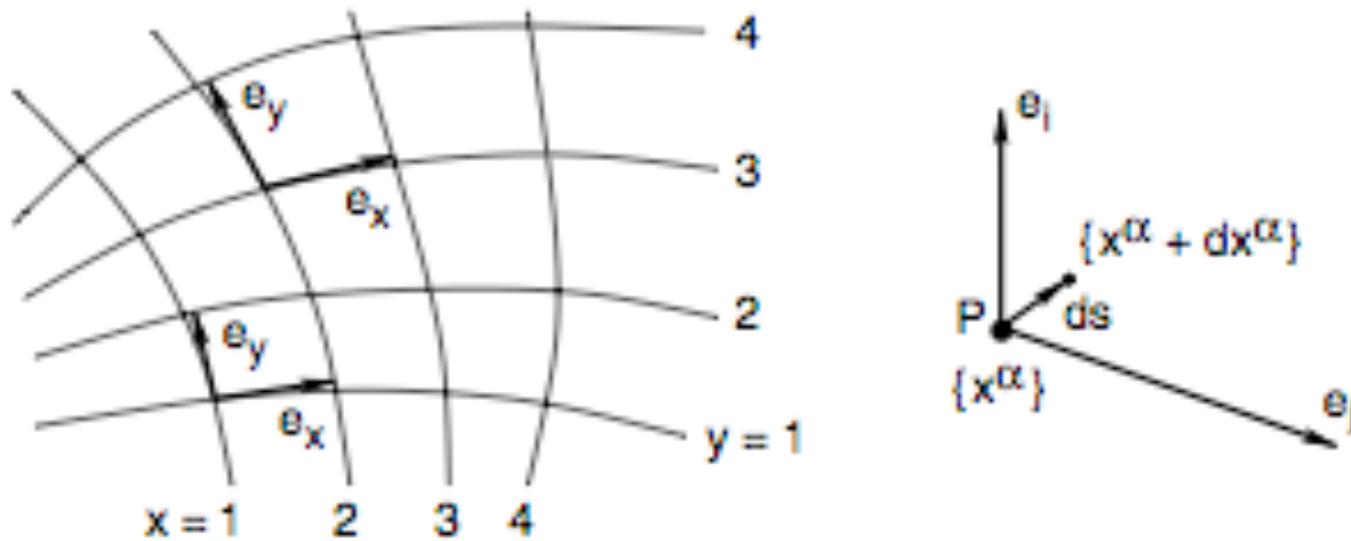
limit on their lifetime is of order $10^{25} - 10^{26}$ s for 100 GeV particles.
These include:

1. the gravitino (the spin-3/2 superpartner of the graviton),
2. states from the hidden sector thought responsible for supersymmetry breaking, and
3. the axino (the spin-1/2 superpartner of the axion).

General Relativity Basics

- The Metric:

The fundamental quantity is the metric $g_{\alpha\beta}$. Consider the curves $\{x^\alpha(p)\}$ through a point P in Riemann space ($p =$ curve parameter). At any given point there will be a set of tangent vectors that indicate how x^α is changing as you move along the curve: $ds = dx^\alpha e_\alpha$.



The length of ds (along the curve p) is given by

$$ds^2 = ds \cdot ds = (dx^\alpha \mathbf{e}_\alpha) \cdot (dx^\beta \mathbf{e}_\beta) = (\mathbf{e}_\alpha \cdot \mathbf{e}_\beta) dx^\alpha dx^\beta \equiv g_{\alpha\beta} dx^\alpha dx^\beta. \quad (8)$$

An arbitrary vector can be defined at a given location P in terms of the tangent vectors at location P : $\mathbf{A} = A^\mu \mathbf{e}_\mu$.

- **Parallel Transport:**

At any given point on the curve, we can define a vector $\mathbf{A}(\mathbf{P})$ in terms of the current tangent basis: $\mathbf{A}(P) = A^\mu \mathbf{e}_\mu$. If we now move to a new point Q on the curve, $\mathbf{A}(Q)$ will in general change, but the question is by how much and how we should define that change. Parallel transport of \mathbf{A} is defined by saying that $\mathbf{A}(Q)$ should have the same orientation (and length) with respect to the new tangent basis as did $\mathbf{A}(P)$ with respect to the previous tangent basis.

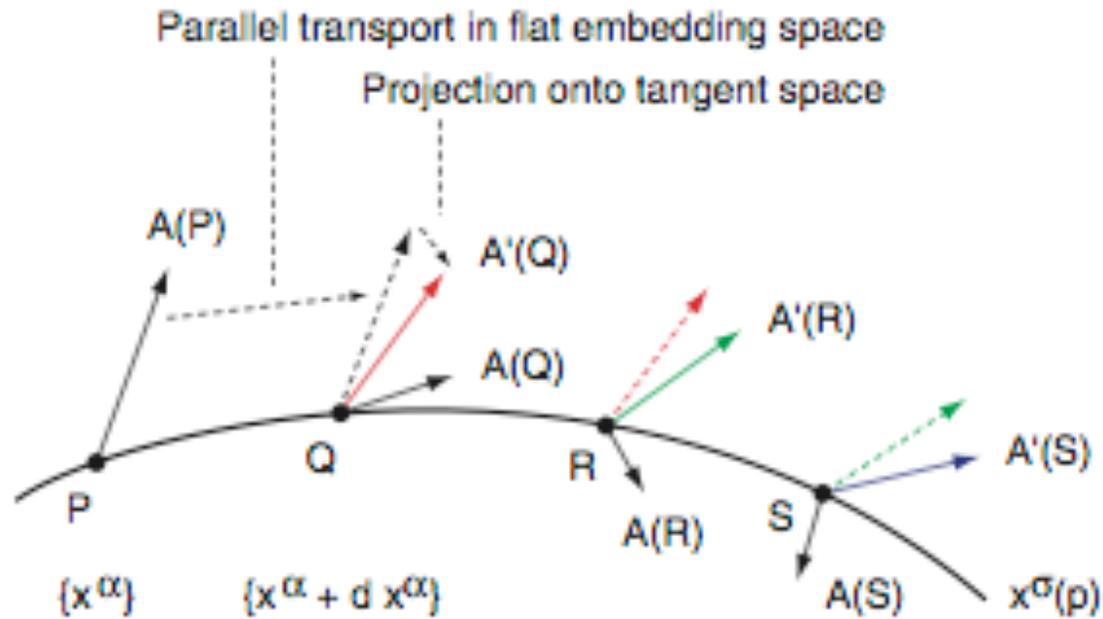


Figure 1: Conceptual definition of parallel displacement of a vector along a curve $x^\sigma(p)$ in Riemann space. First, an ordinary parallel displacement in the flat embedding space (resulting in the dashed arrows). Then followed by projection on the local tangent space. The process is repeated in infinitesimal steps.

The amount of change of \mathbf{A} not associated with maintaining same orientation with respect to the local tangent space is defined as

$$d\mathbf{A} = \mathbf{A}(Q) - \mathbf{A}'(Q) = d(A^\mu \mathbf{e}_\mu) = (dA^\mu) \mathbf{e}_\mu + A^\mu (d\mathbf{e}_\mu). \quad (9)$$

In other words, $d\mathbf{A}$ may be interpreted as the intrinsic change of \mathbf{A} , after correction for the “irrelevant” change in the orientation of the tangent space.

In any case, the important point is that after the move to the new location on the curve, there will be a new set of basis vectors. The change in the tangent vectors is defined via the Christoffel symbol:

$$d\mathbf{e}_\mu = \Gamma_{\mu\beta}^\alpha dx^\beta \mathbf{e}_\alpha . \quad (10)$$

In terms of the Christoffel symbols, we have

$$d\mathbf{A} = (dA^\mu + \Gamma_{\nu\sigma}^\mu A^\nu dx^\sigma) \mathbf{e}_\mu \equiv (DA^\mu) \mathbf{e}_\mu . \quad (11)$$

Parallel transport would correspond to maintaining $DA^\mu/dp = 0$ as we move along the curve parametrized by p .

- Geodesics:

A geodesic is a line that is “as straight as possible” on a curved surface. We say that a curve $x^\mu(p)$ is a geodesic when the tangent vector $\dot{x}^\mu \equiv dx^\mu/dp$ remains a tangent vector under parallel transport along $x^\mu(p)$. According to Eq. (11), this means we require

$$\frac{D}{Dp} \left(\frac{dx^\mu}{dp} \right) = 0, \quad \Rightarrow \quad \ddot{x}^\mu + \Gamma_{\nu\sigma}^\mu \dot{x}^\nu \dot{x}^\sigma = 0. \quad (12)$$

Now, Eq. (12) can also be derived from a variational principle. The simplest is $\delta \int \dot{s} dp = 0$ (*i.e.* the shortest path in terms of ds). This is equivalent to $\delta \int \dot{s}^2 dp = 0$, or

$$\delta \int L dp = 0, \quad \text{with} \quad L(x^\alpha, \dot{x}^\beta) = \left(\frac{ds}{dp} \right)^2 = g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta. \quad (13)$$

The solution is determined by the Euler-Lagrange equations

$$\frac{\partial L}{\partial x^\lambda} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{x}^\lambda} \right) = 0. \quad (14)$$

Now, $\partial L / \partial x^\lambda = g_{\alpha\beta,\lambda} \dot{x}^\alpha \dot{x}^\beta$ since only $g_{\alpha\beta}$ depends on $\{x^\mu\}$. Then, by using $\partial \dot{x}^\alpha / \partial \dot{x}^\lambda = \delta_\lambda^\alpha$ one gets $\partial L / \partial \dot{x}^\lambda = 2g_{\alpha\lambda} \dot{x}^\alpha$. Substituting these in to the Euler Lagrange equation gives:

$$g_{\alpha\beta,\lambda} \dot{x}^\alpha \dot{x}^\beta = 2(\dot{g}_{\alpha\lambda} \dot{x}^\alpha) = 2(g_{\alpha\lambda,\beta} \dot{x}^\beta \dot{x}^\alpha + g_{\alpha\lambda} \ddot{x}^\alpha), \quad (15)$$

or

$$g_{\alpha\lambda} \ddot{x}^\alpha + \frac{1}{2}(2g_{\alpha\lambda,\beta} - g_{\alpha\beta,\lambda}) \dot{x}^\alpha \dot{x}^\beta = 0. \quad (16)$$

It is useful to write

$$2g_{\lambda\alpha,\beta} \dot{x}^\alpha \dot{x}^\beta = g_{\lambda\alpha,\beta} \dot{x}^\alpha \dot{x}^\beta + g_{\lambda\beta,\alpha} \dot{x}^\beta \dot{x}^\alpha = (g_{\lambda\alpha,\beta} + g_{\lambda\beta,\alpha}) \dot{x}^\alpha \dot{x}^\beta. \quad (17)$$

Then, multiplying Eq. (16) by $g^{\mu\lambda}$ and substituting the above, we obtain

$$\ddot{x}^\mu + \frac{1}{2}g^{\mu\lambda}(g_{\lambda\alpha,\beta} + g_{\lambda\beta,\alpha} - g_{\alpha\beta,\lambda})\dot{x}^\alpha\dot{x}^\beta = 0. \quad (18)$$

Comparing to Eq. (12), we conclude that the Christoffel symbols are intimately connected to the metric tensor, the relation being

$$\Gamma_{\alpha\beta}^\mu = \frac{1}{2}g^{\mu\lambda}(g_{\lambda\alpha,\beta} + g_{\lambda\beta,\alpha} - g_{\alpha\beta,\lambda}). \quad (19)$$

The Lagrangian method provides the quickest derivation of the above relation and we have sketched it so that we can use it again when we come to the Robertson-Walker metric appropriate to the Early Universe.

- Riemann Tensor, Ricci Tensor,

Various subsidiary quantities employed in General Relativity can be defined in terms of the Christoffel symbol.

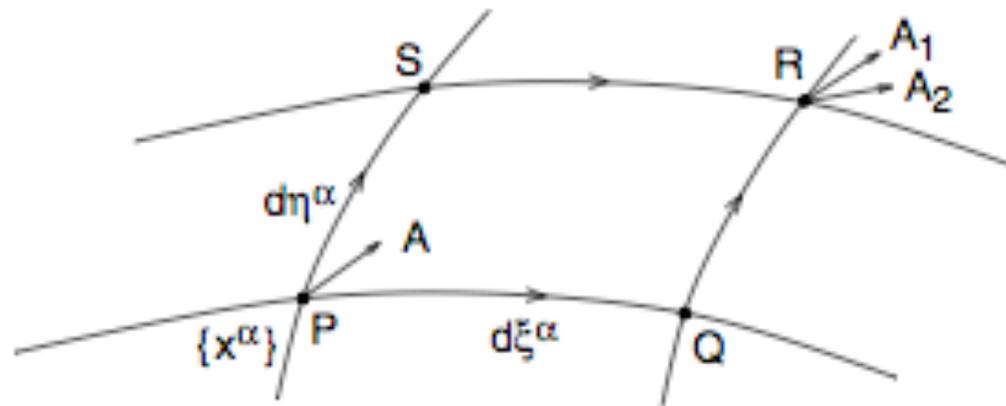
First, there is the Riemann Tensor:

$$R^\alpha_{\beta\gamma\delta} = \Gamma^\alpha_{\beta\delta,\gamma} - \Gamma^\alpha_{\beta\gamma,\delta} + \Gamma^\mu_{\beta\delta}\Gamma^\alpha_{\mu\gamma} - \Gamma^\mu_{\beta\gamma}\Gamma^\alpha_{\mu\delta}. \quad (20)$$

From the Riemann Tensor one finally obtains the Ricci Tensor (or curvature tensor) and the scalar curvature:

$$R_{\mu\kappa} = R^\lambda_{\mu\lambda\kappa}, \quad \mathcal{R} = R^\mu_{\mu}. \quad (21)$$

These are all related to the intrinsic curvature of space as follows. Let us imagine parallel transport of a vector \mathbf{A} from P to R along two paths: $1 = (PQR)$ and $2 = (PSR)$ (see figure).



For each piece we write $dA^\mu = -f_\sigma(x)dx^\sigma$ with $f_\sigma(x) = \Gamma_{\nu\sigma}^\mu A^\nu$.
Then,

$$\begin{aligned}
& A_1^\mu - A_2^\mu \\
&= -f_\sigma(x)d\xi^\sigma - f_\sigma(x+d\xi)d\eta^\sigma + f_\sigma(x)d\eta^\sigma + f_\sigma(x+d\eta)d\xi^\sigma \\
&\simeq -f_\sigma d\xi^\sigma - f_\sigma d\eta^\sigma - f_{\sigma,\lambda}d\xi^\lambda d\eta^\sigma + f_\sigma d\eta^\sigma + f_\sigma d\xi^\sigma + f_{\sigma,\lambda}d\eta^\lambda d\xi^\sigma \\
&= (f_{\sigma,\lambda} - f_{\lambda,\sigma})d\xi^\sigma d\eta^\lambda \\
&= R_{\alpha\lambda\sigma}^\mu A^\alpha d\xi^\sigma d\eta^\lambda.
\end{aligned} \tag{22}$$

One easily verifies that $R_{\mu\nu\sigma}^\alpha$ is zero in a flat space for any choice of the co-ordinates. The above equation then implies that parallel transport along a *closed* path leaves a vector unchanged. But, in a curved space the orientation of the vector will have changed. In 4 dimensions, after using symmetries, one finds that $R_{\nu\rho\sigma}^\alpha$ has 20 independent components. Further, all contractions of $R_{\nu\rho\sigma}^\alpha$, *i.e.* $R_{\mu\nu}$, are either zero or equal, apart from a sign.

- Einstein's Equations:

Einstein's equations are stated in terms of the Einstein Tensor

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\mathcal{R}. \quad (23)$$

They are

$$G_{\mu\nu} - \Lambda g_{\mu\nu} = -8\pi G T_{\mu\nu} \quad (24)$$

where Λ is the cosmological constant.

There are a number of solutions to Einstein's equations for $T_{\mu\nu} = 0$.

- One is of course the standard Minkowski metric.
- However, we are interested in solutions that have intrinsic curvature.

The Basics of the Universe

- An important observation is that the universe is isotropic. The distribution of matter in space is *statistically* the same in all directions, and also as a function of distance, *i.e.* within redshift subclasses.
- There are obvious *evolution effects*. The morphology of the systems changes gradually with distance.
- Hubble, in 1929, demonstrated that the universe expands with time. All galaxies move away from us on average with a velocity proportional to the distance, but independent of direction. This universal expansion is referred to as the Hubble flow:

$$v = H_0 d \quad (25)$$

with $H_0 = 100h \text{ km s}^{-1} \text{ Mpc}^{-1}$, where $h = 0.71 \pm 0.04$ as measured by WMAP. In physical units,

$$H_0 = (2.3 \pm 0.1) \times 10^{-18} \text{ s}^{-1}. \quad (26)$$

The peculiar velocities of the systems within the Universe, *i.e.* the deviations from the Hubble flow, are generally small, $\lesssim 500 \text{ km s}^{-1}$. The Hubble flow is thus 'cold' and this is because the universe cools adiabatically.

- Coordinates:

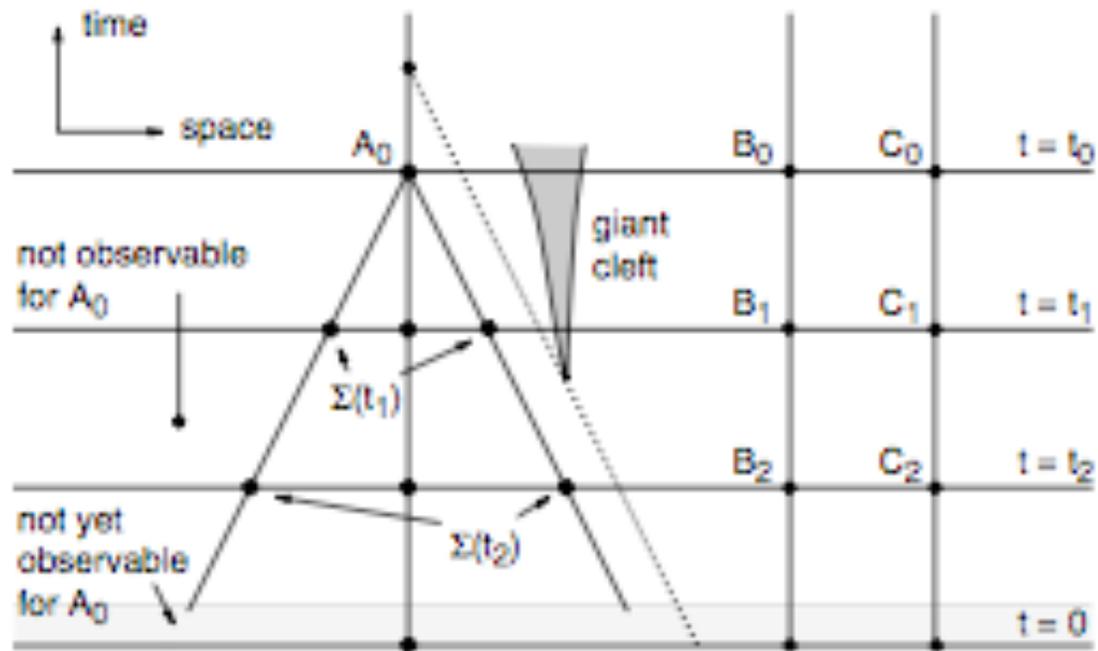


Figure 2: A picture of the spacetime of the universe. Our present position is A_0 . Also indicated is our world line and our past light-cone. The worldlines of a few other galaxies (vertical lines), B and C , are also shown. Finally, a hypothetical inhomogeneity (“giant cleft”) that we might get to see in the future is shown.

- We are only able to see events located within or on our past light-cone. We experience our light-cone as a series of nested, ever-larger concentric spherical shells around us, showing an increasingly younger section of the universe.

- Because of the observed isotropy, each shell $\Sigma(t_i)$ must be on average homogenous.
- Due to our limited technological capabilities, we have not yet been able to detect signals from the early universe, *i.e.* from the most distant shells (the shaded region at the bottom of the figure).
- We now make assumptions about the part of space-time that is outside our past light-cone and therefore unobservable.

To that end, we use the **Cosmological Principle**, which states that we (A_0) occupy no special position in the universe, and that other observers B_0 and C_0 see on average the same universe as we do.

Hence, if we translate our light-cone sideways, the aspect of the shells $\Sigma(t_i)$ would not change, apart from statistical fluctuations (the so-called cosmic variance).

The implication is then that every subspace at $t = \text{const.}$ is isotropic and homogeneous on average.

Further, the Cosmological Principle and the isotropy of the universe imply then that the universe as a whole is homogeneous.

- **The definition of “rest”**: We are free to adopt any definition we like, but there is one that stands out as very natural: a test mass is at **rest** if it does not move with respect to the Hubble flow.

That means the spatial co-ordinates of galaxies are constant (ignoring their peculiar velocities).

Their worldlines are straight vertical lines in the figure, which is a coordinate picture and contains no information about the geometry.

Due to the expansion of the universe, the *geometrical* distance between B_0 and C_0 is larger than between B_1 and C_1 .

It remains possible that the spacetime that we shall see in the future contains huge inhomogeneities, and that the Cosmological Principle will eventually prove to be incorrect, *e.g.* the giant cleft of the figure

could appear.

Presently, however, the assumption that every subspace $t = \text{const.}$ is homogeneous and isotropic is adequate.

- The first step to a definite coordinate system:

It can be shown that one can always define a time that is separate from spatial slices in such a way that

$$ds^2 = (dx^0)^2 + g_{ik} dx^i dx^j . \quad (27)$$

These are called Gaussian co-ordinates.

The essence of Gaussian co-ordinates is that the world lines of a selected set of freely falling test masses are taken as the co-ordinate lines of the co-ordinate system and these lines remain always orthogonal to the sub-spaces $t = \text{const.}$

In cosmology, the sections $t = \text{const}$ are snapshots of the homogeneous and isotropic universe, and the selected test masses are the galaxies.

Because these are at rest ($dx^i = 0$) it follows that $d\tau \equiv ds = dt$. This must be so because otherwise a subspace $t = \text{const}$ would not be homogeneous.

- **Metric and spatial structure:**

In order to describe an expanding universe, it is clear that the metric must depend on x^0 , and that dependence must be the same for every g_{ik} as otherwise anisotropies would develop. The implication is that we can write

$$ds^2 = (dx^0)^2 + S^2(t)a_{ik}dx^i dx^k, \quad (28)$$

with $a_{ik} = \text{const}$ in time.

We may simplify a_{ik} by noting that the space is certainly spherically symmetric around an (arbitrarily chosen) origin. The result (after a

bit of argumentation and rescalings that I omit) is that we can write

$$ds^2 = (dx^0)^2 - S^2(t) \left(e^{2\lambda(r)} dr^2 + r^2 d\Omega \right), \quad (29)$$

To find $\lambda(r)$ we compute the total (spatial) curvature ${}^3\mathcal{R} = R_i^i$ of the $t = \text{const}$ subspace when $S(t) = 1$. One finds (using techniques of geodesics, that we will come to shortly, in order to get the Christoffel symbols)

$${}^3\mathcal{R} = 2 \left(\frac{2\lambda'}{r} - \frac{1}{r^2} \right) e^{-2\lambda} + \frac{2}{r^2} = \frac{2}{r^2} \left(1 - \frac{d}{dr} [r e^{-2\lambda}] \right). \quad (30)$$

From this it follows that

$$\frac{d}{dr} [r e^{-2\lambda}] = 1 - \frac{1}{2} {}^3\mathcal{R} r^2. \quad (31)$$

Now ${}^3\mathcal{R}$ must be constant as a function of r because the space

$t = \text{const}$ is homogeneous.

We can then integrate to obtain

$$e^{-2\lambda} = 1 - \frac{1}{6} {}^3\mathcal{R} r^2 + \frac{A}{r}. \quad (32)$$

The integration constant A should be 0; otherwise the co-ordinates would not be locally flat at $r = 0$. Thus,

$$e^{2\lambda} = \frac{1}{1 - kr^2}, \quad (33)$$

where we have defined ${}^3\mathcal{R} \equiv 6k$. In this way, we arrive at ...

- **The Robertson-Walker (RW) metric:** It takes the form (in terms of the “co-moving” coordinates (r, t, θ, ϕ))

$$ds^2 = dt^2 - S^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \right] \quad (34)$$

This is the metric for a space with homogeneous and isotropic spatial sections. $S(t)$ is the “cosmic scale factor”. With an appropriate rescaling of coordinates, we can choose $k = +1, -1$ or 0 for spaces of constant positive, negative or zero curvature, respectively.

The coordinate r is dimensionless and ranges from $r = 0$ to $r = 1$ if $k = 1$. In this case, there is a singularity at $r = 1$ — one cannot consider “distances” $r \times S(t)$ larger than the cosmic scale factor, $S(t)$, at time t . For $k = 1$,

- the circumference of a one-sphere (a circle at constant ϕ and r) is as expected — $2\pi S(t)r$;
- the area of a two-sphere at constant r is as expected — $4\pi S^2(t)r^2$;
- however, the physical radius of such one and two spheres is defined in terms of $\int ds = S(t) \int_0^r \frac{dr'}{\sqrt{1-kr'^2}}$ rather than $S(t)r$.

The time coordinate being employed is just the proper (or clock) time measured by an observer at rest in the comoving frame, *i.e.*

$(r, \theta, \phi) = \text{const.}$. As stressed earlier, observers at rest in the comoving frame remain at rest, *i.e.* (r, θ, ϕ) remain unchanged, and observers initially moving with respect to this frame will eventually come to rest in it.

Further, if one introduces a homogeneous, isotropic fluid initially at rest in this frame, the $t = \text{const}$ hypersurfaces will always be orthogonal to the fluid flow, and will always coincide with the hypersurfaces of both spatial homogeneity and constant fluid density.

The above RW form gives the metric entries:

$$g_{00} = 1, \quad g_{rr} = -\frac{S^2(t)}{(1 - kr^2)}, \quad (35)$$

$$g_{\theta\theta} = -S^2(t)r^2, \quad g_{\phi\phi} = -S^2(t)r^2 \sin^2 \theta. \quad (36)$$

Finally, from the RW metric form, one finds the Ricci tensor

components and Ricci scalar that we will shortly need:

$$R_{00} = 3\frac{\ddot{S}}{S}, \quad R_{ij} = \left[\frac{\ddot{S}}{S} + 2\frac{\dot{S}^2}{S^2} + \frac{2k}{S^2} \right] g_{ij}, \quad \mathcal{R} = 6 \left[\frac{\ddot{S}}{S} + \frac{\dot{S}^2}{S^2} + \frac{k}{S^2} \right] \quad (37)$$

where $\dot{a} \equiv \frac{\partial a}{\partial t}$. However, in the derivation below, we will need to denote (as before) $\dot{a} = \frac{da}{dp}$ and temporarily use a' for $\frac{\partial a}{\partial t}$.

Derivation:

One can employ the brute force approach of computing the Christoffel symbols directly from Eq. (19).

Alternatively, we can employ the definition of the Christoffel symbol in terms of geodesics. Recall that a geodesic is defined by $\delta \int L dp = 0$, where

$$L = g_{\alpha\beta} \dot{x}^\alpha \dot{x}^\beta = (\dot{x}^0)^2 - \frac{S^2 \dot{r}^2}{1 - kr^2} - S^2 r^2 \dot{\theta}^2 - S^2 r^2 \sin^2 \theta \dot{\phi}^2. \quad (38)$$

We emphasize again that $x^0 = ct$, $x^1 = r$, $x^2 = \theta$ and $x^3 = \phi$ are considered to be functions of the curve parameter p . The scale factor S depends on t , *i.e.* on x^0 . All x^0 dependence of L is in S and we will also encounter $S' = dS/dx^0$ and $S'' = d^2S/dx^0{}^2$.

The Euler Lagrange equations resulting from the variational principle are those given earlier:

$$\frac{\partial L}{\partial x^\lambda} - \frac{d}{dp} \left(\frac{\partial L}{\partial \dot{x}^\lambda} \right) = 0. \quad (39)$$

We apply these in turn to get the $\Gamma_{\alpha\beta}^\lambda$ Christoffel symbols.

1. $\Gamma_{\alpha\beta}^0$.

Applying for $\lambda = 0$, *i.e.* requiring $\partial L / \partial x^0 = \frac{d}{dp}(\partial L / \partial \dot{x}^0)$ gives

$$-2SS' \left(\frac{\dot{r}^2}{1 - kr^2} + r^2 \dot{\theta}^2 + r^2 \sin^2 \theta \dot{\phi}^2 \right) = \frac{d}{dp}(2\dot{x}^0) = 2\ddot{x}^0 \quad (40)$$

or

$$\ddot{x}^0 + SS' \left(\frac{\dot{r}^2}{1 - kr^2} + r^2 \dot{\theta}^2 + r^2 \sin^2 \theta \dot{\phi}^2 \right) = 0. \quad (41)$$

This may be compared to the definition of the Christoffel symbol in

$$\ddot{x}^\mu + \Gamma_{\nu\sigma}^\mu \dot{x}^\nu \dot{x}^\sigma = 0 \quad (42)$$

for the case of $\mu = 0$, yielding (using the notation $1 = r$, $2 = \theta$ and $3 = \phi$)

$$\Gamma_{11}^0 = \frac{SS'}{1 - kr^2}, \quad \Gamma_{22}^0 = SS'r^2, \quad \Gamma_{33}^0 = SS'r^2 \sin^2 \theta, \quad (43)$$

and all other $\Gamma_{\alpha\beta}^0 = 0$. Note that $\Gamma_{ij}^0 = -\frac{S'}{S}g_{ij}$.

2. $\Gamma_{\alpha\beta}^2$.

For this, we employ the Euler Lagrange equation for $\lambda = 2$:
 $\partial L / \partial \theta = \frac{d}{dp}(\partial L / \partial \dot{\theta})$ which gives

$$\begin{aligned} -2S^2 r^2 \sin \theta \cos \theta \dot{\phi}^2 &= \frac{d}{dp}(-2S^2 r^2 \dot{\theta}) \\ &= -4S^2 r \dot{r} \dot{\theta} - 2S^2 r^2 \ddot{\theta} - 4SS' \dot{x}_0 r^2 \dot{\theta} \end{aligned} \quad (44)$$

Dividing by $-2S^2 r^2$ in order to have $1 \times \ddot{\theta}$, the above equation reduces to

$$\ddot{\theta} + \frac{2}{r} \dot{r} \dot{\theta} - \sin \theta \cos \theta \dot{\phi}^2 + 2 \frac{S'}{S} \dot{x}_0 \dot{\theta} = 0, \quad (45)$$

from which we read

$$\Gamma_{12}^2 = \Gamma_{21}^2 = \frac{1}{r}, \quad \Gamma_{33}^2 = -\sin \theta \cos \theta, \quad \Gamma_{02}^2 = \Gamma_{20}^2 = \frac{S'}{S}, \quad (46)$$

all other Γ_{\dots}^2 being zero.

3. $\Gamma_{\alpha\beta}^3$.

For $\lambda = 3$, Euler-Lagrange reads $\partial L/\partial\phi = \frac{d}{dp}(\partial L/\partial\dot{\phi})$ which gives

$$\begin{aligned}
 0 &= \frac{d}{dp}(-2S^2r^2 \sin^2 \theta \dot{\phi}) \\
 &= -4S^2r\dot{r} \sin^2 \theta \dot{\phi} - 4S^2r^2 \sin \theta \cos \theta \dot{\theta} \dot{\phi} \\
 &\quad - 2S^2r^2 \sin^2 \theta \ddot{\phi} - 4SS'r^2 \sin^2 \theta \dot{x}_0 \dot{\phi}.
 \end{aligned} \tag{47}$$

Dividing by $-2S^2r^2 \sin^2 \theta$ gives

$$\ddot{\phi} + \frac{2}{r}\dot{r}\dot{\phi} + 2\frac{S'}{S}\dot{x}_0\dot{\phi} + 2\frac{\cos \theta}{\sin \theta}\dot{\theta}\dot{\phi} = 0, \tag{48}$$

from which we read

$$\Gamma_{13}^3 = \Gamma_{31}^3 = \frac{1}{r}, \quad \Gamma_{23}^3 = \Gamma_{32}^3 = \cot \theta, \quad \Gamma_{03}^3 = \Gamma_{30}^3 = \frac{S'}{S}, \tag{49}$$

with all other $\Gamma_{\dots}^3 = 0$.

4. $\Gamma_{\alpha\beta}^1$.

This is the messiest case. We have

$$\begin{aligned} \frac{\partial L}{\partial r} &= -S^2 \dot{r}^2 (-1) \frac{-2kr}{(1-kr^2)^2} - S^2 2r \dot{\theta}^2 - S^2 2r \sin^2 \theta \dot{\phi}^2 \\ \frac{d}{dp} \left(\frac{\partial L}{\partial \dot{r}} \right) &= \frac{d}{dp} \left(\frac{-2S^2 \dot{r}}{(1-kr^2)} \right) \\ &= \frac{-2S^2 \ddot{r}}{(1-kr^2)} - 2S^2 \dot{r} (-1) \frac{-2k\dot{r}r}{(1-kr^2)^2} - \frac{4SS' \dot{x}_0 \dot{r}}{(1-kr^2)} \end{aligned} \quad (51)$$

Equating and isolating \ddot{r} we arrive at

$$\ddot{r} + \frac{kr}{(1-kr^2)} \dot{r}^2 - (1-kr^2)r(\dot{\theta}^2 + \sin^2 \theta \dot{\phi}^2) + 2 \frac{S'}{S} \dot{x}_0 \dot{r} = 0, \quad (52)$$

from which we read

$$\begin{aligned}\Gamma_{11}^1 &= \frac{kr}{1-kr^2}, & \Gamma_{22}^1 &= -(1-kr^2)r, \\ \Gamma_{33}^1 &= -(1-kr^2)r \sin^2 \theta, & \Gamma_{01}^1 = \Gamma_{10}^1 &= \frac{S'}{S},\end{aligned}\quad (53)$$

all other $\Gamma_{\dots}^1 = 0$.

From the above results for Γ_{\dots}^i , you will find that $\Gamma_{0j}^i = \frac{S'}{S} \delta_j^i$.

I won't go into deriving the Riemann and Ricci tensors. Results for the latter were already given earlier. For the Riemann tensor one finds the following

$$\begin{aligned}R_{0101} &= -\frac{SS''}{1-kr^2} & R_{0202} &= -Sr^2 S'' \\ R_{0303} &= -Sr^2 \sin^2 \theta S'' & R_{1212} &= \frac{S^2 r^2 (S'^2 + k)}{1-kr^2} \\ R_{1313} &= \frac{S^2 r^2 \sin^2 \theta (S'^2 + k)}{1-kr^2} & R_{2323} &= S^2 r^4 \sin^2 \theta (S'^2 + k)\end{aligned}\quad (54)$$

Finally, let me remind you again that the prime indicates derivative with respect to time above.

If we denote the spatial part of the metric as $d\vec{l}^2 = g_{ij}dx^i dx^j$. We can work out the three-dimensional Riemann tensor, Ricci tensor and Ricci scalar, denoted by ${}^3R_{ijkl}$, ${}^3R_{ij}$ and ${}^3\mathcal{R}$ with the results

$$\begin{aligned} {}^3R_{ijkl} &= \frac{k}{S^2(t)}(g_{ik}g_{jl} - g_{il}g_{kj}) \\ {}^3R_{ij} &= \frac{2k}{S^2(t)}g_{ij} \\ {}^3\mathcal{R} &= \frac{6k}{S^2(t)}. \end{aligned} \tag{55}$$

The last of these agrees with the identification of ${}^3\mathcal{R} = 6k$ for $S(t) = 1$ used when we derived the RW metric. The latter two results are also obtained by simply dropping all time derivatives in $R_{\mu\nu}$ and \mathcal{R} .

- Kolb-Turner Notation:

Since I will be following some of their material, I note that they have a notation in which $S(t)$ is replaced by $R(t)$ and they write Einstein's equations as $G_{\mu\nu} = +8\pi GT_{\mu\nu}$, but their conventions are such that their $G_{\mu\nu}$ has a sign that is opposite that employed earlier and so the final equations are the same. Of course, in their book and from here on $\dot{a} = \frac{\partial a}{\partial t}$.

At this point, we will switch to Kolb-Turner notation, *i.e.* $S \rightarrow R$, but maintain the preceding conventions for the signs of the Friedmann Tensor, Ricci Tensor and curvature and Einstein equations.

In the next few sections, I will largely follow the relevant chapters 3 and 5 of the Kolb and Turner monograph. I have also used some notes on this same kind of material prepared by B. Grzadkowski (http://www.fuw.edu.pl/~boh dang/wyklady/Cosmology/cosmo_09_10.html) and by A. Lewis (http://cosmologist.info/teaching/EU/notes_EU1_thermo.pdf). The material up to this

point came from the relevant reviews in the Reviews of Particle Properties and from a variety of standard texts. In many places I have updated some things and added additional material/explanations.

- Time, Kinematics and Red-Shift:

A fundamental question is “for a comoving observer with coordinates (r_0, θ_0, ϕ_0) , for what values of (r, θ, ϕ) would a light signal emitted at $t = 0$ reach the the observer at, or before, time t ?” The answer is determined by the metric. The furthest distance is referred to as the Horizon, with H as a subscript.

A light signal satisfies the geodesic equation $ds^2 = 0$. Because of the homogeneity of space, we may choose $r_0 = 0$. Geodesics passing through $r_0 = 0$ are lines of constant θ and ϕ . Thus, a light signal emitted from position (r_H, θ_0, ϕ_0) at $t = 0$ will reach $r_0 = 0$ in a time t determined by setting $ds^2 = 0$ as we move along the geodesic

which means requiring

$$dt = \frac{R(t)dr}{\sqrt{1 - kr^2}} \quad (56)$$

as we move along the geodesic. Integrating, we get the following consistency requirement for t :

$$\int_0^t \frac{dt'}{R(t')} = \int_0^{r_H} \frac{dr}{\sqrt{1 - kr^2}}. \quad (57)$$

The proper distance to the horizon measured at time t ,

$$d_H(t) = \int_0^{r_H} \sqrt{g_{rr}} dr \quad (58)$$

is related to the above by

$$d_H(t) = R(t) \int_0^t \frac{dt'}{R(t')}. \quad (59)$$

For standard cosmology, $d_H(t) \sim t$ is finite.

So, now let us look at a photon travelling along a geodesic. We describe its 4-momentum as usual as $p^\mu = (E, \vec{p})$, where we can assume that \vec{p} is entirely in the radial direction. The evolution of p^μ is then given by the geodesic equation

$$dp^\mu = -\Gamma_{\nu\alpha}^\mu p^\nu dx^\alpha . \quad (60)$$

For $\mu = 0$ and $\vec{p} = \hat{r}p^r$, this reduces to

$$dE = -\Gamma_{rr}^0 p^r dr = -\frac{R\dot{R}}{(1 - kr^2)} p^r dr . \quad (61)$$

Now, we must be careful to relate p^r to E correctly by using

$$g_{\mu\nu} p^\mu p^\nu = E^2 - g_{rr} p^r p^r = E^2 - \frac{R^2}{(1 - kr^2)} p^r p^r = 0 ., \quad (62)$$

which yields $p^r = E\sqrt{1 - kr^2}/R$. Substituting this and the geodesic constant θ, ϕ relation of (see Eq. (56)) $dr = dt\sqrt{1 - kr^2}/R$ into the dE equation, we obtain

$$dE = -\frac{R\dot{R}}{1 - kr^2} \left(\frac{E\sqrt{1 - kr^2}}{R} \right) \left(\frac{dt\sqrt{1 - kr^2}}{R} \right) \Rightarrow \frac{\dot{E}}{E} = -\frac{\dot{R}}{R}. \quad (63)$$

The solution is obviously that $E \propto \frac{1}{R}$. Since for a photon $\lambda \propto 1/E$, we find $\lambda \propto R$.

The standard “red-shift”, z , is defined in terms of the ratio of the detected wavelength (at t_0) to the emitted wavelength at the earlier time t_1 :

$$1 + z \equiv \frac{\lambda_0}{\lambda_1} = \frac{R(t_0)}{R(t_1)}. \quad (64)$$

Any increase in $R(t)$ with t for $t_0 > t_1$ leads to a red-shift of the light from sources that are distant from the observer, since such sources

emitted light at earlier times when $R(t)$ was smaller..

For a particle with mass, for $u^\mu \equiv dx^\mu/ds = (\gamma, \gamma\vec{v})$ (\vec{v} being the usual velocity) a similar procedure gives

$$\frac{|\dot{\vec{u}}|}{|\vec{u}|} = -\frac{\dot{R}}{R}, \quad \Rightarrow |\vec{u}| \propto R^{-1}. \quad (65)$$

implying that a particle with mass will eventually come to rest in the comoving frame.

- **The stress energy tensor:** To be consistent with the symmetries of the metric, $T_{\mu\nu}$ must be diagonal. By isotropy the spatial components must be equal. The simplest realization is

$$T^\mu{}_\nu = \text{diag}(\rho, -p, -p, -p) \quad (66)$$

i.e. as for a perfect fluid characterized by a time-dependent energy density $\rho(t)$ and pressure $p(t)$.

- **The 1st law of thermodynamics:** Given the above form, the first law follows from $T^{\mu\nu}_{;\nu} = 0$ and takes the physically expected form

$$d(\rho R^3) = -pd(R^3), \quad \text{or equivalently} \quad d\rho R^3 = -(p + \rho)d(R^3). \quad (67)$$

The first form simply says that the change in energy in a co-moving volume element, $d(\rho R^3)$, is equal to minus the pressure times the change in volume, $-pd(R^3)$.

- **Equation of state implications:**

One writes $p = w\rho$. In simple cases, w is independent of time. If so, then Eq. (67) reduces to $\frac{d\rho}{\rho} = -(1 + w)\frac{d(R^3)}{R^3}$, for which the solution is

$$\rho \propto R^{-3(1+w)}. \quad (68)$$

Important cases are:

1. Radiation Dominance (RD): $(p = \frac{1}{3}\rho)$ *i.e.* $w = 1/3 \Rightarrow \rho \propto R^{-4}$

2. Matter Dominance (MD): ($p = 0$) *i.e.* $w = 0 \Rightarrow \rho \propto R^{-3}$
3. Vacuum Energy Dominance (VD): ($T^\mu_\nu \propto g^\mu_\nu = \text{diag}(1, 1, 1, 1)$) for which $p = -\rho$, *i.e.* $w = -1 \Rightarrow \rho \propto \text{constant}$

At the earliest times after inflation the universe was radiation dominated. By radiation here, we mean any relativistic object, including relativistic matter as well as photons.

At intermediate times, earlier than the present epoch, the universe was matter dominated.

The final situation depends upon whether there is a cosmological constant or equivalent. If not, then the universe would remain matter dominated (or curvature dominated if $k \neq 1$). However, measurements suggest that something like a small cosmological constant is taking over slowly.

During inflation, the universe was dominated by vacuum energy (not the current vacuum energy, but something much larger).

- The Einstein equation:

The $0 - 0$ and $i - i$ components of

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\mathcal{R} = -8\pi GT_{\mu\nu} \quad (69)$$

read:

$$\frac{\dot{R}^2}{R^2} + \frac{k}{R^2} = \frac{8\pi G}{3}\rho, \quad (70)$$

$$2\frac{\ddot{R}}{R} + \frac{\dot{R}^2}{R^2} + \frac{k}{R^2} = -8\pi G\rho. \quad (71)$$

The first equation above is the “Friedmann Equation”. The difference between the above two equations yields:

$$\frac{\ddot{R}}{R} = -\frac{4\pi G}{3}(\rho + 3p). \quad (72)$$

Today, $\dot{R} \geq 0$. If in the past $\rho + 3p$ was always > 0 (radiation or matter) then \ddot{R} was always < 0 , and thus at some finite time in the past R must have been equal to 0. This time corresponds to the “big-bang” and is usually identified with $t = 0$.

Eqs. (67), (70) and (71) are not independent (being related by Bianchi identities). Usually, it is most convenient to take Eqs. (67) and (70) as the independent equations.

- Hubble parameter and related:

The Hubble parameter is $H \equiv \frac{\dot{R}}{R}$ in terms of which the Friedmann equation,

$$\frac{\dot{R}^2}{R^2} + \frac{k}{R^2} = \frac{8\pi G}{3}\rho, \quad (73)$$

can be rewritten in the form

$$\frac{k}{H^2 R^2} = \frac{\rho}{3H^2/8\pi G} - 1 \equiv \Omega - 1, \quad (74)$$

where

$$\Omega \equiv \frac{\rho}{\rho_c} \quad \text{with} \quad \rho_c = \frac{3H^2}{8\pi G}. \quad (75)$$

Since $H^2 R^2 \geq 0$, we have the correspondence

$$\begin{aligned} k = +1 &\Rightarrow \Omega > 1 && \text{closed} \\ k = 0 &\Rightarrow \Omega = 1 && \text{flat} \\ k = -1 &\Rightarrow \Omega < 1 && \text{open.} \end{aligned} \quad (76)$$

Now, it is important to note that H is not independent of time. We will denote the current value by H_0 . We will also use R_0 for the current scale factor $R(t_0)$ at current time $t = t_0$. As we saw earlier, $\frac{R_0}{R} \equiv 1 + z$, $R = R(t)$ being the scale factor at some earlier time, t . Of course, $1 + z$ is then the red-shift we see for light coming from some earlier time t from a distant galaxy. In the $\Omega \geq 1$ cases, it is conventional to define the zero of time to be that time when $R \rightarrow 0$ (and $z \rightarrow \infty$). Of course, technology and physics prevent us from

seeing all the way back to $t = 0$.

Now, $\frac{k}{H^2 R^2} = \Omega - 1$ is valid for all times. However, $\Omega - 1$ and $\rho_c = \frac{3H^2}{8\pi G}$ are not constant. They change as the Universe expands.

At early times, we will see that a self-consistent solution to the Friedmann equation will imply that the curvature term k/R^2 is negligible compared to $(\dot{R}/R)^2 = H^2$.

Assuming this for the moment, the Friedmann equation written as in Eq. (70) gives $\frac{\dot{R}^2}{R^2} \sim \frac{8\pi G}{3}\rho$ which gives $H^2 \propto R^{-3}$ in the (MD) case or $H^2 \propto R^{-4}$ in the (RD) case (see earlier summary), giving $k/(H^2 R^2) \propto R$ (MD), R^2 (RD).

Using the above scalings, at early times $\frac{k}{H^2 R^2} = \Omega - 1$ with $k = +1$ gives

$$|\Omega - 1| \sim \begin{cases} R/R_0 = (1+z)^{-1} & (MD) \\ (R_{EQ}/R_0)(R/R_{EQ})^2 \simeq 10^4(1+z)^{-2} & (RD) \end{cases} \quad (77)$$

where $R_{EQ} \simeq 10^{-4}R_0$ is the value of R at the transition between matter domination and radiation domination. At early times, $1+z \gg 1$, $\Omega \sim 1$ and the Universe was *very* nearly critical. Of course, the above equation can't be used for small z .

Note that since the above gives $\frac{k}{H^2 R^2} \ll 1$ at early times, this implies that $\frac{k}{R^2} \ll H^2$ at early times so that our original assumption of neglecting the $\frac{k}{R^2}$ curvature term compared to H^2 in the Friedmann equation at early times is self-consistent.

Another little detail we need to get straight is the relationship of k to actual spatial curvature.

You need to recall that we defined ${}^3\mathcal{R} = \frac{6k}{R^2(t)}$ earlier below Eq. (33). (There we did the calculation for $S(t)$, now called $R(t)$, $= 1$ and got ${}^3\mathcal{R} = 6k$ — clearly ${}^3\mathcal{R}$ scales as $1/R^2(t)$, see Eq. (55).) Using $\frac{k}{H^2 R^2} = \Omega - 1$ this can be re-expressed as ${}^3\mathcal{R} = 6H^2(\Omega - 1)$.

From the form of the RW metric, it is clear that the effects of

spatial curvature become very significant for $r \sim |k|^{-1/2}$, so we define a physical “radius of curvature” of the Universe, $R_{curv} \equiv R(t)|k|^{-1/2} = \left(\frac{6}{|{}^3\mathcal{R}|}\right)^{1/2}$ that is related to the Hubble radius, H^{-1} , by

$$R_{curv} = \frac{H^{-1}}{|\Omega - 1|^{1/2}}. \quad (78)$$

When $|\Omega - 1|$ is of order unity, R_{curv} and H^{-1} are comparable; when $|\Omega - 1|$ is very small, $R_{curv} \gg H^{-1}$. What this means is that if the Universe is close to critical density, then it has very large curvature, *i.e.* is very flat.

In particular, since $|\Omega - 1|$ must have been very small at early epochs (see above), it is safe to ignore spatial curvature in the early Universe.

Note that for the closed models, $k > 0$, R_{curv} is just the physical radius of the 3-sphere defined by $kr^2 = 1$ with actual radius $rR(t) = k^{-1/2}R(t)$.

Finally, we note while we have scaled r such that $k = +1, 0$ or -1 , there are an infinity of RW models, characterized by different k but *the same* physical curvature radius, $R_{curv} = R(t)|k|^{-1/2}$, at some specified epoch.

That is to say, all physics is contained in the ratio k/R^2 and not separately in k .

- **The Age of the Universe:**

We will now show how the Friedmann equation can be integrated to give the age of the Universe in terms of present cosmological parameters.

Given that

$$\rho/\rho_0 = (R/R_0)^{-3(1+w)} \sim R^{-3} \quad (MD), \quad R^{-4} \quad (RD), \quad (79)$$

the Friedmann equation, after multiplying by R^2/R_0^2 , can be rewritten

in the form

$$\left(\frac{\dot{R}}{R_0}\right)^2 + \frac{k}{R_0^2} = \frac{8\pi G}{3}\rho_0\frac{R_0}{R} = H_0^2\Omega_0\frac{R_0}{R} \quad (MD) \quad (80)$$

$$\left(\frac{\dot{R}}{R_0}\right)^2 + \frac{k}{R_0^2} = \frac{8\pi G}{3}\rho_0\left(\frac{R_0}{R}\right)^2 = H_0^2\Omega_0\left(\frac{R_0}{R}\right)^2 \quad (RD) \quad (81)$$

where the last forms used the fact that $(8\pi G/3)\rho_c = H^2$ (see Eq. (75)) applies at any time and in particular today so that $(8\pi G/3)\rho_0 = (H_0^2/\rho_{0c})\rho_0 = H_0^2\frac{\rho_0}{\rho_{0c}} = H_0^2\Omega_0$. Defining $x = R/R_0$ and recalling that $R_0/R = 1 + z$ we then obtain (assuming the Universe started with a big-bang at $R = 0$, or at least some very small value, in order to set the lower limit in the following)

$$\begin{aligned}
t &\equiv \int_0^{R(t)} \frac{dR'}{\dot{R}'} \\
&= H_0^{-1} \int_0^{(1+z)^{-1}} \frac{dx}{[1 - \Omega_0 + \Omega_0 x^{-1}]^{1/2}} \quad (MD) \\
&= H_0^{-1} \int_0^{(1+z)^{-1}} \frac{dx}{[1 - \Omega_0 + \Omega_0 x^{-2}]^{1/2}} \quad (RD). \quad (82)
\end{aligned}$$

If we now input the fact that currently $\Omega_0 \sim 1$ and use the (present) measured value of $H_0 \sim h_0 \times (9.77 \text{ Gyr})^{-1}$ with $h_0 \sim 0.7$, then for a matter dominated universe we would have

$$t \sim 9.77 \text{ Gyr} \times h_0^{-1} \frac{2}{3} (1+z)^{-3/2} \sim 9(1+z)^{-3/2} \text{ Gyr}. \quad (83)$$

Setting $z = 0$ to obtain the current age of the universe gives $t_0 \sim 9 \text{ Gyr}$ which is too small. The reason is that we need to include

the vacuum energy or cosmological constant. The result is that the present age is

$$t_0 \sim \frac{2}{3} H_0^{-1} \Omega_{vac}^{-1/2} \ln \left[\frac{1 + \Omega_{vac}^{1/2}}{(1 - \Omega_{vac})^{1/2}} \right], \quad (84)$$

where $\Omega_{vac} + \Omega_{matter} = 1$ was assumed. Plugging in $\Omega_{vac} = 0.75$ gives $t_0 = 14.1 \text{ Gyr}$, nicely consistent with what we think we know.

- **The deceleration parameter:**

As an aside, we now know that the deceleration parameter, called q_0 , defined by

$$q_0 \equiv - \left(\frac{\ddot{R}(t_0)}{R(t_0)} \right) \frac{1}{H_0^2}, \quad (85)$$

is quite interesting. By taking the ratio of $\frac{\ddot{R}}{R} = -\frac{4\pi G}{3}(\rho + 3p)$ to the Friedmann equation $\frac{\dot{R}^2}{R^2} + \frac{k}{R^2} = \frac{8\pi G}{3}\rho$ and using the definition of

Ω_0 , we find (neglecting k/R_0^2 , a good approximation at the current epoch)

$$q_0 = \frac{1}{2}\Omega_0 \left[1 + 3\frac{p}{\rho} \right] = \frac{1}{2}\Omega_0(1 + 3w). \quad (86)$$

This gives:

$$\begin{aligned} q_0 &= \frac{1}{2}\Omega_0 & (MD, w = 0) \\ q_0 &= \Omega_0 & (RD, w = 1/3) \\ q_0 &= -\Omega_0 & (VD, w = -1). \end{aligned} \quad (87)$$

In particular, we see the standard result that in a vacuum-dominated model the expansion is accelerating, $\ddot{R}_0 > 0$, since $\Omega_0 \sim 1 > 0$.

- **Equilibrium Thermodynamics:**

In terms of the phase space distribution function $f(\vec{p})$ we have:

$$n = \frac{g}{(2\pi)^3} \int f(\vec{p}) d^3p \quad (88)$$

$$\rho = \frac{g}{(2\pi)^3} \int E(\vec{p}) f(\vec{p}) d^3p \quad (89)$$

$$p = \frac{g}{(2\pi)^3} \int \frac{|\vec{p}|^2}{3E(\vec{p})} f(\vec{p}) d^3p, \quad (90)$$

where g is the number of internal degrees of freedom and $E^2(\vec{p}) = |\vec{p}|^2 + m^2$.

For a species in **kinetic equilibrium** the phase space occupancy is given by

$$f(\vec{p}) = \frac{1}{e^{(E-\mu)/T} \pm 1} \quad (91)$$

where μ is the chemical potential and $+1$ is for fermions and -1 is for bosons. If the species is also in chemical equilibrium then its μ is related to the chemical potentials of other species with which it interacts. For example, if chemical equilibrium holds for $i + j \leftrightarrow k + l$, we have

$$\mu_i + \mu_j = \mu_k + \mu_l \quad (92)$$

In the relativistic limit $T \gg m$ and for $T \gg \mu$ the integrals are simple and we find

$$n = \begin{cases} \left(\frac{\zeta(3)}{\pi^2}\right) gT^3 & (\text{bosons}) \\ \frac{3}{4} \left(\frac{\zeta(3)}{\pi^2}\right) gT^3 & (\text{fermions}) \end{cases} \quad (93)$$

$$\rho = \begin{cases} \left(\frac{\pi^2}{30}\right) gT^4 & (\text{bosons}) \\ \frac{7}{8} \left(\frac{\pi^2}{30}\right) gT^4 & (\text{fermions}) \end{cases} \quad (94)$$

$$p = \frac{\rho}{3}. \quad (95)$$

If there are any relativistic species then it is a good approximation to use only them since the contributions from non-relativistic species are very small (exponentially suppressed) in comparison. Thus, we

typically employ to good approximation

$$\rho_R = \frac{\pi^2}{30} g_* T^4, \quad p_R = \frac{\rho_R}{3}, \quad (96)$$

with

$$g_* = \sum_{i=\text{bosons}} g_i \left(\frac{T_i}{T} \right)^4 + \frac{7}{8} \sum_{i=\text{fermions}} g_i \left(\frac{T_i}{T} \right)^4. \quad (97)$$

Of course, once T falls below m_i we stop including particle i in the sum. We will find that $T_i = T$ for all particles except neutrinos.

During the radiation dominated epoch (roughly $t \lesssim 4 \times 10^{10}$ sec) we have $\rho = \frac{\pi^2}{30} g_* T^4$. Also recall that $\frac{\rho}{(3H^2/8\pi G)} = \Omega$, so that for $\Omega \simeq 1$ ($k \simeq 0$) we obtain

$$H = \left[\frac{8\pi G}{3} \rho \right]^{1/2} = \left[\frac{8\pi G \pi^2}{3 \cdot 30} g_* T^4 \right]^{1/2} = 1.66 \frac{g_*^{1/2}}{M_{\text{P}}} T^2, \quad (98)$$

where M_{P} is the Planck mass defined as $M_{\text{P}} = \sqrt{\frac{\hbar c}{G}} = 1.22 \times 10^{19} \frac{\text{GeV}}{c} = 1.22 \times 10^{22} \frac{\text{MeV}}{c}$.

For the radiation dominated Universe we found earlier that $R(t) \propto t^{1/2}$ with the consequence that $H \equiv \frac{\dot{R}}{R} = \frac{1}{2t}$. Plugging this in above and solving for t gives

$$t = 0.30 \frac{M_{\text{P}}}{g_*^{1/2} T^2} \sim \frac{2.4}{g_*^{1/2}} \left(\frac{1 \text{ MeV}}{T(\text{in MeV})} \right)^2 \text{ sec}, \quad (99)$$

where 1 MeV is a temperature that will frequently appear in our discussion.

Note: The above formula can't be used to compute the age of the universe since it is only valid for smallish times.

Units

Perhaps this is as good a time as any to make sure everyone has

the system of units under control. Everything we have written has assumed $\hbar = c = 1$.

Now $\hbar c = 197.3 \text{ MeV fm}$ where $1 \text{ fm} = 10^{-13} \text{ cm}$. Then, $\hbar c = 1$ implies $1 \text{ cm} = \frac{10^{13}}{197.3} \text{ MeV}^{-1}$.

Further $c = 1$ is equivalent to $3 \times 10^{10} \text{ cm} = 1 \text{ sec}$.

Combining, we get $1 \text{ sec} = \frac{3 \times 10^{23}}{197.3} \text{ MeV}^{-1}$.

Thus,

$$\begin{aligned} 0.3 \frac{M_{\text{P}}}{\text{MeV}^2} &= 0.3 \times 1.22 \times 10^{22} \text{ MeV}^{-1} \\ &= \frac{0.3 \times 1.22 \times 10^{22}}{\left(\frac{3 \times 10^{23}}{197.3}\right)} \text{ sec} \\ &= 2.4 \text{ sec} \end{aligned} \tag{100}$$

If you are unfamiliar with the $\hbar c = 197.3 \text{ MeV fm}$ equivalence, please look work it out for yourself. It is roughly saying that 0.2 GeV

(the typical energy scale associated with the proton bound state) is equivalent to $1/fm$ where the typical size of a proton is of order a fermi.

Some other useful conversion factors are the following: $1 K = 4.3668 \text{ cm}^{-1} = 8.6170 \cdot 10^{-14} \text{ GeV} = 1.5361 \cdot 10^{-37} \text{ g}$ (coming from the $k_B = 1$ convention implicit in our $f(\vec{p})$ formulae); $1 \text{ Mpc} = 1.5637 \cdot 10^{38} \text{ GeV}^{-1}$; $G = 6.7065 \cdot 10^{-39} \text{ GeV}^{-2}$; and $H_0 = h \times 2.1317 \cdot 10^{-42} \text{ GeV}$.

In particular, it is worth noting that the CMB temperature of $2.73 K$ is equivalent to $2.73 \times 8.6170 \cdot 10^{-14} \text{ GeV} \sim 2.35 \cdot 10^{-10} \text{ MeV}$.

Particle Counting

1. For $T \ll \text{MeV}$, only the 3 neutrino species (that we now know are very light) and the photon are relativistic.

Since $T_\nu = (4/11)^{1/3}T_\gamma$ (will discuss later),

$$g_*(\ll \text{ MeV}) = \frac{7}{8}(3 \times 2) \left(\frac{4}{11}\right)^{4/3} + 2 \simeq 3.36, \quad (101)$$

where we have taken account of the facts that the photon has 2 spin directions and that each neutrino has an anti-neutrino partner, but that each neutrino has only a left-handed spin direction in the SM (and each anti-neutrino a right-handed spin direction).

2. For $1 \text{ MeV} \lesssim T \lesssim 100 \text{ MeV}$, the electron and positron are relativistic, each having two spin directions, and $T_\nu = T_\gamma = T_e$, implying

$$g_* = \frac{7}{8}(3 \times 2) + 2 + \frac{7}{8}(2 \times 2) = 10.75. \quad (102)$$

3. For $T > 300 \text{ GeV}$, all the species in the standard model — neutrinos, photon, 8 gluons, W^\pm, Z , 3 generations of quarks (each with 3 colors and each having an anti-quark partner) and leptons (plus anti-leptons), and 1 spin-0 Higgs boson — should be

relativistic yielding

$$g_* = \text{neutrinos} + \text{photon} + \text{charged-leptons} + \text{gluons} + (W^\pm, Z) + \text{quarks} + \text{Higgs} \quad (103)$$

as given below:

$$g_* = \frac{7}{8}(3 \times 2) + 2 + \frac{7}{8}(3 \times 2 \times 2) + 8 \times 2 + 3 \times 3 + \frac{7}{8}(3 \times 3 \times 2 \times 2 \times 2) + 1 = 106.75. \quad (104)$$

For example, for the quarks there are 3 families, 3 colors, 2 for up+down, 2 for quark+antiquark and 2 for two spin directions.

You might ask why the gluons are not treated as massless. This is because they are confined and have an effective mass for counting purposes of order a GeV.

When we come to supersymmetry, there will be additional particles to account for, especially in the very early universe when T was very large.

Without supersymmetry the plot is

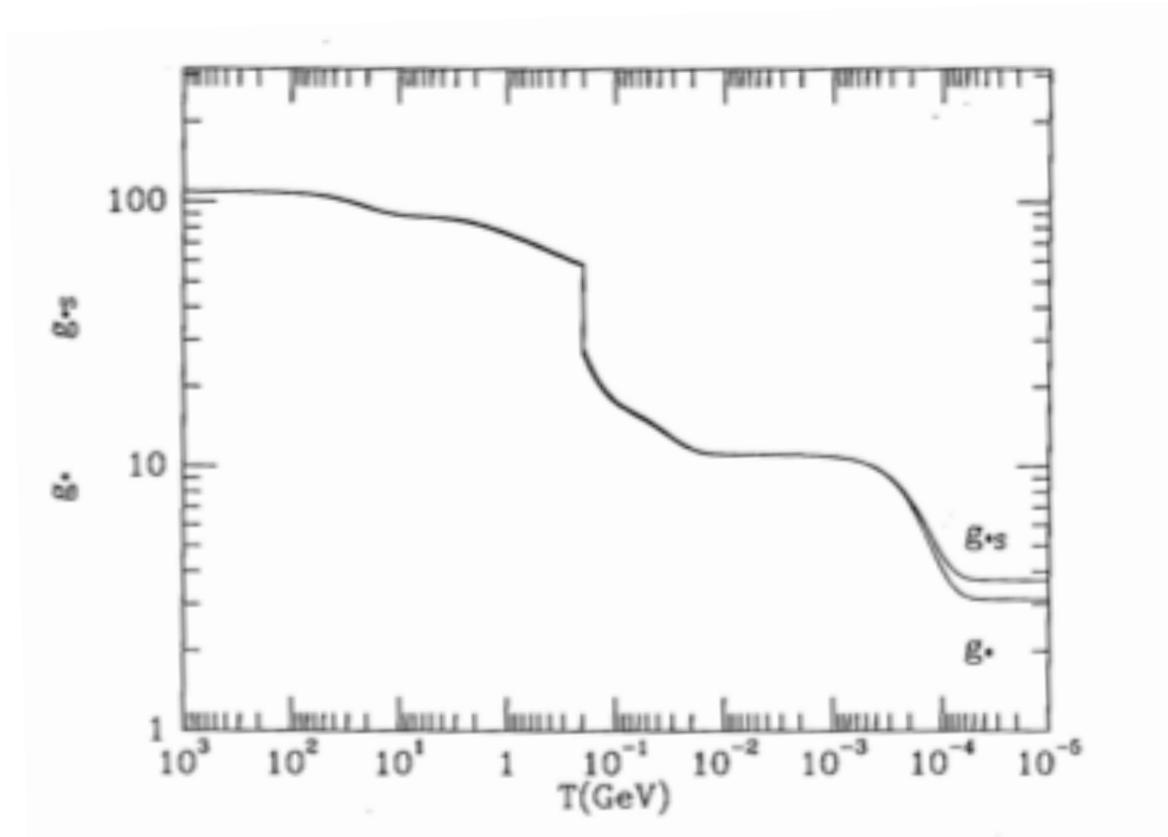


Figure 3: The evolution of g_* with T is shown below. We will discuss g_{*S} shortly.

- Entropy

For most of the history of the Universe (in particular at very early times) the reaction rates of particles in the thermal bath, Γ_{int} ,

were much greater than the expansion rate, H , and local thermal equilibrium should have applied.

In this case, the entropy per comoving volume element remains constant, and provides a very useful “fiducial” quantity during the expansion of the Universe.

For unit coordinate volume ($r = 1$) the physical volume is $V = R^3$. The 2nd law of thermodynamics states that (U is internal energy and W is work)

$$TdS = dU - dW = d(\rho V) + pdV = d[(\rho + p)V] - Vd\rho, \quad (105)$$

where ρ and p are the equilibrium energy density and pressure and are only function of T (see explicit forms given earlier).

Now, in general,

$$dS(V, T) = \left. \frac{\partial S(V, T)}{\partial V} \right|_T dV + \left. \frac{\partial S(V, T)}{\partial T} \right|_V dT. \quad (106)$$

Since ρ and T depend only on T , matching this to the $TdS = d(\rho V) + pdV$ form of the 2nd law given in Eq. (105) leads to

$$\left. \frac{\partial S}{\partial V} \right|_T = \frac{1}{T}(\rho + p), \quad \left. \frac{\partial S}{\partial T} \right|_V = \frac{V}{T} \frac{d\rho}{dT}. \quad (107)$$

The integrability condition

$$\frac{\partial^2 S}{\partial T \partial V} = \frac{\partial^2 S}{\partial V \partial T} \quad (108)$$

then takes the form

$$\left. \frac{\partial}{\partial T} \right|_V \left[\frac{1}{T}(\rho + p) \right] = \left. \frac{\partial}{\partial V} \right|_T \left[\frac{V}{T} \frac{d\rho}{dT} \right] \quad (109)$$

which implies

$$-\frac{1}{T^2}(\rho + p) + \frac{1}{T} \frac{d}{dT}(\rho + p) = \frac{1}{T} \frac{d\rho}{dT} \quad \Rightarrow \quad \frac{dp}{dT} = \frac{1}{T}(\rho + p). \quad (110)$$

This can also be obtained by direct computation using the forms of $p(T)$ and $\rho(T)$ given earlier. Inserting $dp = \frac{dT}{T}(\rho + p)$ into the 2nd law form $dS = \frac{1}{T}d[(\rho + p)V] - V dp$ gives

$$dS = \frac{1}{T}d[(\rho + p)V] - V(\rho + p)\frac{dT}{T^2} = d \left\{ \frac{V}{T}(\rho + p) + \text{const.} \right\}, \quad (111)$$

so that up to a constant

$$S(V, T) = \frac{V}{T}(\rho + p). \quad (112)$$

Returning to the 1st law $d[(\rho + p)V] - V dp = 0$ and inserting $dp = \frac{dT}{T}(\rho + p)$ as obtained above, gives, after removing an overall

factor of T ,

$$d \left[\frac{(\rho + p)V}{T} \right] = 0, \quad \Rightarrow dS = 0, \quad (113)$$

which is to say that (in thermal equilibrium) the entropy per comoving volume, which is what S is, is conserved.

Note: above we implicitly assumed $|\mu_k| \ll T$ for all particle species — a very good approximation.

We can also define the entropy density

$$s \equiv \frac{S}{V} = \frac{\rho + p}{T}. \quad (114)$$

Using the earlier derived results

$$\rho_R = \frac{\pi^2}{30} g_* T^4, \quad p_R = \frac{\rho_R}{3}, \quad (115)$$

we find

$$s = \frac{2\pi^2}{45} g_* s T^3 \quad (116)$$

where

$$g_* s = \sum_{i=bosons} g_i \left(\frac{T_i}{T}\right)^3 + \frac{7}{8} \sum_{i=fermions} g_i \left(\frac{T_i}{T}\right)^3. \quad (117)$$

For most of the history of the Universe all particle species had a common T and $g_* s = g_*$.

A useful normalization is the number of photons,

$$n_\gamma = \frac{\zeta(3)}{\pi^2} 2T^3. \quad (118)$$

Inserting this into the expression for s we find

$$s = \frac{\pi^4}{45\zeta(3)} g_* s n_\gamma \simeq 1.8 g_* s n_\gamma. \quad (119)$$

Today, $g_* s \simeq \frac{7}{8}(3 \times 2)\frac{4}{11} + 2 = \frac{43}{11}$ leading to $s \simeq 7.04n_\gamma$.

Now, conservation of S implies $sR^3 \sim \text{const}$. Combining with $s \propto g_* s T^3$ we obtain

$$g_* s T^3 R^3 = \text{const} \quad (120)$$

as the Universe expands. Note that this implies that

$$T \propto g_* s^{-1/3} R^{-1}, \quad (121)$$

yielding the familiar result that $T \propto R^{-1}$ whenever $g_* s$ is constant.

Of course, whenever a particle species becomes non-relativistic and “disappears” (from chemical equilibrium), its entropy is transferred to the other relativistic particles still present in the thermal plasma, causing T to decrease slightly less slowly.

It is useful to note that $s \propto R^{-3} \Rightarrow R^3 \propto s^{-1}$. Thus, the number of

some species in a comoving volume, $N = R^3 n$, is $\propto \frac{n}{s}$. In fact, it is conventional to define

$$N \equiv \frac{n}{s}. \quad (122)$$

For example, for a boson with $n = \left(\frac{\zeta(3)}{\pi^2}\right) g T^3$ (see earlier), using Eq. (116) one finds

$$N = \frac{45\zeta(3)g}{2\pi^4 g_* s}, \quad T \gg m, \mu. \quad (123)$$

- Decoupling of massless particles

The above does not apply to massless particles (massless with respect to the decoupling temperature). Massless particles, that are already decoupled from the heat bath (because they interact too weakly — defined in a moment — with other particles) will not share in the entropy transfer as T drops below the mass threshold for some process.

Instead, the T for the massless species scales as $T \propto R^{-1}$.

To see this, consider a massless particle species initially in LTE which decouples at time t_{dec} , temperature T_{dec} and scale factor R_{dec} . The phase space distribution **at decoupling** is

$$f(\vec{p}, t_{dec}) = \frac{1}{e^{(E/T_{dec})} \pm 1}. \quad (124)$$

After decoupling, the energy of each massless particle is red shifted (just like a photon) by the expansion of the Universe: $E(t) = E(t_{dec})[R(t_{dec})/R(t)]$.

But, since the particle species has decoupled there is no longer any “temperature” defined by thermal equilibrium and so T_{dec} remains as shown in the above formula.

In addition, the number density of particles decreases due to the expansion as $n \propto R^{-3}$.

These facts imply that in the the phase space distribution function, $f(\vec{p}) = d^3n/d^3p$, the factors of R^{-3} cancel out between numerator and denominator and $f(\vec{p}, t)$ at time t will be precisely that of a species in LTE with temperature $T(t) = T_{dec}R_{dec}/R(t)$:

$$f(\vec{p}, t) = f(\vec{p}_{dec}, t_{dec}) = f\left(\vec{p}\frac{R}{R_{dec}}, t_{dec}\right) \quad (125)$$

$$= \frac{1}{e^{(\frac{ER}{R_{dec}})/T_{dec}} \pm 1} = \frac{1}{e^{(E/T)} \pm 1}. \quad (126)$$

Thus, the f for a massless particle species remains self-similar as the universe expands, with the temperature red-shifting as R^{-1} :

$$T = T_{dec}\frac{R_{dec}}{R} \propto R^{-1}, \quad (127)$$

and not as $g_*^{-1/3}R^{-1}$.

- Neutrinos — an example of decoupling:

Let's now discuss the decoupling of neutrinos at about $T \sim 1$ MeV and how this leads to neutrinos currently having a lower temperature than the CMB photons.

First, the equilibrium reaction just before neutrino decoupling was $e^+ + e^- \leftrightarrow \nu_i + \bar{\nu}_i$. The cross section for neutrino production (via W exchange) is roughly (for $s \ll m_W^2$)

$$\sigma_{weak}(e^+ + e^- \rightarrow \nu_i + \bar{\nu}_i) \sim \left(\frac{g_{weak}^2}{4\pi} \right)^2 \frac{s}{m_W^4}. \quad (128)$$

Now $s \sim (\langle E \rangle)^2$ and $\langle E \rangle \sim 3T$, so that for $T \ll m_W$

$$\sigma_{weak}(e^+ + e^- \rightarrow \nu_i + \bar{\nu}_i) \sim \left(\frac{g_{weak}^2}{4\pi} \right)^2 \frac{T^2}{m_W^4}. \quad (129)$$

The interaction rate $\Gamma_{int} \equiv t_{collision}^{-1} = n\sigma v$, so with $n \sim T^3$ and

$v \sim 1$, we find

$$\Gamma_{int} \sim \frac{\alpha_{weak}^2 T^5}{m_W^4} T^5 \sim G_F^2 T^5, \quad (130)$$

where $G_F = 1.1664 \times 10^{-5} \text{ GeV}^{-2}$ is the Fermi constant.

This interaction rate should be compared to the expansion rate $H \sim g_*^{1/2} T^2 / M_P$ [see Eq. (98)]:

$$\frac{\Gamma_{int}}{H} \sim \frac{G_F^2 T^5}{g_*^{1/2} T^2 / M_P} \sim \left(\frac{T}{0.7 \text{ MeV}} \right)^3. \quad (131)$$

Thus, at $T \lesssim 1 \text{ MeV}$ the interactions are too slow to provide an equilibrium between leptons and neutrinos. Neutrinos decouple from the other SM particles and evolve separately. As we have seen above, the neutrinos will then have a different temperature as compared to

other particles, in particular the photons. For the neutrinos we have:

$$T_\nu = T_{dec} \frac{R_{dec}}{R} \sim \frac{1}{R}. \quad (132)$$

Now, for those species still in equilibrium,

$$g_{*S}(RT)^3 = \text{const.} \Rightarrow T \sim (g_{*S})^{-1/3} \frac{1}{R} \quad (133)$$

The implication of the above is that the neutrino distribution will be the same as if the neutrinos were still in thermal equilibrium with the photons **as long as g_{*S} does not change.**

However, slightly below the \sim MeV temperature at which the neutrinos decouple, the electrons become non-relativistic, *i.e.* at $T \sim m_e \simeq 0.5$ MeV, and they annihilate via $e^+e^- \rightarrow \gamma\gamma$ (the inverse process being suppressed as the average energy decreases

below roughly $2m_e$). As a result, the number of relativistic degrees of freedom drops:

- for $T \gtrsim 2m_e$, the number of particle species in equilibrium with photons include the photon ($g = 2$) and e^\pm pairs ($g = 4$) for a value of

$$g_* = \sum_{\text{bosons}} g_i + \frac{7}{8} \sum_{\text{fermions}} g_i = 2 + \frac{7}{8} \times 4 = \frac{11}{2}; \quad (134)$$

- for $T \ll 2m_e$, the electrons and positrons have annihilated and only the photons are in equilibrium with themselves, yielding

$$g_* = 2. \quad (135)$$

For the particles in equilibrium with the photons, $g_*(RT)^3$ (where $T = T_\gamma$) remains constant during expansion. This implies that the entropy that resided in the e^\pm for $T \gtrsim 2m_e$ must be transferred to

the photons when T falls below $2m_e$ and g_* decreases.

We have (below “before” refers to before e^\pm annihilation and “after” refers to after e^\pm annihilation)

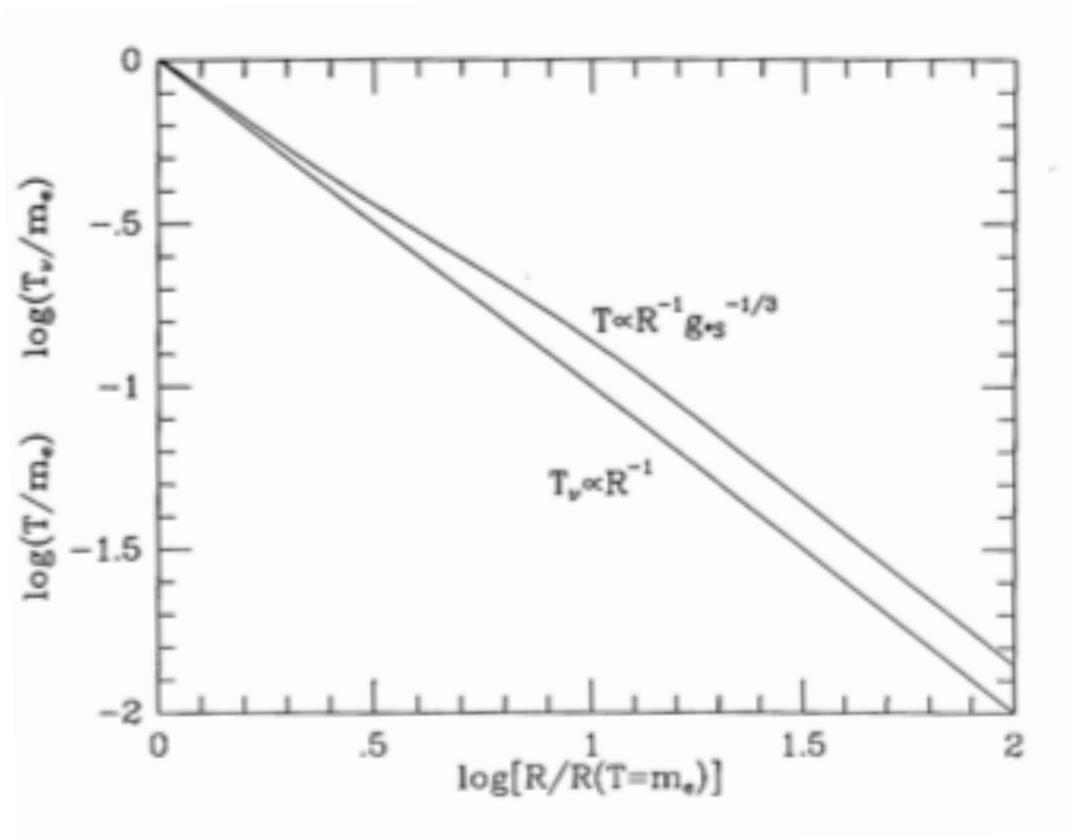
$$[g_*(RT)^3]_{before} = [g_*(RT)^3]_{after} \Rightarrow \frac{11}{2}(RT)_{before}^3 = 2(RT)_{after}^3. \quad (136)$$

Thus, the e^\pm entropy transfer increases (RT_γ) by a factor of $(11/4)^{1/3}$, while RT_ν remains constant since the neutrinos had already decoupled.

$$\Rightarrow \frac{T}{T_\nu} = \left(\frac{11}{4}\right)^{1/3} = 1.40 \quad (137)$$

just after annihilation and this ratio continues until today. Using the currently observed CMB temperature $T = T_\gamma = 2.73 \text{ K}$, we get $T_\nu = 1.95 \text{ K}$. This will create a difference between g_* and the g_{*S} that includes the entropy contribution of the neutrinos — g_{*S} must

increase relative to g_* since the entropy should not change as T is decreased. Of course, in actuality, there is no sudden jump. The decrease in g_* occurs in a continuous fashion and does not lead to an actual increase in T , but rather causes T to decrease less slowly than R^{-1} . A plot is below (the $g_* s$ indicated there includes entropy in the neutrinos)



Using the above temperatures, we find (assuming 3 neutrino species)

$$\begin{aligned}g_*(today) &= 2 + \frac{7}{8} \times 2 \times 3 \times \left(\frac{4}{11}\right)^{4/3} = 3.36, \\g_{*S}(today) &= 2 + \frac{7}{8} \times 2 \times 3 \times \left(\frac{4}{11}\right) = 3.91. \quad (138)\end{aligned}$$

We emphasize that since the photon and neutrinos species are decoupled, their entropies are separately conserved (which fact we used implicitly above).

Using the above results, we can compute the present energy density and entropy density of the photons and neutrinos:

	γ	ν
$\rho = \frac{\pi^2}{30} g_* T^4$	$4.67 \cdot 10^{-34} \text{ g cm}^{-3}$	$3.18 \cdot 10^{-34} \text{ g cm}^{-3}$
$s = \frac{2\pi^2}{45} g_* s T^3$	1486 cm^{-3}	1419 cm^{-3}
$n = \frac{\zeta(3)}{\pi^2} g_n T^3$	413 cm^{-3}	338 cm^{-3}
$\Omega h^2 = \rho \frac{8\pi G}{3(H_0/h)^2}$	$2.49 \cdot 10^{-5}$	$1.70 \cdot 10^{-5}$

(139)

where I have defined

$$\begin{aligned}
 g_n &= \sum_{\text{bosons}} \left(\frac{T_i}{T} \right)^3 + \sum_{\text{fermions}} \frac{3}{4} \left(\frac{T_i}{T} \right)^3 \\
 &= 2 \text{ photons} \\
 &= \frac{3}{4} \times 2 \times 3 \times \left(\frac{4}{11} \right) \text{ neutrinos.}
 \end{aligned}
 \tag{140}$$

Additional remarks

Notice that the above situation applies whenever a massive particle decouples when it is still relativistic, $T_{dec} \gg m$. That is the distribution function is “frozen” in the form of the equilibrium distribution function f_{eq} of massless particles.

These decoupled massive particles will eventually become non-relativistic when the temperature of the thermal bath drops below their mass, $T < m$, and their energy will then be $E \simeq m$.

The distribution function and number density of the particles will still be given by the frozen-in form corresponding to relativistic particles, but the energy density will be that of non-relativistic particles, $\rho = nm$.

This is exactly what happens for massive neutrinos.

If the particles decouple when they are already non-relativistic, then

$$\begin{aligned} f_{t>t_{dec}}(E) &= f_{eq}(ER/R_{dec}, T_{dec}) \\ &\simeq \frac{g}{(2\pi)^3} \exp\left[-\frac{(m-\mu)}{T_{dec}}\right] \\ &\quad \times \exp\left[-\frac{p^2}{2mT_{dec}} \left(\frac{R}{R_{dec}}\right)^2\right]. \end{aligned} \quad (141)$$

That is to say, the distribution function has the same form as that of a non-relativistic Maxwell-Boltzmann gas with a temperature given by $T = T_{dec}(R_{dec}/R)^2$ and chemical potential given by $\mu(t) = m + (\mu_{dec} - m)(T/T_{dec})$.

The Expanding Universe, Non-Equilibrium processes and the Boltzmann Equation

- Thermal equilibrium applied for much of the early history of the Universe.
- But, there have been crucial and very interesting departures
 1. neutrino decoupling
 2. decoupling of the background radiation
 3. primordial nucleosynthesis
 4. inflation?
 5. baryogenesis?
 6. relic WIMPs?
 7. axions?

If it were not for such departures from LTE, the present state of the

Universe would be completely specified by the present T .

The decouplings are interesting because they leave behind relics.

Our focus = relic WIMPS and related

- Once decoupled $n \propto R^{-3}$ and $E, p \propto R^{-1}$.

It is the evolution of particle distributions around the epoch of decoupling that is particularly challenging but also particularly interesting

Recall again the relations:

$$\Gamma_{int} \gtrsim H = \text{coupled}, \quad \Gamma_{int} \lesssim H = \text{decoupled}. \quad (142)$$

But to properly predict relic abundances as a function of the interactions and properties of the relevant particles requires a microscopic treatment using

- the Boltzmann Equation:

$$\hat{\mathbf{L}}[f] = \mathbf{C}[f], \quad (143)$$

where \mathbf{C} is the collision operator and $\hat{\mathbf{L}}$ is the Liouville operator. The covariant, relativistic version of the Liouville operator is

$$\hat{\mathbf{L}} = p^\alpha \frac{\partial}{\partial x^\alpha} - \Gamma_{\beta\gamma}^\alpha p^\beta p^\gamma \frac{\partial}{\partial p^\alpha}. \quad (144)$$

Specializing to the RW metric and using the fact that f is spatially homogeneous and isotropic, $f = f(|\vec{p}|, t)$, the $\alpha = 0$ component (the only non-trivial component) becomes

$$\hat{\mathbf{L}}[f] = E \frac{\partial f}{\partial t} - \frac{\dot{R}}{R} |\vec{p}|^2 \frac{\partial f}{\partial E}. \quad (145)$$

We can now write the Boltzmann equation in terms of the number

density

$$n = 4\pi \int dp p^2 \frac{g}{(2\pi)^3} f(E, t) \quad (146)$$

by integrating the preceding equation over the momentum. Defining $g_\pi = \frac{g}{(2\pi)^3}$, we find

$$g_\pi 4\pi \int dp p^2 \frac{\hat{\mathbf{L}}[f]}{E} = \frac{dn}{dt} - g_\pi H 4\pi \int dp \frac{p^4}{E} \frac{\partial f}{\partial E} \quad (147)$$

$$= \frac{dn}{dt} - g_\pi H 4\pi \int dE E \frac{p^3}{E} \frac{\partial f}{\partial E} \quad (148)$$

$$= \frac{dn}{dt} + g_\pi H 4\pi \int dE \frac{\partial(p^3)}{\partial E} f \quad (149)$$

$$= \frac{dn}{dt} + g_\pi H 4\pi \int dE 3p^2 \frac{\partial p}{\partial E} f \quad (150)$$

$$= \dot{n} + H g_\pi 4\pi \int dE 3p^2 \frac{E}{p} f \quad (151)$$

$$= \dot{n} + g_\pi H 4\pi \int 3dp p^2 f \quad (152)$$

$$= \dot{n} + 3Hn, \quad (153)$$

where we identified $H = \frac{\dot{R}}{R}$ and used $m^2 = E^2 - p^2$ so that $pdp = EdE$ (used in several places) and integrated by parts.

Notice that in the absence of interactions, the Boltzmann equation would reduce to:

$$\dot{n} + 3Hn = 0, \quad (154)$$

which is just the conservation of particles per comoving volume, $d(R^3n)/dt = 0$, which would apply when there are no interactions or decays. The $3H$ term takes care of the dilution that comes from the Hubble expansion.

For very massive particles (so $\rho = mn$, $p = 0$), the equation is just the energy conservation equation.

In summary, the lhs of the integrated Boltzmann's equation is simply

$$\frac{g}{(2\pi)^3} 4\pi \int dp p^2 \frac{\hat{\mathbf{L}}[f]}{E} = \frac{1}{R^3} \frac{d}{dt} (nR^3), \quad (155)$$

which is the rate of change of the number of particles per comoving volume divided by volume.

- **Inputting Collisions:**

On the rhs of the equation, we have

$$\frac{g}{(2\pi)^3} \int \frac{d^3p}{E} \mathbf{C}[f]. \quad (156)$$

For a general collision operator $\mathbf{C}[f]$ we get a very complicated set of coupled equations: n above would refer to that for any given particle and $\int \frac{d^3p}{E} \mathbf{C}[f]$ would involve it and all other particles.

In practice, we will be interested in focusing on a given particle, let us say a WIMP, denoted χ , and the collision operator will be well approximated by keeping only the χ and a limited set of particles that remain in equilibrium with one another (only the χ is going out of equilibrium).

It is also particularly relevant to consider two particle into two particle processes, say $\chi + A \rightarrow B + C$. The general case gets rather messy.

Focusing on the χ , we want the rhs of our Boltzmann equation to correspond to minus the rate of loss of χ 's plus the rate of production of χ 's, where the rate corresponds to the number of interactions per unit volume per unit time.

Now, simplifying our notation to $12 \rightarrow 34$, reaction rates for $1 + 2 \rightarrow 3 + 4$ are related to the corresponding cross section by (see p. 100, Peskin and Schroeder)

$$rate(12 \rightarrow 34) = n_1 n_2 |v_1 - v_2| \sigma_{12 \rightarrow 34} \quad (157)$$

where n_1 and n_2 are the number densities of particles 1 and 2; in our case $1 = \chi$, $2 = A$, $3 = B$ and $4 = C$.

Meanwhile (see Eq. (4.79) of PS),

$$\sigma_{12 \rightarrow 34} = \frac{1}{2E_1 2E_2 |v_1 - v_2|} \int d\Pi_3 d\Pi_4 |\mathcal{M}|_{12 \rightarrow 34}^2 (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4) \quad (158)$$

with $d\Pi = \frac{g}{(2\pi)^3} \frac{d^3 p}{2E}$, where g counts the degrees of freedom of the particle and $|\mathcal{M}|^2$ is defined in the convention that it is averaged over the degrees of freedom of all particles involved.

In our case, the number densities will be differentially defined as for example $n_1 = \int dn_1 = \int \frac{g_1}{(2\pi)^3} d^3 p_1 f_1$. Putting the above together gives

$$rate(12 \rightarrow 34) = \int d\Pi_1 d\Pi_2 d\Pi_3 d\Pi_4 (2\pi)^4 \delta^4(\dots) |\mathcal{M}|_{12 \rightarrow 34}^2 f_1 f_2. \quad (159)$$

Note how the $2E_1$ and $2E_2$ factors in the denominator of the cross section form neatly combined with the d^3p_1 and d^3p_2 to give the $\frac{d^3p_1}{2E_1}$ and $\frac{d^3p_2}{2E_2}$ factors that are part of $d\Pi_1$ and $d\Pi_2$. And, of course, the $\frac{g_{1,2}}{(2\pi)^3}$ that were part of $dn_{1,2}$ also work nicely to give the rest of the $d\Pi_{1,2}$.

The net result is that the rhs of the Boltzmann equation for n_1 is

$$\begin{aligned}
 & -rate(12 \rightarrow 34) + rate(34 \rightarrow 12) \\
 & = - \int d\Pi_1 d\Pi_2 d\Pi_3 d\Pi_4 (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4) \\
 & \quad \times [|\mathcal{M}|_{12 \rightarrow 34}^2 f_1 f_2 - |\mathcal{M}|_{34 \rightarrow 12}^2 f_3 f_4] . \quad (160)
 \end{aligned}$$

Now, T invariance (equivalently CP invariance) implies

$$|\mathcal{M}|_{12 \rightarrow 34}^2 = |\mathcal{M}|_{34 \rightarrow 12}^2 \quad (161)$$

so that we can rewrite Eq. (160) in the form

$$\begin{aligned}
 & -rate(12 \rightarrow 34) + rate(34 \rightarrow 12) \\
 & = - \int d\Pi_1 d\Pi_2 d\Pi_3 d\Pi_4 (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4) \\
 & \quad \times |\mathcal{M}|_{12 \rightarrow 34}^2 [f_1 f_2 - f_3 f_4] . \tag{162}
 \end{aligned}$$

In the above derivations, we have neglected any Fermi blocking or similar effects. The invariant matrix elements squared are averaged over initial and final spins, and include the appropriate symmetry factors for identical particles in the initial or final states.

There is an additional subtlety. The expressions for the cross sections given above only apply in a collinear frame where \vec{v}_1 and \vec{v}_2 are parallel. In thermal averaging, this is usually not the case. What really should enter everywhere is $|v_{Mol}|$ in place of $|v_1 - v_2|$. We will discuss v_{Mol} shortly.

In addition, we will use the simplification of assuming that we can use Maxwell-Boltzmann statistics for all species. This results in the further simplification that in equilibrium

$$f_3^{eq} f_4^{eq} = K \exp[-(E_3 + E_4)/T] = K \exp[-(E_1 + E_2)/T] = f_1^{eq} f_2^{eq}. \quad (163)$$

Of course, in our situation we are going to assume that 3 and 4 remain in equilibrium but that 1, and usually 2, drops out of equilibrium.

With these inputs, we arrive at the Boltzmann equation form

$$\begin{aligned} \dot{n}_1 + 3Hn_1 &= - \int d\Pi_1 d\Pi_2 d\Pi_3 d\Pi_4 (2\pi)^4 \delta^4(p_1 + p_2 - p_3 - p_4) \\ &\quad \times |\mathcal{M}|_{12 \rightarrow 34}^2 [f_1 f_2 - f_1^{eq} f_2^{eq}] \\ &= - \int \frac{g_1}{(2\pi)^3} d^3 p_1 \frac{g_2}{(2\pi)^3} d^3 p_2 (|v_{Mol}| \sigma_{12 \rightarrow 34}) [f_1 f_2 - f_1^{eq} f_2^{eq}] \\ &\equiv - \langle |v_{Mol}| \sigma_{12 \rightarrow 34} \rangle [n_1 n_2 - n_1^{eq} n_2^{eq}], \end{aligned} \quad (164)$$

where the 2nd equality simply results from the original definition of the cross section back in Eq. (158) (with $|v_1 - v_2| \rightarrow |v_{Mol}|$) and

the final identity defines the thermally averaged cross section times velocity.

Cross Sections

Let us return to the basic Peskin and Schroeder, Eq. (4.79) for parallel \vec{v}_1 and \vec{v}_2 . For $1 + 2 \rightarrow \{f\}$,

$$d\sigma = \frac{1}{2E_1 2E_2 |v_1 - v_2|} \left(\prod_f \frac{d^3 p_f}{(2\pi)^3 2E_f} \right) \times |\mathcal{M}(p_1, p_2 \rightarrow \{p_f\})|^2 (2\pi)^4 \delta^4(p_1 + p_2 - \sum_f p_f). \quad (165)$$

The above form for the cross section assumes that 1 and 2 are colliding collinearly with 1 travelling in one direction and 2 in the opposite direction (or one can be rest). A covariant way of writing

this prefactor is to use

$$2E_1 2E_2 |v_1 - v_2| = 4 \sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2}. \quad (166)$$

In a non-collinear situation

$$\frac{\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2}}{E_1 E_2} \equiv |v_{Mol}| = [|\vec{v}_1 - \vec{v}_2|^2 - |\vec{v}_1 \times \vec{v}_2|^2]^{1/2}, \quad (167)$$

where v_{Mol} is called the Moller velocity. Sometimes, the distinction between $|v_1 - v_2|$ and $|v_{Mol}|$ is ignored in the literature, at least for pedagogical purposes (*e.g.* in Kolb and Turner). Since $d\sigma$ is a relativistic covariant (by the way it is defined), one should use the relativistically covariant form of the prefactor. After all, in the process of thermal averaging not all momenta of the colliding dark matter particles are collinear. And, one does find that this difference is important numerically in some cases.

Anyway, the fully covariant cross section form is

$$d\sigma = \frac{1}{4\sqrt{(p_1 \cdot p_2)^2 - m_1^2 m_2^2}} \left(\prod_f \frac{d^3 p_f}{(2\pi)^3 2E_f} \right) \times |\mathcal{M}(p_1, p_2 \rightarrow \{p_f\})|^2 (2\pi)^4 \delta^4(p_1 + p_2 - \sum_f p_f). \quad (168)$$

The procedure for doing the proper thermal averaging was developed by Gondolo and Gelmini (Nucl. Phys. B360, p. 145). I sketch it below. By definition

$$\langle \sigma |v_{Mol}| \rangle = \frac{\int \sigma |v_{Mol}| e^{-E_1/T} e^{-E_2/T} d^3 p_1 d^3 p_2}{\int e^{-E_1/T} e^{-E_2/T} d^3 p_1 d^3 p_2}. \quad (169)$$

Writing

$$d^3 p_1 d^3 p_2 = 4\pi p_1 E_1 dE_1 4\pi p_2 E_2 dE_2 \frac{1}{2} d \cos \theta \quad (170)$$

and then changing variables to (for simplicity, assume $m_1 = m_2 = m$)

$$E_+ = E_1 + E_2, \quad E_- = E_1 - E_2, \quad s = 2m^2 + 2E_1E_2 - 2p_1p_2 \cos \theta \quad (171)$$

yields

$$d^3p_1 d^3p_2 = 4\pi^2 E_1 E_2 dE_+ dE_- ds. \quad (172)$$

In terms of these new variables, the integration region ($E_1 > m, E_2 > m, |\cos \theta| \leq 1$) transforms into

$$|E_-| \leq \sqrt{1 - \frac{4m^2}{s}} \sqrt{E_+^2 - s}, \quad E_+ \geq \sqrt{s}, \quad s \geq 4m^2, \quad (173)$$

and

$$|v_{Mol}| E_1 E_2 = \sqrt{(p_1 \cdot p_2)^2 - m^4} = \frac{1}{2} \sqrt{s(s - 4m^2)}. \quad (174)$$

is a function of s only. The numerator is then computed as

$$\begin{aligned}
& \int \sigma |v_{Mol}| e^{-E_1/T} e^{-E_2/T} d^3 p_1 d^3 p_2 \\
&= 2\pi^2 \int dE_+ \int dE_- \int ds \sigma |v_{Mol}| E_1 E_2 e^{-E_+/T} \\
&= 4\pi^2 \int_{4m^2}^{\infty} ds \sigma \frac{1}{2} \sqrt{s(s-4m^2)} \sqrt{1 - \frac{4m^2}{s}} \int_{\sqrt{s}}^{\infty} dE_+ e^{-E_+/T} \sqrt{E_+^2 - s} \\
&= 2\pi^2 T \int_{4m^2}^{\infty} ds \sigma (s-4m^2) \sqrt{s} K_1(\sqrt{s}/T). \tag{175}
\end{aligned}$$

Meanwhile, the denominator is

$$\int e^{-E_1/T} d^3 p_1 \int e^{-E_2/T} d^3 p_2 = [4\pi m^2 T K_2(m/T)]^2. \tag{176}$$

The K_i are the modified Bessel functions of order i .

Of course, the cross section itself has a prefactor proportional to $1/|v_{Mol}|$ so that $\sigma |v_{Mol}|$ will behave like $|\mathcal{M}(p_1, p_2 \rightarrow \{p_f\})|^2$. For

example, in the $1 + 2 \rightarrow 3 + 4$ case, we have

$$\frac{d\sigma}{dt} = \frac{1}{64\pi s} \frac{1}{p_{1\text{ cm}}^2} |\mathcal{M}|^2 \quad (177)$$

where $s = (p_1 + p_2)^2$ and $t = (p_1 - p_3)^2$ are the usual Mandelstam invariants. In the cm frame, one can write

$$dt = -2p_{1\text{ cm}}p_{3\text{ cm}}d\cos\theta_{\text{cm}} \quad (178)$$

so that

$$\frac{d\sigma}{d\Omega_{\text{cm}}} = \frac{1}{64\pi^2 s} \frac{p_{3\text{ cm}}}{p_{1\text{ cm}}} |\mathcal{M}|^2 = \frac{1}{64\pi^2 s} \frac{\beta_f^{\text{cm}}}{\beta_i^{\text{cm}}} |\mathcal{M}|^2, \quad (179)$$

where

$$\beta_f = \left[1 - (m_3 + m_4)^2/s\right]^{1/2} \left[1 - (m_3 - m_4)^2/s\right]^{1/2} \quad (180)$$

$$\beta_i = \left[1 - (m_1 + m_2)^2/s\right]^{1/2} \left[1 - (m_1 - m_2)^2/s\right]^{1/2}, \quad (181)$$

the latter reducing to $\beta_i = \sqrt{1 - 4m^2/s}$ for $m_1 = m_2 = m$, are entirely expressed in terms of the relativistic invariant s , as will be $|\mathcal{M}|^2$. Using the $m_1 = m_2 = m$ form of β_i , Eq. (175) reduces to

$$\begin{aligned} & \int \sigma |v_{Mol}| e^{-E_1/T} e^{-E_2/T} d^3 p_1 d^3 p_2 \\ &= 2\pi^2 T \int_{4m^2}^{\infty} ds \frac{1}{64\pi^2 s} \beta_f \int d\Omega_{cm} |\mathcal{M}|^2 s \sqrt{s - 4m^2} K_1(\sqrt{s}/T) \end{aligned} \quad (182)$$

Ultimately, what will be really important is the behavior of $|\mathcal{M}|^2$. In the absence of a Sommerfeld enhancement effect (which could introduce an extra $1/|v_{Mol}|$)² it will behave as $|v_{Mol}|^p$, with $p = 0$ for S-wave annihilation, $p = 2$ for P-wave annihilation, and so forth.

²For example, Sommerfeld enhancement can arise from the exchange of a very light gauge boson in the t -channel between the dark matter particles prior to their annihilation.

So, now let us consider a particular case.

$$\chi + \bar{\chi} \leftrightarrow B + \bar{B}$$

Referring back to Eq. (164) we have in the initial state $1 = \chi$ and $2 = \bar{\chi}$ and in the final state $3 = B$ and $4 = \bar{B}$. At this point we then have a Boltzmann equation that reads (using $v \equiv |v_{Mol}^{\chi\bar{\chi}}|$ and assuming $n_\chi = n_{\bar{\chi}}$)

$$\dot{n}_\chi + 3Hn_\chi = -\langle \sigma_{\chi\bar{\chi} \rightarrow B\bar{B}} |v| \rangle [n_\chi^2 - (n_\chi^{eq})^2], \quad (183)$$

Before proceeding further, as an aside let me sketch a **simpler treatment**. We could have said that physically

$$\dot{n}_\chi + 3Hn_\chi = -R(\chi\bar{\chi} \rightarrow B\bar{B})n_\chi n_{\bar{\chi}} + R(B\bar{B} \rightarrow \chi\bar{\chi})n_B n_{\bar{B}}. \quad (184)$$

where the R 's are appropriate rate factors.

In equilibrium, the rhs must be zero to have no net change in particle numbers, implying that

$$R(\chi\bar{\chi} \rightarrow B\bar{B})n_{\chi}^{eq}n_{\bar{\chi}}^{eq} = R(B\bar{B} \rightarrow \chi\bar{\chi})n_B^{eq}n_{\bar{B}}^{eq}. \quad (185)$$

This relation is called **detailed balance**. It gives the same net effect as the energy conservation plus time reversal gave in the more detailed approach above.

We then appeal to the physical argument that $R(\chi\bar{\chi} \rightarrow B\bar{B})$ would have to be given (dimensionally at any rate) by $\langle \sigma_{\chi\bar{\chi} \rightarrow B\bar{B}} |v| \rangle$.

Anyway, let us now continue with the development of the formalism.

The structure developed above generalizes in a very natural way to $\chi + A \rightarrow F$ where A and F are *systems* of particles. The general Boltzmann equation for this case looks like:

$$\dot{n}_{\chi} + 3Hn_{\chi} = - \int d\Pi_{\chi} d\Pi_A d\Pi_F |\mathcal{M}|_{\chi+A \rightarrow F}^2 (2\pi)^4 \delta^4(p_{\chi} + p_A - p_F) [f_{\chi} f_A - f_F], \quad (186)$$

where now

$$\begin{aligned}d\Pi_A &= \prod_{i=\text{other initial state particles}} d\Pi_i \\d\Pi_F &= \prod_{k=\text{all final state particles}} d\Pi_k \\f_A &= \prod_{i=\text{other initial state particles}} f_i \\f_F &= \prod_{k=\text{all final state particles}} f_k.\end{aligned}\tag{187}$$

Again, we had to use time reversal symmetry and assume absence of Fermi blocking factors to get $[f_\chi f_A - f_F]$ to factor. And, once again, if we use Maxwell Boltzmann statistics and energy conservation $f_F^{eq} = f_\chi^{eq} f_A^{eq}$.

$$\chi + A \leftrightarrow B$$

This is the case where χ and A can collide to create an on-shell stable particle or narrow resonance B . I will not give details here, but simply summarize results for this case.

Here, A denotes collectively any set of particles other than the χ . In

this case we have

$$\dot{n}_\chi + 3Hn_\chi = -R(\chi A \rightarrow B)n_\chi n_A + R(B \rightarrow \chi A)n_B \quad (188)$$

where the $R(\dots)$ are forward and backward rate coefficients. For the first term, if A is a single particle (the usual case) $R = \langle \sigma_{\chi+A \rightarrow B} |v| \rangle$. The second term only depends on the number of B particles around and $\Gamma(B \rightarrow \chi A)$ is the spontaneous decay rate.

Again, we have assumed occupation numbers are low (*e.g.* all massive particles), so that there are no fermi blocking or bose enhancement effects. In this approximation, the rate of producing B is independent of the existing n_B .

Again, we see from the above equation that for $\Gamma \ll H$ the collision term is small compared to the Hubble expansion term, which means that the system will go out of equilibrium.

- Calculation of the relic abundance:

Let us consider further the important case of $\chi + \bar{\chi} \leftrightarrow B + \bar{B}$, where B is a single particle, but where in general we must sum over all B 's.

For any species of particles which is not being created or destroyed, $n \propto R^{-3}$, and we can assign a conserved number $Y \propto nR^3$. Since $s \propto R^{-3}$, we can define this number to be

$$Y \equiv n/s. \quad (189)$$

Scaling of entropy

We will be discussing the decoupling of χ from other species that remain in equilibrium. Above, we made the statement that entropy scales like R^{-3} . This assumes the preceding statement that so long as a species of particle is not being created or destroyed, it is included in $n \propto R^{-3}$ and will be counted in the entropy.

The DM annihilation will simply shut off and will not heat up the

photons and other things that remain in thermal equilibrium.

The effective g_{*S} for the things that remain in equilibrium will drop, but the entropy WILL NOT drop. One must continue to count in the DM objects that are still present (they did not annihilate away).

The point is that there is still simply the combinatorial entropy of the "heavy billiard balls" that the DM is. It is given by the number density of DM particles, *i.e.* their energy density divided by their mass — if you like, one bit per particle, roughly.

The combinatorial entropy from the DM particles still scales as $1/R$ and so it seems to scale as T . And, it is normalized to $g_{DM}T^3$ because this is what it was BEFORE decoupling.

In other words, both before and after the DM has decoupled it is correct to write

$$g_{*S} = g_{*S}^{\text{non-DM equilibrium stuff}} + g_{*S}^{DM}, \quad (190)$$

This total $g_* s$ does not change — s remains continuous. It does not jump, as it cannot thanks to the 2nd law of thermodynamics.

You just wouldn't use the standard relativistic formulas to calculate the entropy.

This can be compared to, for instance, supersymmetric particles. These do not simply decouple. When one crosses below their threshold, they actually annihilate to less massive particles and one really does not continue to count them — $g_* s$ does decrease, but the entropy is conserved because the temperature does increase due to the annihilation feeding into the lower mass particles.

In summary, during **decoupling** vs. annihilation s stays smooth.

And, when you calculate it, you can use the standard relativistic formulas for all things as computed when they were in equilibrium at $HIGH T \gg all masses$, and the evolution takes care of the rest.

Of course, it is important that the DM particles be included in

computing the total g_{*S} when performing computations.

For example, if the χ is quite heavy and all the SM particles are still around ($T > 300 \text{ GeV}$) and in addition we have $\chi, \bar{\chi}$, one needs to modify the old formula of Eq. (517) that gave $g_{*S} = g_* = 106.75$ by adding in $(7/8)2 \times 2$ (2 spins, 2 for particle plus antiparticle, $7/8$ being present if we assume the χ is a fermion).³

Taking this kind of modification into account, we can now proceed using the standard scaling laws.

In thermal equilibrium, for a relativistic species χ ($m_\chi \ll T, \mu_\chi \ll T$) we have

$$Y_\chi = \frac{45}{2\pi^4} \zeta(3) \frac{g_\chi}{g_{*S}}, \quad (191)$$

³In supersymmetry, the χ is its own antiparticle so the 2nd factor of 2 is not present. Also, in supersymmetry, it is assumed that the χ is the lightest supersymmetry partner, so that all other superparticles would have annihilated by the time T reaches the value at which the χ decouples.

whereas for a non-relativistic species ($m_\chi \gg T$) we have

$$Y_\chi = \frac{45}{4\sqrt{2\pi^7}} \frac{g_\chi}{g_* S} \left(\frac{m_\chi}{T}\right)^{3/2} \exp[-(m_\chi - \mu_\chi)/T] . \quad (192)$$

Here, T is generally dominated by the relativistic species, as they contribute by far the biggest part to the entropy. At this point it is useful to remind ourselves that $T_{CMB} \sim 2.35 \times 10^{-4}$ eV.

If the species χ is stable, then the dominant process which can change the number of particles in a comoving volume are the annihilation and inverse annihilation processes $\chi + \bar{\chi} \leftrightarrow B + \bar{B}$.

Assuming, as before, that there is no asymmetry between particles and antiparticles, and that B and \bar{B} remain in thermal equilibrium throughout the freeze out of the χ 's we have

$$n_\chi = n_{\bar{\chi}}, \quad n_B = n_{\bar{B}} = n_B^{eq} . \quad (193)$$

To reemphasize, the idea behind the B, \bar{B} remaining in equilibrium is that the particles B, \bar{B} will usually have additional interactions (beyond those with $\chi, \bar{\chi}$) which are “stronger” than their interactions with the χ 's, so that they will remain in equilibrium even as the χ 's fall out of equilibrium. An example would be $\chi, \bar{\chi} = \nu, \bar{\nu}$ and $B, \bar{B} = e^-, e^+$; the neutrinos only have weak interactions whereas the e^\pm 's have weak *and* electromagnetic interactions.

Inputting Eq. (193) and the detailed balance relation of Eq. (185) into Eq. (184) we obtain

$$\dot{n}_\chi + 3Hn_\chi = -R(\bar{\chi}\chi \rightarrow \bar{B}B) [n_\chi^2 - (n_\chi^{eq})^2] , \quad (194)$$

where the annihilation rate $R(\bar{\chi}\chi \rightarrow \bar{B}B) = \langle \sigma_{\chi\bar{\chi} \rightarrow B\bar{B}} |v| \rangle$. After summing over all possible B 's we obtain $\langle \sigma_\chi |v| \rangle$, where σ_χ is the total annihilation cross section and $|v| = |v_{Mol}^{\chi\bar{\chi}}|$. The $\langle \dots \rangle$ indicates thermal averaging as defined in Eq. (164).

We now note that for $Y_\chi \equiv n_\chi/s$ we find (using $s \propto T^3$ with $T \propto 1/R$ so that $s = KR^{-3}$ for some constant K — see earlier discussion which discussed why the scaling of s does not change during decoupling)

$$\begin{aligned} \dot{Y}_\chi &= \frac{\dot{n}_\chi}{s} - n_\chi \frac{\dot{s}}{s^2} = \frac{\dot{n}_\chi}{s} - n_\chi \frac{-3KR^{-4}\dot{R}}{K^2R^{-6}} \\ &= \frac{\dot{n}_\chi}{s} + 3n_\chi \frac{R^2\dot{R}}{K} = \frac{\dot{n}_\chi}{s} + 3n_\chi \left(\frac{\dot{R}}{R}\right) \left(\frac{R^3}{K}\right) = \frac{\dot{n}_\chi}{s} + \frac{3n_\chi H}{s}. \end{aligned} \quad (195)$$

which in turn implies that

$$s\dot{Y}_\chi = \dot{n}_\chi + 3Hn_\chi. \quad (196)$$

This is the reason why normalization of n_χ to s is useful.

To proceed further, we introduce the dimensionless parameter $x \equiv m_\chi/T$ which will replace our time variable through the fact that $T = T(t)$. During the radiation dominated epoch, x and t are related

by [c.f. Eq. (99)]

$$t \simeq 0.301 g_*^{-1/2} M_{\text{P}} / T^2 = 0.301 g_*^{-1/2} \frac{M_{\text{P}}}{m^2} x^2. \quad (197)$$

Now, using the fact that $T = \frac{K}{R}$, where K is some constant (not the same as that used above), and defining $Y'_\chi \equiv \frac{dY_\chi}{dx}$ we find

$$\begin{aligned} \dot{Y}_\chi &= \frac{dx}{dt} Y'_\chi = \frac{dx}{dT} \frac{dT}{dt} Y'_\chi = \left[\frac{-m_\chi}{T^2} \right] \left[\frac{-K \dot{R}}{R^2} \right] Y'_\chi = \left[-m_\chi \frac{R^2}{K^2} \right] \left[\frac{-K \dot{R}}{R^2} \right] Y'_\chi \\ &= m_\chi \frac{\dot{R}}{K} Y'_\chi = \frac{m_\chi T \dot{R}}{T K} Y'_\chi = x \frac{\dot{R}}{R} Y'_\chi = x H Y'_\chi. \end{aligned} \quad (198)$$

Using this, the Boltzmann equation becomes

$$Y'_\chi = -\langle \sigma_\chi | v | \rangle \frac{s}{H x} [Y_\chi^2 - (Y_\chi^{eq})^2]. \quad (199)$$

If we now multiply by x/Y_χ^{eq} we obtain (using $s = s^{eq} = \frac{n_\chi^{eq}}{Y_\chi^{eq}}$)

$$\frac{x}{Y_\chi^{eq}} Y'_\chi = -\langle \sigma_\chi | v \rangle \frac{n_\chi^{eq}}{Y_\chi^{eq}} \frac{1}{HY_\chi^{eq}} [Y_\chi^2 - (Y_\chi^{eq})^2] = -\frac{\Gamma_\chi}{H} \left[\left(\frac{Y_\chi}{Y_\chi^{eq}} \right)^2 - 1 \right], \quad (200)$$

where we have defined $\Gamma_\chi \equiv n_\chi \langle \sigma_\chi | v \rangle$.

Clearly, Γ_χ/H describes the “efficiency” of the annihilations relative to the Hubble expansion when in equilibrium. The rate at which the number of χ 's per comoving volume changes is controlled by this efficiency factor times a measure of the deviation from equilibrium.

1. When $\Gamma_\chi \gg H$, the interactions are fast enough that the χ 's thermalize and $Y_\chi \rightarrow Y_\chi^{eq}$.
2. When $\Gamma_\chi \ll H$, the rhs “turns off” and $Y'_\chi = 0$, implying that the abundance Y_χ “freezes in” to the value $Y_\chi^{eq}(x_f)$, where $x_f = m_\chi/T_f$ with T_f being the “freeze out” temperature.

The equilibrium form of Y_χ is different in the case where χ is still relativistic at the time of freeze out vs. that when the χ is non-relativistic at the time of freeze out. One finds from Eqs. (191) and (192), respectively

$$Y_\chi^{eq} = 0.278 \frac{g_\chi}{g_* S(x)} \quad \text{if } x \ll 1 \quad (201)$$

$$Y_\chi^{eq} = 0.145 \frac{g_\chi}{g_* S(x)} x^{3/2} e^{-x} \quad \text{if } x \gg 1. \quad (202)$$

In the above, g_χ is the value for the number density, n_χ , which for fermions is the standard counting factor times $3/4$ and for bosons is simply equal to the standard counting factor. As noted earlier, $g_* S = g_*^{\text{non-DM particles remaining in eq}} + g_\chi$ does not actually change during the decoupling process.

We now consider these two cases in a bit more detail.

Hot relics

If the species decouples when still relativistic, say $x_f = m_\chi/T_f < 3$, then at that time Y_χ^{eq} is constant in time and the final, asymptotic value of Y_χ is quite insensitive to the details of freeze out (*i.e.* the precise value of x_f and precise behavior of $\langle \sigma_\chi |v| \rangle$): from Eq. (201)

$$Y_\infty \equiv Y_\chi(x \rightarrow \infty) \sim Y^{eq}(x_f) = 0.278 \frac{g_\chi}{g_* s(x_f)}. \quad (203)$$

That is, the species freezes out with order unity abundance relative to s [which is roughly the number of photons — $s = 1.8g_* s n_\gamma$, see Eq. (119)].

If we assume that there is no further entropy production following the decoupling of the χ 's (*i.e.* no actual annihilations), then both s

and n_χ will behave as $1/R^3$ and the abundance of χ 's today is

$$n_\chi^0 = s_0 Y_\infty \simeq 2905 Y_\infty \text{ cm}^{-3} = 808 \left[\frac{g_\chi}{g_* s(x_f)} \right] \text{ cm}^{-3}, \quad (204)$$

where 2905 comes from the Table of Eq. (139).

If, after freeze out, the entropy per comoving volume of the Universe should increase, say by a factor of γ (presumably due to some annihilations), then the present abundance of χ 's in a comoving volume would be diminished by γ : $Y_\infty = Y(x_f)/\gamma$.

A species that decouples when it is relativistic is often called a “hot relic”. The present relic mass density contributed by a (once) hot (but now cold, $m_\chi \gg T_0$) relic of mass m_χ is simply

$$\rho_\chi^0 = n_\chi^0 m_\chi = 808 \left[\frac{g_\chi}{g_* s(x_f)} \right] \text{ cm}^{-3} \left(\frac{m_\chi}{\text{eV}} \right) \text{ eV} \quad (205)$$

$$\Omega_\chi^0 h_0^2 = \frac{\rho_\chi^0}{\rho_c^0} = 7.65 \times 10^{-2} \left[\frac{g_\chi}{g_* s(x_f)} \right] \left(\frac{m_\chi}{\text{eV}} \right), \quad (206)$$

where we recall that $H_0 = 100h_0 \text{ km sec}^{-1} \text{ Mpc}^{-1} \simeq \frac{h_0}{9.25 \times 10^{27} \text{ cm}}$, so that

$$\rho_c^0 = \frac{3H_0^2}{8\pi G} = h_0^2 1.057 \times 10^4 \frac{\text{eV}}{\text{cm}^3}. \quad (207)$$

where I used G in the form

$$G = 6.71 \times 10^{-39} \text{ GeV}^{-2} = 6.71 \times 10^{-39} \text{ GeV}^{-1} 0.197 \times 10^{-13} \text{ cm} = 1.32 \times 10^{-61} \frac{\text{eV}}{\text{cm}}. \quad (208)$$

We know that for sure $\Omega_0 h_0^2 \lesssim 1$ at the present day. Combined with the above value for $\Omega_\chi^0 h_0^2$, we conclude that

$$m_\chi \lesssim 13 \text{ eV} \left[\frac{g_* s(x_f)}{g_\chi} \right]. \quad (209)$$

Light mass ($m_\chi \lesssim \text{MeV}$) neutrinos decouple when $T \sim \text{few} \times \text{MeV}$ at which point $g_* s = g_* = 10.75$, not including the $\chi, \bar{\chi}$, but this **does** include our SM neutrinos. For a single, extra 2-component neutrino species (plus anti-neutrino) we have $g_\chi = (3/4) \times 2 = 1.5$,

so that $g_\chi/g_{*S}(x_f) = 0.122$ if we include $g_{*S\chi} = (7/8)g_\chi$ in g_{*S} . KT, however, focus on applying this game to the SM neutrinos that are already included in $g_{*S} = 10.75$. In this case, the appropriate factor is $\frac{g_\chi}{g_{*S}} = \frac{1.5}{10.75} \sim 0.140$. Plugging this in, we find that

$$\Omega_{\nu\bar{\nu}}^0 h_0^2 \sim \frac{m_\nu}{93.5 \text{ eV}}, \quad m_\nu \lesssim 93.5 \text{ eV}. \quad (210)$$

Of course, we actually have stronger constraints on $\Omega_{\nu\bar{\nu}}^0 h_0^2$, coming from all the other things that we know contribute to $\Omega^0 h_0^2$ such as cold dark matter. Looking back at earlier tabulations we learn that $\Omega_m^0 \simeq 0.136$ while $\Omega_{cdm}^0 \simeq 0.113$. This leaves room for $\Omega_{hdm}^0 \sim 0.023$. Now multiplying by $h_0^2 \sim (0.7)^2$ gives $\Omega_{hdm}^0 h_0^2 < 0.0115$.

So, very roughly we now know that

$$\Omega_{\nu\bar{\nu}}^0 h_0^2 \lesssim 0.01 \quad (211)$$

resulting in a bound of

$$m_\nu \lesssim 0.935 \text{ eV} \quad (212)$$

on any single neutrino. If there are also axions, then this bound is strengthened. However, we also know that the 3 neutrinos are pretty degenerate, yielding

$$m_\nu < 0.31 \text{ eV} \quad (213)$$

for each. This is the bound that is relevant for the current data.

If the overall $\Omega_{\nu\bar{\nu}}^0 h_0^2 \lesssim 0.01$ bound decreases with future data, one could arrive at a situation where the lightest neutralino would have to be nearly massless given the measured Δm^2 between neutrinos. The point at which this happens depends upon whether the neutrinos have a normal hierarchy with $m_3 > m_2 \gtrsim m_1$ ($1 = e, 2 = \mu, 3 = \tau$) or inverted hierarchy with $m_2 \gtrsim m_1 > m_3$. The key input is the measured value of $\Delta m_{32}^2 \sim 2.43 \times 10^{-3} \text{ eV}^2$. Recall also that

$\Delta m_{12}^2 \sim 7.6 \times 10^{-5} \text{ eV}^2$ is much smaller.

1. In the normal hierarchy case, $m_2 \sim m_1 \sim 0$, the measured Δm_{23}^2 would yield $m_3 \sim 0.049 \text{ eV}$. Since 1 and 2 are much lighter than 3, this is a single “heavy” neutrino case. It could be tested by cosmological data if $\Omega_{\nu\bar{\nu}}^0 h_0^2 \lesssim 0.0005$ could be probed.
2. In the inverted hierarchy case, at the limit $m_3 \sim 0$ one would have $m_2 \sim m_1 \sim 0.049 \text{ eV}$, *i.e.* two “heavy” neutrinos, which could be tested if $\Omega_{\nu\bar{\nu}}^0 h_0^2 \lesssim 0.001$ could be probed.

In either case, LSST is potentially capable of obtaining the required precision on $\Omega_{\nu\bar{\nu}}^0 h_0^2$, but I don't believe it will be reached by experiments planned between now and then.

Of course, we can also consider a ψ that decouples much earlier on, in particular when $T \gtrsim 300 \text{ GeV}$ (requiring that the ψ interact even more weakly than a neutrino). Let us assume that the ψ has $g_\psi = 1.5$. Then $g_{*S}(x_f) \simeq g_*(x_f) \sim 106.75 + (7/8) \times 1.5 \sim 108.06$

and $g_\psi/g_{*S} \simeq 1.5/108.06 \simeq 0.0139$. Then, the present contribution to the energy density is very roughly

$$\Omega_{\psi\bar{\psi}}^0 h_0^2 = \frac{m_\psi}{910 \text{ eV}}, \quad (214)$$

which is about a factor of 10 less than that of a conventional neutrino species of the same mass.

The present number density of such a species is given by the standard formula, see Eq. (204),

$$n_\psi^0 \sim 808 \left[\frac{g_\psi}{g_{*S}(x_f)} \right] \text{ cm}^{-3} \sim 11 \text{ cm}^{-3}, \quad (215)$$

so that $n_\psi^0 \ll n_\gamma^0 \sim 413 \text{ cm}^{-3}$.

Thus, a species that decouples when $g_{*S} \gg 1$ has a present abundance that is much less than that of the microwave photons, and if the species is massless, a temperature much less than the photon

temperature, $T_\psi \simeq (3.91/g_{*S}(x_f))^{1/3}T$. For the latter reason, such a relic is often referred to as **warm relic**.

In this case, the temperatures for the ψ and for the photons, $T = T_\gamma$, diverge precisely because a lot of the SM particles do in fact annihilate and feed entropy into the photons. Entropy conservation requires that $T = T_\gamma$ increases when annihilations occur and g_{*S} decreases.

Examples of such a warm relic include a light gravitino, or a light photino, where light means $m_\psi \lesssim \text{keV}$.

Cold relics

The definition of “cold” is that freeze out occurs when the species is non-relativistic ($x_f \gtrsim 3$).

This is a more difficult case than the hot relic case. At the time of freeze out Y^{eq} is decreasing exponentially with x . As a result the precise details of freeze out *are* important.

Gondolo and Gelmini proceed by expanding

$$\sigma_\chi |v_{Mol}| = \sum_{n=0}^{\infty} \frac{a_n}{n!} \epsilon^n, \quad (216)$$

where

$$\epsilon = \frac{s - 4m_\chi^2}{4m_\chi^2} \quad (217)$$

is the kinetic energy per unit mass in the laboratory frame. For s

close to threshold, $s \sim 4m_\chi^2$, $\epsilon \sim (\beta_i^{c.m.})^2$, but it is more correct to define it as above.

They then use the formulas given earlier to show that

$$\langle \sigma_\chi |v_{Mol}| \rangle = a_0 + \frac{3}{2}a_1x^{-1} + \left[\frac{9}{2}a_1 + \frac{15}{8}a_2 \right] x^{-2} + \dots, \quad (218)$$

where, as before, $x = m_\chi/T$. They note that this compares to the non-relativistic approximation implicit in Kolb and Turner, which gives

$$\langle \sigma_\chi |v_{Mol}| \rangle_{n.r.} = a_0 + \frac{3}{2}a_1x^{-1} + \frac{15}{8}a_2x^{-2} + \dots. \quad (219)$$

The expressions only begin to differ at order x^{-2} . In the non-relativistic limit, one can think of $\sigma_\chi |v_{Mol}| \propto |v_{Mol}|^p$, where $p = 0, 2$ corresponds to S-wave, P-wave annihilation.

Since $\langle |v_{Mol}| \rangle \propto T^{1/2}$, $\langle \sigma_\chi |v_{Mol}| \rangle \propto T^n$, with $n = 0, 1$ for S-wave, P-wave annihilation, respectively.

In any case, regardless of whether we use the full relativistic treatment with v_{Mol} or the non-relativistic treatment, we may write (as in KT)

$$\langle \sigma_\chi |v| \rangle \equiv \sigma_0 (T/m_\chi)^n = \sigma_0 x^{-n}, \quad \text{for } x \gtrsim 3. \quad (220)$$

Meanwhile, we should recall that $s \propto T^3 \propto x^{-3}$ and $H = \frac{\dot{R}}{R} \propto t^{-1} \propto x^{-2}$ [c.f. Eq. (197)]. Using the above parameterization for $\langle \sigma_\chi |v| \rangle$ and these scalings, the Boltzmann equation, Eq. (199),

$$Y'_\chi = -\langle \sigma_\chi |v| \rangle \frac{s}{Hx} [Y_\chi^2 - (Y_\chi^{eq})^2], \quad (221)$$

can be rewritten in the form

$$\frac{dY_\chi}{dx} = -\lambda x^{-(n+2)} [Y_\chi^2 - (Y_\chi^{eq})^2], \quad (222)$$

where

$$\lambda = \left[\frac{\langle \sigma_\chi | v | \rangle s}{xH(x)} \right]_{x=1} \stackrel{\text{Eqs. (116,98)}}{=} 0.264(g_* s/g_*^{1/2}(x))M_{\text{P}}m_\chi\sigma_0 \quad (223)$$

$$Y_\chi^{eq} \stackrel{\text{Eq. (202)}}{=} 0.145(g_\chi/g_* s(x))x^{3/2}e^{-x}. \quad (224)$$

The first two equations referenced above were

$$s = \frac{2\pi^2}{45}g_* s T^3$$

and

$$H = \left[\frac{8\pi G \pi^2}{3 \cdot 30}g_* T^4 \right]^{1/2} = 1.66g_*^{1/2} \frac{T^2}{M_{\text{P}}},$$

while Eq. (202) came from the earlier expression of Eq. (192), namely

for non-relativistic particles

$$Y_\chi = \frac{45}{4\sqrt{2\pi^7}} \frac{g_\chi}{g_* S} \left(\frac{m_\chi}{T}\right)^{3/2} \exp[-(m_\chi - \mu_\chi)/T] .$$

and we have taken $T = m_\chi$ ($x = 1$) for the computation of s and H for getting λ .

To be absolutely clear, $\frac{\langle\sigma_A|v|\rangle s}{xH}$ as a function of x takes the form

$$\begin{aligned} \frac{\langle\sigma_A|v|\rangle s}{xH} &= \frac{\sigma_0 x^{-n} \left(\frac{2\pi^2}{45} g_* S \frac{m_\chi^3}{x^3}\right)}{x \left(\frac{8\pi G}{3}\right)^{1/2} g_*^{1/2} \frac{m_\chi^2}{x^2}} \\ &= m_\chi \frac{2\pi^2}{45} \underbrace{\left(\frac{3}{8\pi G}\right)^{-1/2} \frac{g_* S}{g_*^{1/2}} \sigma_0}_{\lambda} x^{-(n+2)} \\ &\equiv \lambda x^{-(n+2)} . \end{aligned} \tag{225}$$

Note that because of the M_{P} factor in λ , $\lambda \gg 1$.

We are now ready to solve the Boltzmann equation using some approximate techniques.

Define $\Delta = Y - Y^{eq}$. We then have from Eq. (222)

$$\Delta' = -Y^{eq'} - \lambda x^{-(n+2)} \Delta (2Y^{eq} + \Delta), \quad (226)$$

where (assuming constant g_{*S} during decoupling) Eq. (224) implies

$$Y^{eq'} = Y^{eq} \left(\frac{3}{2x} - 1 \right) \simeq -Y^{eq}, \quad (227)$$

if the relevant x values are large, as will turn out to be the case.

With this approximation for $Y^{eq'}$, Eq. (226) takes the form

$$\Delta' = Y^{eq} - \lambda x^{-(n+2)} \Delta (2Y^{eq} + \Delta), \quad (228)$$

At early times ($1 < x \ll x_f$), Y tracks Y^{eq} very closely, and both Δ and $|\Delta'|$ are small. So an approximate solution is obtained by setting $\Delta' = 0$, yielding

$$\Delta \simeq \lambda^{-1} x^{n+2} Y^{eq} / (2Y^{eq} + \Delta) \quad (229)$$

$$\simeq \frac{1}{2\lambda} x^{n+2}. \quad (230)$$

At late times ($x \gg x_f$), Y tracks Y^{eq} very poorly: $\Delta \simeq Y \gg Y^{eq}$, and the terms in Eq. (226) involving $Y^{eq'}$ and Y^{eq} can be safely neglected, so that

$$\Delta' = -\lambda x^{-(n+2)} \Delta^2. \Rightarrow \frac{d\Delta}{\Delta^2} = -dx \frac{\lambda}{x^{(n+2)}}. \quad (231)$$

Upon integration of this latter equation from $x = x_f$ to $x = \infty$, we

obtain

$$\frac{1}{\Delta_\infty} - \frac{1}{\Delta(x_f)} = \frac{\lambda}{(n+1)x_f^{n+1}}. \quad (232)$$

It remains to determine x_f .

We recall that $x = x_f$ is the time when Y ceases to track Y^{eq} , or equivalently, when Δ becomes of order Y^{eq} . Let us define x_f by the criterion $\Delta(x_f) = cY^{eq}(x_f)$, where c is a numerical constant of order unity. Substituting this into Eq. (229) evaluated at $x = x_f$ yields

$$\Delta(x_f) \simeq \frac{x_f^{n+2}}{\lambda(2+c)}. \quad (233)$$

Using $\Delta(x_f) = cY^{eq}(x_f) = cax_f^{3/2}e^{-x_f}$ with $a = 0.145(g_\chi/g_*s)$, see Eq. (224), we have the freeze out condition

$$cax_f^{3/2}e^{-x_f} = \frac{x_f^{n+2}}{\lambda(2+c)}, \quad (234)$$

which has the approximate solution

$$x_f \simeq \ln[(2+c)\lambda ac] - \left(n + \frac{1}{2}\right) \ln\{\ln[(2+c)\lambda ac]\}. \quad (235)$$

Note that x_f depends only logarithmically upon the numerical condition for freeze out, *i.e.* the value of c , as will the final abundance — see below.

Meanwhile, plugging $\Delta(x_f) \simeq \frac{x_f^{n+2}}{\lambda(2+c)}$ into Eq. (232) yields an expression for $\frac{1}{\Delta_\infty}$:

$$\frac{1}{\Delta_\infty} = \frac{\lambda(2+c)}{x_f^{n+2}} + \frac{\lambda}{(n+1)x_f^{n+1}}. \quad (236)$$

Assuming x_f will turn out to be large we may drop the first term and obtain

$$Y_\infty \simeq \Delta_\infty \simeq \frac{n+1}{\lambda} x_f^{n+1}, \quad (237)$$

where we have used $Y_{\infty}^{eq} \propto e^{-\infty} = 0$ so that $\Delta_{\infty} = Y_{\infty} - Y_{\infty}^{eq} = Y_{\infty}$.

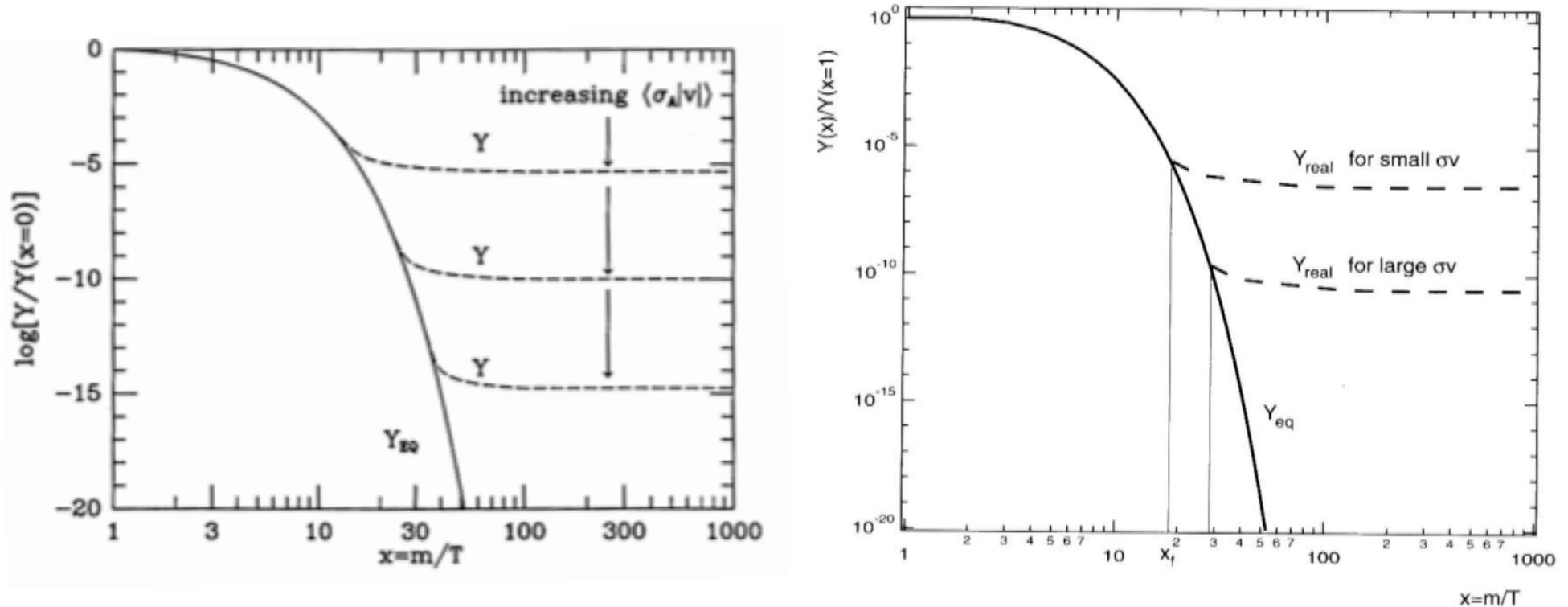


Figure 4: Freeze-out plots from Kolb and Turner and from Bergstrom and Goobar. This latter seems to have correction normalization label.

One can of course perform a numerical integration of the Boltzmann equation with the result shown. Choosing $c(c+2) = n+1$ gives the best fit to the numerical results for the final abundance Y_{∞} (to

better than 5% for any $x_f \geq 3$).

With the above choice of c our formula for x_f simplifies to

$$\begin{aligned} x_f &= \ln[\lambda a] - \left(n + \frac{1}{2}\right) \ln \{\ln[\lambda a]\} \\ &= \ln \left[0.038(n+1)(g_\chi/g_*^{1/2})M_{\text{P}}m_\chi\sigma_0 \right] \\ &\quad - \left(n + \frac{1}{2}\right) \ln \left\{ \ln \left[0.038(n+1)(g_\chi/g_*^{1/2})M_{\text{P}}m_\chi\sigma_0 \right] \right\}, \end{aligned} \tag{238}$$

where $0.038 = 0.264 \times 0.145$ comes from the coefficients in λ and a , Eqs. (223) and (224), respectively. Note the big M_{P} in the x_f expression which means that x_f will typically be big unless σ_0 is really tiny (in appropriate inverse mass-squared units).

Using Eq. (237) and keeping only the first term in the x_f expression

above, gives

$$Y_\infty = \frac{3.79(n+1)x_f^{n+1}}{(g_* s/g_*^{1/2})M_{\text{P}}m_\chi\sigma_0} \quad (239)$$

Just as for hot relics, we use

$$n_\chi^0 = s_0 Y_\infty = 2905 Y_\infty \text{ cm}^{-3} \quad (240)$$

but with the above expression for Y_∞ , yielding

$$n_\chi^0 = 1.101 \times 10^4 \frac{(n+1)x_f^{n+1}}{(g_* s/g_*^{1/2})M_{\text{P}}m_\chi\sigma_0} \quad (241)$$

with corresponding results of

$$\rho_\chi^0 = n_\chi^0 m_\chi = 1.101 \times 10^4 \frac{(n+1)x_f^{n+1}}{(g_* s/g_*^{1/2})M_{\text{P}}\sigma_0}$$

$$\Omega_{\chi}^0 h_0^2 = \left(\frac{\rho_{\chi}^0}{\rho_c^0} \right) h_0^2 = 1.042 \times 10^9 \frac{(n+1)x_f^{n+1} \text{ GeV}^{-1}}{(g_{*S}/g_{*}^{1/2}) M_{\text{P}} \sigma_0} \quad (242)$$

where in this case we wrote ρ_c^0 in GeV units (c.f. Eq. (207)):

$$\rho_c^0 = \frac{3H_0^2}{8\pi G} = h_0^2 1.057 \times 10^{-5} \frac{\text{GeV}}{\text{cm}^3}. \quad (243)$$

Obviously, x_f and the final $\Omega_{\chi}^0 h_0^2$ depend on σ_0 . A few examples are in order.

Heavy ($m \gg \text{MeV}$), stable neutrino species

Because of its large mass, such a neutrino, call it N , will decouple when it is non-relativistic (to be verified) and the formulae for a cold relic apply.

Annihilation for such a species proceeds through Z exchange to final states $i\bar{i}$ with $i = \nu_L, e, \mu, \tau, u, d, s, \dots$. The annihilation cross section

depends upon whether the N is of the Dirac or Majorana type. For $T \lesssim m_N \lesssim m_Z$, we find

$$\langle \sigma_N | v \rangle_{Dirac} = \frac{G_F^2 m_N^2}{2\pi} \sum_i (1 - z_i^2)^{1/2} \times \left[(C_{V_i}^2 + C_{A_i}^2) \left(1 + \frac{1}{2} z_i^2 \right) \right] \quad (244)$$

$$\langle \sigma_N | v \rangle_{Maj} = \frac{G_F^2 m_N^2}{2\pi} \sum_i (1 - z_i^2)^{1/2} \times \left[(C_{V_i}^2 + C_{A_i}^2) \frac{8\beta_i^2}{3} + C_{A_i}^2 2z_i^2 \right], \quad (245)$$

where $z_i \equiv m_i/m_N$, β is the relative velocity in the cm, and C_V and C_A are given in terms of the weak isospin I_3 , the electric charge q and the Weinberg angle θ_W by $C_A = I_3$, $C_V = I_3 - 2q \sin^2 \theta_W$. The sum ranges over all quark and lepton species lighter than m_N .⁴

⁴We are assuming that $m_N < m_Z$ and are therefore using the “contact” Fermi 4-point style of interaction, equivalent to the approximation $1/[(s - m_Z^2)^2 + \Gamma_Z^2/4] \sim 1/m_Z^4$.

Focusing on the Dirac case, annihilations proceed through the S-wave and $\langle \sigma_N |v| \rangle$ is velocity independent:

$$\sigma_0 \simeq c_2 \frac{G_F^2 m_N^2}{2\pi}, \quad (246)$$

where $c_2 \sim 5$ after performing the sums. Taking $g_N = 2$ (2 spins but no 2nd 2 for \bar{N} since Boltzmann focuses on only N , or \bar{N} , on its own) and $g_* \sim 60$ (no $W, Z, t, b, H \Rightarrow 106.75 - 31 = 75.75?$), from our formulae for x_f and Y_∞ we find

$$\begin{aligned} x_f &\simeq 15 + 3 \ln(m_N / \text{GeV}) + \ln(c_2/5) \\ Y_\infty &\simeq 6 \times 10^{-9} \left(\frac{m_N}{\text{GeV}} \right)^{-3} \left[1 + \frac{3 \ln(m_N / \text{GeV})}{15} + \frac{\ln(c_2/t)}{15} \right] \end{aligned} \quad (247)$$

from which we obtain

$$\Omega_{N\bar{N}}^0 h_0^2 \simeq 3 \left(\frac{m_N}{\text{GeV}} \right)^{-2} \left[1 + \frac{3 \ln(m_N / \text{GeV})}{15} \right], \quad (248)$$

where we included the factor of 2 for $N + \bar{N}$ abundance.

We observe that freeze out takes place at

$$T_f \simeq m_N / 15 \simeq 70 \text{ MeV} (m_N / \text{GeV}), \quad (249)$$

which, in particular, is before the light neutrinos freeze out. The heavy neutrinos annihilate and become rare early on, and the annihilation process quenches.

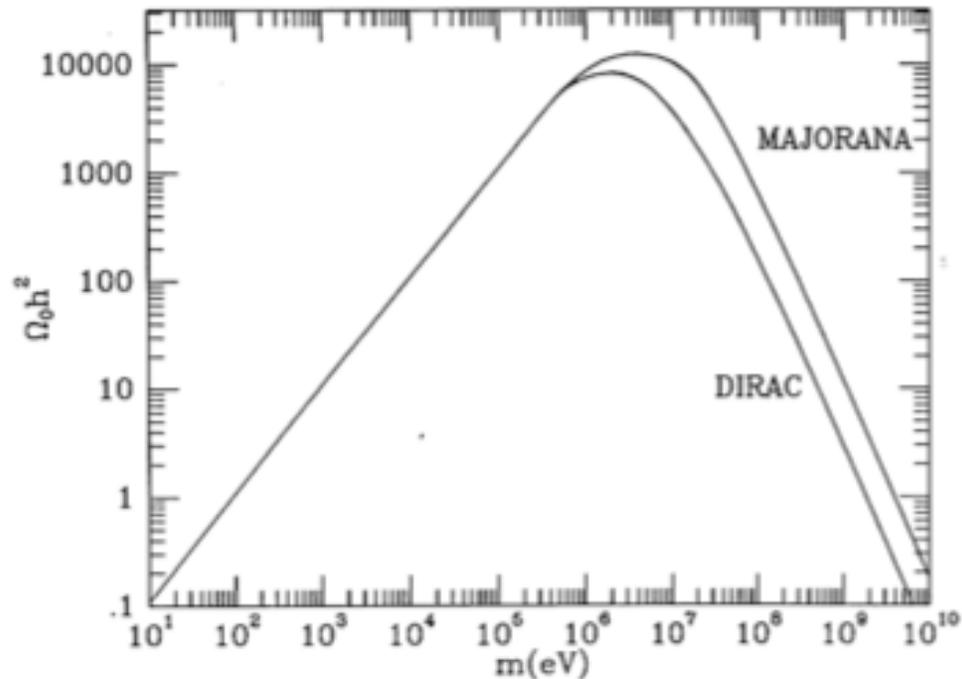
If we require that the N, \bar{N} not overclose the Universe, $\Omega_{N\bar{N}}^0 h_0^2 \leq 1$, we obtain the “Lee-Weinberg” bound of

$$m_N \gtrsim 2 \text{ GeV}. \quad (250)$$

Of course, nowadays we might ask that the N (which in the mass range being considered should be non-relativistic) should give the observed (mainly cold) dark matter density: $\Omega_{N\bar{N}}^0 h_0^2 \sim 0.11$. This would work for $m_N \sim \sqrt{30}$ GeV.

The problem is that since the N that we are envisioning couples to the Z , it would have appeared at LEP in Z decays. LEP excludes such a heavy neutrino for $m_N < m_Z/2 \sim 45$ GeV; for $m_N > 45$ GeV, $\Omega_{N\bar{N}}^0 h_0^2 < 0.0015$.

For the Majorana case, annihilation proceeds through both the S-wave and the P-wave. However, the formulae for x_f , Y_∞ and $\Omega_{N\bar{N}}^0 h_0^2$ are similar. The plot below gives a comparison of the two cases.



We note:

- For $m_N \lesssim \text{MeV}$, $\Omega_{N\bar{N}}^0 h_0^2 \propto m_N$ because the relic abundance is constant.
- For $m_N \gtrsim \text{MeV}$, $\Omega_{N\bar{N}}^0 h_0^2 \propto m_N^{-2}$ due to the fact that the relic abundance is decreasing like m_N^{-3} .
- $\Omega_{N\bar{N}}^0 h_0^2$ achieves its maximum for $m_N \sim \text{MeV}$.
- Requiring $\Omega_{N\bar{N}}^0 h_0^2 \lesssim 0.01$ (which is beyond the plot limits) implies

$m_N \lesssim 1 \text{ eV}$ on the low side.

In this region, Z decay limits imply that there is no room for a new N in addition to the standard 3 neutrino species.

- As mentioned above, on the high side $m_N \gtrsim 45 \text{ GeV}$ is required by LEP and extending the graph would show that $\Omega_{N\bar{N}}^0 h_0^2 \sim 0.0015 \left(\frac{45 \text{ GeV}}{m_N} \right)^2$ in the Dirac case.

LSP of MSSM

For this, we need to make an excursion into supersymmetry. However, from the previous discussion, we already see that the LSP cannot be that closely analogous to a heavy N if it is to supply the observed cold dark matter.

This is because a heavy-neutrino N does not give enough dark matter. What seems to be required is that the $\langle \sigma \chi | v | \rangle$ be substantially smaller than that for a recurrence of a neutrino with SM-like couplings.

This can be accomplished in the context of supersymmetry provided the χ LSP has the correct composition in terms of its bino, wino, and higgsino (and in the NMSSM, singlino) content.

This shows the limitation of the WIMP miracle argument. We are quite sensitive to the actual cross section and so the WIMP miracle in the $\Omega_{DM}^0 h_0^2$ sense is only a miracle within a factor of 100 or so, which is already quite an achievement, but more precision is needed.

So, at this point, we must learn more about Supersymmetry.

Supersymmetry

As I mentioned at the beginning of the quarter, I will not be deriving the rules for constructing a supersymmetric Lagrangian, I will simply state those rules and give the procedures for employing such a Lagrangian. I will be following Steve Martin's "Supersymmetry Primer" for notation since that is easily available on the arXiv. Other possible references include the old Haber-Kane Physics Report, John Ternings "Modern Supersymmetry", and the Wess-Bagger "Supersymmetry and Supergravity". For detailed Feynman rule derivations and spinor techniques, see Dreiner, Haber and Martin, arXiv:0812.1594. The 246 Supersymmetry Barcelona lecture attachment on my home page focuses on the phenomenology of supersymmetry at colliders. Although it is a bit old now, the general phenomenology reviewed there is mostly still relevant.

General Structure of a Supersymmetric Theory

The extension of the Lorentz group to supersymmetry requires (see Martin) an equal number of bosons and fermions:

$$n_B = n_F . \quad (251)$$

The simplest possibility for a supermultiplet consistent with Eq. (251) has a single Weyl fermion (with two spin helicity states, so $n_F = 2$) and two real scalars (each with $n_B = 1$). It is natural to assemble the two real scalar degrees of freedom into a complex scalar field; as we will see below this provides for convenient formulations of the supersymmetry algebra, Feynman rules, supersymmetry-violating effects, etc. This combination of a two-component Weyl fermion and a complex scalar field is called a *chiral* or *matter* or *scalar* supermultiplet.

The next-simplest possibility for a supermultiplet contains a spin-1 vector boson. If the theory is to be renormalizable, this must be a gauge boson that is massless, at least before the gauge symmetry is spontaneously broken.

A massless spin-1 boson has two helicity states, so the number of bosonic degrees of freedom is $n_B = 2$. Its superpartner is therefore a massless spin-1/2 Weyl fermion, again with two helicity states, so $n_F = 2$. (If one tried to use a massless spin-3/2 fermion instead, the theory would not be renormalizable.)

Gauge bosons must transform as the adjoint representation of the gauge group, so their fermionic partners, called *gauginos*, must also. Since the adjoint representation of a gauge group is always its own conjugate, the gaugino fermions must have the same gauge transformation properties for left-handed and for right-handed components. Such a combination of spin-1/2 gauginos and spin-1 gauge bosons is called a *gauge* or *vector* supermultiplet.

If we include gravity, then the spin-2 graviton (with 2 helicity states, so $n_B = 2$) has a spin-3/2 superpartner called the gravitino. The gravitino would be massless if supersymmetry were unbroken, and so it has $n_F = 2$ helicity states.

There are other possible combinations of particles with spins that can satisfy Eq. (251). However, these are always reducible to combinations⁵ of chiral and gauge supermultiplets if they have renormalizable interactions, except in certain theories with “extended” supersymmetry.

The ordinary, non-extended, phenomenologically viable supersymmetric model is sometimes called $N = 1$ supersymmetry, with N referring to the number of supersymmetries (the number of distinct copies of the generators Q, Q^\dagger of supersymmetry transformations).

⁵For example, if a gauge symmetry were to spontaneously break without breaking supersymmetry, then a massless vector supermultiplet would “eat” a chiral supermultiplet, resulting in a massive vector supermultiplet with physical degrees of freedom consisting of a massive vector ($n_B = 3$), a massive Dirac fermion formed from the gaugino and the chiral fermion ($n_F = 4$), and a real scalar ($n_B = 1$).

Thus, In a supersymmetric extension of the Standard Model each of the known fundamental particles is in either a chiral or gauge supermultiplet, and must have a superpartner with spin differing by $1/2$ unit.

The first step in understanding the exciting consequences of this prediction is to decide exactly how the known particles fit into supermultiplets, and to give them appropriate names.

Summary of the Chiral Supermultiplets

A crucial observation here is that **only what are called chiral supermultiplets can contain fermions whose left-handed parts transform differently under the gauge group than their right-handed parts.** All of the Standard Model fermions (the known quarks and leptons) have this property, so they must be members of chiral supermultiplets.⁶

⁶In particular, one cannot attempt to make a spin- $1/2$ neutrino be the superpartner of the spin-1 photon; the neutrino is in a doublet, and the photon is neutral, under weak isospin.

The names for the spin-0 partners of the quarks and leptons are constructed by prepending an “s”, for scalar. So, generically they are called *squarks* and *sleptons* (short for “scalar quark” and “scalar lepton”), or sometimes *sfermions*.

The left-handed and right-handed pieces of the quarks and leptons are separate two-component Weyl fermions with different gauge transformation properties in the Standard Model, so each must have its own complex scalar partner.

The symbols for the squarks and sleptons are the same as for the corresponding fermion, but with a tilde ($\tilde{}$) used to denote the superpartner of a Standard Model particle. For example, the superpartners of the left-handed and right-handed parts of the electron Dirac field are called left- and right-handed selectrons, and are denoted \tilde{e}_L and \tilde{e}_R .

It is important to keep in mind that the “handedness” here does not

refer to the helicity of the selectrons (they are spin-0 particles) but to that of their superpartners.

A similar nomenclature applies for smuons and staus: $\tilde{\mu}_L, \tilde{\mu}_R, \tilde{\tau}_L, \tilde{\tau}_R$. The Standard Model neutrinos (neglecting their very small masses) are always left-handed, so the sneutrinos are denoted generically by $\tilde{\nu}$, with a possible subscript indicating which lepton flavor they carry: $\tilde{\nu}_e, \tilde{\nu}_\mu, \tilde{\nu}_\tau$.

Finally, a complete list of the squarks is \tilde{q}_L, \tilde{q}_R with $q = u, d, s, c, b, t$. The gauge interactions of each of these squark and slepton fields are the same as for the corresponding Standard Model fermions; for instance, the left-handed squarks \tilde{u}_L and \tilde{d}_L couple to the W boson, while \tilde{u}_R and \tilde{d}_R do not.

It seems clear that the Higgs scalar boson must reside in a chiral supermultiplet, since it has spin 0. Actually, it turns out that just one chiral supermultiplet is not enough.

One reason for this is that if there were only one Higgs chiral supermultiplet, the electroweak gauge symmetry would suffer a gauge anomaly, and would be inconsistent as a quantum theory. This is because the conditions for cancellation of gauge anomalies include $\text{Tr}[T_3^2 Y] = \text{Tr}[Y^3] = 0$, where T_3 and Y are the third component of weak isospin and the weak hypercharge, respectively, in a normalization where the ordinary electric charge is $Q_{\text{EM}} = T_3 + Y$. The traces run over all of the left-handed Weyl fermionic degrees of freedom in the theory.

In the Standard Model, these conditions are already satisfied, somewhat miraculously, by the known quarks and leptons.

Now, a fermionic partner of a Higgs chiral supermultiplet must be a weak isodoublet with weak hypercharge $Y = 1/2$ or $Y = -1/2$. In either case alone, such a fermion will make a non-zero contribution to the traces and spoil the anomaly cancellation. This can be avoided if there are two Higgs supermultiplets, one $Y = +1/2$ and one with

$Y = -1/2$, so that the total contribution to the anomaly traces from the two fermionic members of the Higgs chiral supermultiplets vanishes by cancellation.

As we will see later, two Higgs chiral supermultiplets are also necessary for another completely different reason: because of the structure of supersymmetric theories, only a $Y = 1/2$ Higgs chiral supermultiplet can have the Yukawa couplings necessary to give masses to charge $+2/3$ up-type quarks (up, charm, top), and only a $Y = -1/2$ Higgs can have the Yukawa couplings necessary to give masses to charge $-1/3$ down-type quarks (down, strange, bottom) and to the charged leptons.

We will call the $SU(2)_L$ -doublet complex scalar fields with $Y = 1/2$ and $Y = -1/2$ by the names H_u and H_d , respectively.²

The weak isospin components of H_u with $T_3 = (1/2, -1/2)$ have

²Other notations in the literature have H_1, H_2 or H, \bar{H} instead of H_u, H_d . The notation used here has the virtue of making it easy to remember which Higgs VEVs gives masses to which type of quarks.

electric charges 1, 0 respectively, and are denoted (H_u^+, H_u^0) . Similarly, the $SU(2)_L$ -doublet complex scalar H_d has $T_3 = (1/2, -1/2)$ components (H_d^0, H_d^-) .

The neutral scalar that corresponds to the physical Standard Model Higgs boson is in a linear combination of H_u^0 and H_d^0 .

The generic nomenclature for a spin-1/2 superpartner is to append “-ino” to the name of the Standard Model particle, so the fermionic partners of the Higgs scalars are called higgsinos. They are denoted by \tilde{H}_u, \tilde{H}_d for the $SU(2)_L$ -doublet left-handed Weyl spinor fields, with weak isospin components $\tilde{H}_u^+, \tilde{H}_u^0$ and $\tilde{H}_d^0, \tilde{H}_d^-$.

We have now found all of the chiral supermultiplets of a minimal phenomenologically viable extension of the Standard Model. They are summarized in Table 1, classified according to their transformation properties under the Standard Model gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$, which combines u_L, d_L and ν, e_L degrees of freedom into

$SU(2)_L$ doublets.

Names		spin 0	spin 1/2	$SU(3)_C, SU(2)_L, U(1)_Y$
squarks, quarks ($\times 3$ families)	Q	$(\tilde{u}_L \ \tilde{d}_L)$	$(u_L \ d_L)$	$(\mathbf{3}, \mathbf{2}, \frac{1}{6})$
	\bar{u}	\tilde{u}_R^*	u_R^\dagger	$(\bar{\mathbf{3}}, \mathbf{1}, -\frac{2}{3})$
	\bar{d}	\tilde{d}_R^*	d_R^\dagger	$(\bar{\mathbf{3}}, \mathbf{1}, \frac{1}{3})$
sleptons, leptons ($\times 3$ families)	L	$(\tilde{\nu} \ \tilde{e}_L)$	$(\nu \ e_L)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$
	\bar{e}	\tilde{e}_R^*	e_R^\dagger	$(\mathbf{1}, \mathbf{1}, 1)$
Higgs, higgsinos	H_u	$(H_u^+ \ H_u^0)$	$(\tilde{H}_u^+ \ \tilde{H}_u^0)$	$(\mathbf{1}, \mathbf{2}, +\frac{1}{2})$
	H_d	$(H_d^0 \ H_d^-)$	$(\tilde{H}_d^0 \ \tilde{H}_d^-)$	$(\mathbf{1}, \mathbf{2}, -\frac{1}{2})$

Table 1: Chiral supermultiplets in the Minimal Supersymmetric Standard Model. The spin-0 fields are complex scalars, and the spin-1/2 fields are left-handed two-component Weyl fermions.

Here we follow a standard convention, whereby all chiral supermultiplets are defined in terms of left-handed Weyl spinors. In particular, the *conjugates* of the right-handed quarks and leptons (and their superpartners) appear in Table 1. This protocol for defining chiral

supermultiplets will prove very useful for constructing supersymmetric Lagrangians.

It is also useful to have a symbol for each of the chiral supermultiplets as a whole; these are indicated in the second column of Table 1. Thus, for example, Q stands for the $SU(2)_L$ -doublet chiral supermultiplet containing \tilde{u}_L, u_L (with weak isospin component $T_3 = 1/2$), and \tilde{d}_L, d_L (with $T_3 = -1/2$), while \bar{u} stands for the $SU(2)_L$ -singlet supermultiplet containing $\tilde{u}_R^*, u_R^\dagger$.

There are three families for each of the quark and lepton supermultiplets, Table 1 lists the first-family representatives. A family index $i = 1, 2, 3$ can be affixed to the chiral supermultiplet names (Q_i, \bar{u}_i, \dots) when needed, for example $(\bar{e}_1, \bar{e}_2, \bar{e}_3) = (\bar{e}, \bar{\mu}, \bar{\tau})$. The bar on $\bar{u}, \bar{d}, \bar{e}$ fields is part of the name, and does not denote any kind of conjugation.

The Higgs chiral supermultiplet H_d (containing $H_d^0, H_d^-, \tilde{H}_d^0, \tilde{H}_d^-$) has exactly the same Standard Model gauge quantum numbers as

the left-handed sleptons and leptons L_i , for example $(\tilde{\nu}, \tilde{e}_L, \nu, e_L)$.

Naively, one might therefore suppose that we could have been more economical in our assignment by taking a neutrino and a Higgs scalar to be superpartners, instead of putting them in separate supermultiplets. This would amount to the proposal that the Higgs boson and a sneutrino should be the same particle.

This attempt played a key role in some of the first attempts to connect supersymmetry to phenomenology, but it is now known to not work. Even ignoring the anomaly cancellation problem mentioned above, many insoluble phenomenological problems would result, including lepton-number non-conservation and a mass for at least one of the neutrinos in gross violation of experimental bounds. Therefore, all of the superpartners of Standard Model particles are really new particles, and cannot be identified with some other Standard Model state.

The Vector Supermultiplets

The vector bosons of the Standard Model clearly must reside in gauge supermultiplets. Their fermionic superpartners are generically referred to as gauginos.

The $SU(3)_C$ color gauge interactions of QCD are mediated by the gluon, whose spin-1/2 color-octet supersymmetric partner is the gluino. As usual, a tilde is used to denote the supersymmetric partner of a Standard Model state, so the symbols for the gluon and gluino are g and \tilde{g} respectively.

The electroweak gauge symmetry $SU(2)_L \times U(1)_Y$ is associated with spin-1 gauge bosons W^+, W^0, W^- and B^0 , with spin-1/2 superpartners $\tilde{W}^+, \tilde{W}^0, \tilde{W}^-$ and \tilde{B}^0 , called *winos* and *binos*. After electroweak symmetry breaking, the W^0, B^0 gauge eigenstates mix to give mass eigenstates Z^0 and γ . The corresponding gaugino mixtures of \tilde{W}^0 and \tilde{B}^0 are called zino (\tilde{Z}^0) and photino ($\tilde{\gamma}$); if

supersymmetry were unbroken, they would be mass eigenstates with masses m_Z and 0. Table 2 summarizes the gauge supermultiplets of a minimal supersymmetric extension of the Standard Model.

Names	spin 1/2	spin 1	$SU(3)_C, SU(2)_L, U(1)_Y$
gluino, gluon	\tilde{g}	g	(8 , 1 , 0)
winos, W bosons	$\tilde{W}^\pm \tilde{W}^0$	$W^\pm W^0$	(1 , 3 , 0)
bino, B boson	\tilde{B}^0	B^0	(1 , 1 , 0)

Table 2: Gauge supermultiplets in the Minimal Supersymmetric Standard Model.

General Remarks on the MSSM and Supersymmetry Breaking

The chiral and gauge supermultiplets in Tables 1 and 2 make up the particle content of the Minimal Supersymmetric Standard Model (MSSM).

The most obvious and interesting feature of this theory is that none of the superpartners of the Standard Model particles has been

discovered.

If supersymmetry were unbroken, then there would have to be selectrons \tilde{e}_L and \tilde{e}_R with masses exactly equal to $m_e = 0.511\dots$ MeV. A similar statement applies to each of the other sleptons and squarks, and there would also have to be a massless gluino and photino. These particles would have been extraordinarily easy to detect long ago. Clearly, therefore, *supersymmetry is a broken symmetry* in the vacuum state chosen by Nature.

An important clue as to the nature of supersymmetry breaking can be obtained by returning to the motivation provided by the hierarchy problem.

Supersymmetry forced us to introduce two complex scalar fields for each Standard Model Dirac fermion, which is just what is needed to enable a cancellation of the quadratically divergent (Λ_{UV}^2) one-loop corrections to the Higgs mass. This sort of cancellation also requires

that the associated dimensionless couplings should be related (for example $\lambda_S = |\lambda_f|^2$).

The necessary relationships between couplings indeed occur in unbroken supersymmetry. In fact, unbroken supersymmetry guarantees that the quadratic divergences in scalar squared masses must vanish to all orders in perturbation theory.³

Now, if broken supersymmetry is still to provide a solution to the hierarchy problem even in the presence of supersymmetry breaking, then the relationships between dimensionless couplings that hold in an unbroken supersymmetric theory must be maintained. Otherwise, there would be quadratically divergent radiative corrections to the

³A simple way to understand this is to recall that unbroken supersymmetry requires the degeneracy of scalar and fermion masses. Radiative corrections to fermion masses are known to diverge at most logarithmically in any renormalizable field theory, so the same must be true for scalar masses in unbroken supersymmetry.

Higgs scalar masses of the form

$$\Delta m_H^2 = \frac{1}{8\pi^2}(\lambda_S - |\lambda_f|^2)\Lambda_{UV}^2 + \dots \quad (252)$$

We are therefore led to consider “soft” supersymmetry breaking. This means that the effective Lagrangian of the MSSM can be written in the form

$$\mathcal{L} = \mathcal{L}_{\text{SUSY}} + \mathcal{L}_{\text{soft}}, \quad (253)$$

where $\mathcal{L}_{\text{SUSY}}$ contains all of the gauge and Yukawa interactions and preserves supersymmetry invariance, and $\mathcal{L}_{\text{soft}}$ violates supersymmetry but contains only mass terms and coupling parameters with *positive* mass dimension.

Without further justification, soft supersymmetry breaking might seem like a rather arbitrary requirement. Fortunately, theoretical models for supersymmetry breaking do indeed yield effective Lagrangians with just such terms for $\mathcal{L}_{\text{soft}}$.

If the largest mass scale associated with the soft terms is denoted m_{soft} , then the additional non-supersymmetric corrections to the Higgs scalar squared mass must vanish in the $m_{\text{soft}} \rightarrow 0$ limit, so by dimensional analysis they cannot be proportional to Λ_{UV}^2 .

More generally, these models maintain the cancellation of quadratically divergent terms in the radiative corrections of all scalar masses, to all orders in perturbation theory.

The corrections also cannot go like $\Delta m_H^2 \sim m_{\text{soft}} \Lambda_{\text{UV}}$, because in general the loop momentum integrals always diverge either quadratically or logarithmically, not linearly, as $\Lambda_{\text{UV}} \rightarrow \infty$. So they must be of the form

$$\Delta m_H^2 = m_{\text{soft}}^2 \left[\frac{\lambda}{16\pi^2} \ln(\Lambda_{\text{UV}}/m_{\text{soft}}) + \dots \right]. \quad (254)$$

Here λ is schematic for various dimensionless couplings, and the ellipses stand both for terms that are independent of Λ_{UV} and for higher loop corrections (which depend on Λ_{UV} through powers of

logarithms).

Because the mass splittings between the known Standard Model particles and their superpartners are just determined by the parameters m_{soft} appearing in $\mathcal{L}_{\text{soft}}$, Eq. (254) tells us that the superpartner masses cannot be too huge. Otherwise, we would lose our successful cure for the hierarchy problem, since the m_{soft}^2 corrections to the Higgs scalar squared mass parameter would be unnaturally large compared to the square of the electroweak breaking scale of 174 GeV.

The top and bottom squarks and the winos and bino give especially large contributions to $\Delta m_{H_u}^2$ and $\Delta m_{H_d}^2$, but the gluino mass and all the other squark and slepton masses also feed in indirectly, through radiative corrections to the top and bottom squark masses.

Furthermore, in most viable models of supersymmetry breaking that are not unduly contrived, the superpartner masses do not differ from each other by more than about an order of magnitude. Using $\Lambda_{\text{UV}} \sim$

M_{P} and $\lambda \sim 1$ in Eq. (254), one finds that m_{soft} , and therefore the masses of at least the lightest few superpartners, should be at the most about 1 TeV or so, in order for the MSSM scalar potential to provide a Higgs VEV resulting in $m_W, m_Z = 80.4, 91.2$ GeV without miraculous cancellations.

This is the best reason for the optimism among many theorists that supersymmetry will be discovered at the Fermilab Tevatron or the CERN Large Hadron Collider, and can be studied at a future e^+e^- linear collider.

However, it should be noted that the hierarchy problem was *not* the historical motivation for the development of supersymmetry in the early 1970's. The supersymmetry algebra and supersymmetric field theories were originally concocted independently in various disguises bearing little resemblance to the MSSM.

It is quite impressive that a theory developed for quite different

reasons, including purely aesthetic ones, can later be found to provide a solution for the hierarchy problem.

One might also wonder whether there is any good reason why all of the superpartners of the Standard Model particles should be heavy enough to have avoided discovery so far.

There is. All of the particles in the MSSM that have been found so far have something in common; they would necessarily be massless in the absence of electroweak symmetry breaking. In particular, the masses of the W^\pm, Z^0 bosons and all quarks and leptons are equal to dimensionless coupling constants times the Higgs VEV ~ 174 GeV, while the photon and gluon are required to be massless by electromagnetic and QCD gauge invariance.

Conversely, all of the undiscovered particles in the MSSM have exactly the opposite property; each of them can have a Lagrangian mass term in the absence of electroweak symmetry breaking.

For the squarks, sleptons, and Higgs scalars this follows from a general property of complex scalar fields that a mass term $m^2|\phi|^2$ is always allowed by all gauge symmetries.

For the higgsinos and gauginos, it follows from the fact that they are fermions in a real representation of the gauge group.

So, from the point of view of the MSSM, the discovery of the top quark in 1995 marked a quite natural milestone; the already-discovered particles are precisely those that had to be light, based on the principle of electroweak gauge symmetry.

There is a single exception: one neutral Higgs scalar boson should be lighter than about 135 GeV if the minimal version of supersymmetry is correct. In non-minimal models that do not have extreme fine tuning of parameters, and that remain perturbative up to the scale of apparent gauge coupling unification, the lightest Higgs scalar boson can have a mass up to about 150 GeV.

An important feature of the MSSM is that the superpartners listed in Tables 1 and 2 are not necessarily the mass eigenstates of the theory.

This is because after electroweak symmetry breaking and supersymmetry breaking effects are included, there can be mixing between the electroweak gauginos and the higgsinos, and within the various sets of squarks and sleptons and Higgs scalars that have the same electric charge.

The lone exception is the gluino, which is a color octet fermion and therefore does not have the appropriate quantum numbers to mix with any other particle.

The masses and mixings of the superpartners are obviously of paramount importance to experimentalists. It is perhaps slightly less obvious that these phenomenological issues are all quite directly related to one central question that is also the focus of much of the theoretical work in supersymmetry: “How is supersymmetry broken?”

The reason for this is that most of what we do not already know about the MSSM has to do with $\mathcal{L}_{\text{soft}}$. The structure of supersymmetric Lagrangians allows little arbitrariness, as we will see later.

In fact, all of the dimensionless couplings and all but one mass term in the supersymmetric part of the MSSM Lagrangian correspond directly to parameters in the ordinary Standard Model that have already been measured by experiment.

For example, we will find out that the supersymmetric coupling of a gluino to a squark and a quark is determined by the QCD coupling constant α_S .

In contrast, the supersymmetry-breaking part of the Lagrangian contains many unknown parameters and, apparently, a considerable amount of arbitrariness.

Each of the mass splittings between Standard Model particles and their superpartners correspond to terms in the MSSM Lagrangian

that are purely supersymmetry-breaking in their origin and effect.

These soft supersymmetry-breaking terms can also introduce a large number of mixing angles and CP-violating phases not found in the Standard Model.

Fortunately, as we will later show, there is already strong evidence that the supersymmetry-breaking terms in the MSSM are actually not arbitrary at all. Furthermore, the additional parameters will be measured and constrained as the superpartners are detected.

From a theoretical perspective, the challenge is to explain all of these parameters with a predictive model for supersymmetry breaking.

Conventions, Spinors, etc.

The conventions for supersymmetry, where we talk about Weyl spinors and so forth, have now become quite standard. They are best specified by showing how they correspond to the four-component

spinor language.

A four-component Dirac fermion Ψ_D with mass M is described by the Lagrangian

$$\mathcal{L}_{\text{Dirac}} = i\bar{\Psi}_D \gamma^\mu \partial_\mu \Psi_D - M\bar{\Psi}_D \Psi_D. \quad (255)$$

For our purposes it is convenient to use the specific representation of the 4×4 gamma matrices given in 2×2 blocks by

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}, \quad \gamma_5 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (256)$$

where

$$\begin{aligned} \sigma^0 &= \bar{\sigma}^0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, & \sigma^1 &= -\bar{\sigma}^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \sigma^2 &= -\bar{\sigma}^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, & \sigma^3 &= -\bar{\sigma}^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned} \quad (257)$$

In this representation, a four-component Dirac spinor is written in terms of 2 two-component, complex, anticommuting objects ξ_α and $(\chi^\dagger)^{\dot{\alpha}} \equiv \chi^{\dagger\dot{\alpha}}$ with two distinct types of spinor indices $\alpha = 1, 2$ and $\dot{\alpha} = 1, 2$:

$$\Psi_D = \begin{pmatrix} \xi_\alpha \\ \chi^{\dagger\dot{\alpha}} \end{pmatrix}. \quad (258)$$

It follows that

$$\bar{\Psi}_D = \Psi_D^\dagger \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = (\chi^\alpha \quad \xi_{\dot{\alpha}}^\dagger). \quad (259)$$

Undotted (dotted) indices from the beginning of the Greek alphabet are used for the first (last) two components of a Dirac spinor. The field ξ is called a “left-handed Weyl spinor” and χ^\dagger is a “right-handed Weyl spinor”. The names fit, because

$$P_L \Psi_D = \begin{pmatrix} \xi_\alpha \\ 0 \end{pmatrix}, \quad P_R \Psi_D = \begin{pmatrix} 0 \\ \chi^{\dagger\dot{\alpha}} \end{pmatrix}. \quad (260)$$

The Hermitian conjugate of any left-handed Weyl spinor is a right-handed Weyl spinor:

$$\psi_{\dot{\alpha}}^{\dagger} \equiv (\psi_{\alpha})^{\dagger} = (\psi^{\dagger})_{\dot{\alpha}}, \quad (261)$$

and vice versa:

$$(\psi^{\dagger\dot{\alpha}})^{\dagger} = \psi^{\alpha}. \quad (262)$$

Therefore, any particular fermionic degrees of freedom can be described equally well using a left-handed Weyl spinor (with an undotted index) or by a right-handed one (with a dotted index). By convention, all names of fermion fields are chosen so that left-handed Weyl spinors do not carry daggers and right-handed Weyl spinors do carry daggers, as in Eq. (258).

The heights of the dotted and undotted spinor indices are important; for example, comparing Eqs. (255)-(259), we observe that the matrices $(\sigma^{\mu})_{\alpha\dot{\alpha}}$ and $(\bar{\sigma}^{\mu})^{\dot{\alpha}\alpha}$ defined by Eq. (257) carry indices with

the heights as indicated. The spinor indices are raised and lowered using the antisymmetric symbol $\epsilon^{12} = -\epsilon^{21} = \epsilon_{21} = -\epsilon_{12} = 1$, $\epsilon_{11} = \epsilon_{22} = \epsilon^{11} = \epsilon^{22} = 0$, according to

$$\xi_\alpha = \epsilon_{\alpha\beta}\xi^\beta, \quad \xi^\alpha = \epsilon^{\alpha\beta}\xi_\beta, \quad \chi_{\dot{\alpha}}^\dagger = \epsilon_{\dot{\alpha}\dot{\beta}}\chi^{\dagger\dot{\beta}}, \quad \chi^{\dagger\dot{\alpha}} = \epsilon^{\dot{\alpha}\dot{\beta}}\chi_{\dot{\beta}}^\dagger. \quad (263)$$

This is consistent since $\epsilon_{\alpha\beta}\epsilon^{\beta\gamma} = \epsilon^{\gamma\beta}\epsilon_{\beta\alpha} = \delta_\alpha^\gamma$ and $\epsilon_{\dot{\alpha}\dot{\beta}}\epsilon^{\dot{\beta}\dot{\gamma}} = \epsilon^{\dot{\gamma}\dot{\beta}}\epsilon_{\dot{\beta}\dot{\alpha}} = \delta_{\dot{\alpha}}^{\dot{\gamma}}$.

As a convention, repeated spinor indices contracted like

$$\begin{matrix} \alpha \\ \alpha \end{matrix} \quad \text{or} \quad \begin{matrix} \dot{\alpha} \\ \dot{\alpha} \end{matrix} \quad (264)$$

can be suppressed. In particular,

$$\xi\chi \equiv \xi^\alpha\chi_\alpha = \xi^\alpha\epsilon_{\alpha\beta}\chi^\beta = -\chi^\beta\epsilon_{\alpha\beta}\xi^\alpha = \chi^\beta\epsilon_{\beta\alpha}\xi^\alpha = \chi^\beta\xi_\beta \equiv \chi\xi \quad (265)$$

with, conveniently, no minus sign in the end. [A minus sign appeared in Eq. (265) from exchanging the order of anticommuting spinors, but it disappeared due to the antisymmetry of the ϵ symbol.] Likewise, $\xi^\dagger \chi^\dagger$ and $\chi^\dagger \xi^\dagger$ are equivalent abbreviations for $\chi_{\dot{\alpha}}^\dagger \xi^{\dagger\dot{\alpha}} = \xi_{\dot{\alpha}}^\dagger \chi^{\dagger\dot{\alpha}}$, and in fact this is the complex conjugate of $\xi\chi$:

$$\xi^\dagger \chi^\dagger = \chi^\dagger \xi^\dagger = (\xi\chi)^*. \quad (266)$$

In a similar way, one can check that

$$\xi_{\dot{\alpha}}^\dagger \bar{\sigma}^\mu \chi = -\chi \sigma^\mu \xi^\dagger = (\chi^\dagger \bar{\sigma}^\mu \xi)^* = -(\xi \sigma^\mu \chi^\dagger)^* \quad (267)$$

stands for $\xi_{\dot{\alpha}}^\dagger (\bar{\sigma}^\mu)^{\dot{\alpha}\alpha} \chi_\alpha$, etc. The anti-commuting spinors here are taken to be classical fields; for quantum fields the complex conjugation in the last two equations would be replaced by Hermitian conjugation in the Hilbert space operator sense.

Some other identities that will be useful below include:

$$\xi \sigma^\mu \bar{\sigma}^\nu \chi = \chi \sigma^\nu \bar{\sigma}^\mu \xi = (\chi^\dagger \bar{\sigma}^\nu \sigma^\mu \xi^\dagger)^* = (\xi^\dagger \bar{\sigma}^\mu \sigma^\nu \chi^\dagger)^*, \quad (268)$$

and the Fierz rearrangement identity:

$$\chi_\alpha (\xi \eta) = -\xi_\alpha (\eta \chi) - \eta_\alpha (\chi \xi), \quad (269)$$

and the reduction identities

$$\sigma_{\alpha\dot{\alpha}}^\mu \bar{\sigma}_\mu^{\dot{\beta}\beta} = -2\delta_\alpha^\beta \delta_{\dot{\alpha}}^{\dot{\beta}}, \quad (270)$$

$$\sigma_{\alpha\dot{\alpha}}^\mu \sigma_{\mu\beta\dot{\beta}} = -2\epsilon_{\alpha\beta} \epsilon_{\dot{\alpha}\dot{\beta}}, \quad (271)$$

$$\bar{\sigma}^{\mu\dot{\alpha}\alpha} \bar{\sigma}_\mu^{\dot{\beta}\beta} = -2\epsilon^{\alpha\beta} \epsilon^{\dot{\alpha}\dot{\beta}}, \quad (272)$$

$$[\sigma^\mu \bar{\sigma}^\nu + \sigma^\nu \bar{\sigma}^\mu]_\alpha^\beta = -2\eta^{\mu\nu} \delta_\alpha^\beta, \quad (273)$$

$$[\bar{\sigma}^\mu \sigma^\nu + \bar{\sigma}^\nu \sigma^\mu]^{\dot{\beta}}_{\dot{\alpha}} = -2\eta^{\mu\nu} \delta_{\dot{\alpha}}^{\dot{\beta}}, \quad (274)$$

$$\bar{\sigma}^\mu \sigma^\nu \bar{\sigma}^\rho = -\eta^{\mu\nu} \bar{\sigma}^\rho - \eta^{\nu\rho} \bar{\sigma}^\mu + \eta^{\mu\rho} \bar{\sigma}^\nu + i\epsilon^{\mu\nu\rho\kappa} \bar{\sigma}_\kappa, \quad (275)$$

$$\sigma^\mu \bar{\sigma}^\nu \sigma^\rho = -\eta^{\mu\nu} \sigma^\rho - \eta^{\nu\rho} \sigma^\mu + \eta^{\mu\rho} \sigma^\nu - i\epsilon^{\mu\nu\rho\kappa} \sigma_\kappa, \quad (276)$$

where $\epsilon^{\mu\nu\rho\kappa}$ is the totally antisymmetric tensor with $\epsilon^{0123} = +1$.

With these conventions, the Dirac Lagrangian Eq. (255) can now be rewritten:

$$\mathcal{L}_{\text{Dirac}} = i\xi^\dagger \bar{\sigma}^\mu \partial_\mu \xi + i\chi^\dagger \bar{\sigma}^\mu \partial_\mu \chi - M(\xi\chi + \xi^\dagger \chi^\dagger) \quad (277)$$

where we have dropped a total derivative piece $-i\partial_\mu(\chi^\dagger \bar{\sigma}^\mu \chi)$, which does not affect the action.

A four-component Majorana spinor can be obtained from the Dirac spinor of Eq. (259) by imposing the constraint $\chi = \xi$, so that

$$\Psi_M = \begin{pmatrix} \xi_\alpha \\ \xi^{\dagger\dot{\alpha}} \end{pmatrix}, \quad \bar{\Psi}_M = (\xi^\alpha \quad \xi^\dagger_{\dot{\alpha}}). \quad (278)$$

The four-component spinor form of the Lagrangian for a Majorana fermion with mass M ,

$$\mathcal{L}_{\text{Majorana}} = \frac{i}{2} \bar{\Psi}_M \gamma^\mu \partial_\mu \Psi_M - \frac{1}{2} M \bar{\Psi}_M \Psi_M \quad (279)$$

can therefore be rewritten as

$$\mathcal{L}_{\text{Majorana}} = i \xi^\dagger \bar{\sigma}^\mu \partial_\mu \xi - \frac{1}{2} M (\xi \xi + \xi^\dagger \xi^\dagger) \quad (280)$$

in the more economical two-component Weyl spinor representation. Note that even though ξ_α is anticommuting, $\xi \xi$ and its complex conjugate $\xi^\dagger \xi^\dagger$ do not vanish, because of the suppressed ϵ symbol, see Eq. (265). Explicitly, $\xi \xi = \epsilon^{\alpha\beta} \xi_\beta \xi_\alpha = \xi_2 \xi_1 - \xi_1 \xi_2 = 2 \xi_2 \xi_1$.

More generally, any theory involving spin-1/2 fermions can always be written in terms of a collection of left-handed Weyl spinors ψ_i with

$$\mathcal{L} = i \psi_i^\dagger \bar{\sigma}^\mu \partial_\mu \psi_i + \dots \quad (281)$$

where the ellipses represent possible mass terms, gauge interactions, and Yukawa interactions with scalar fields. Here the index i runs over the appropriate gauge and flavor indices of the fermions; it is raised or lowered by Hermitian conjugation. Gauge interactions are obtained by promoting the ordinary derivative to a gauge-covariant derivative:

$$\mathcal{L} = i\psi^{\dagger i}\bar{\sigma}^{\mu}D_{\mu}\psi_i + \dots \quad (282)$$

with

$$D_{\mu}\psi_i = \partial_{\mu}\psi_i - ig_a A_{\mu}^a T_i^{aj}\psi_j, \quad (283)$$

where g_a is the gauge coupling corresponding to the Hermitian Lie algebra generator matrix T^a with vector field A_{μ}^a .

There is a different ψ_i for the left-handed piece and for the hermitian conjugate of the right-handed piece of a Dirac fermion. Given any

expression involving bilinears of four-component spinors

$$\Psi_i = \begin{pmatrix} \xi_i \\ \chi_i^\dagger \end{pmatrix}, \quad (284)$$

labeled by a flavor or gauge-representation index i , one can translate into two-component Weyl spinor language (or vice versa) using the dictionary:

$$\bar{\Psi}_i P_L \Psi_j = \chi_i \xi_j, \quad \bar{\Psi}_i P_R \Psi_j = \xi_i^\dagger \chi_j^\dagger, \quad (285)$$

$$\bar{\Psi}_i \gamma^\mu P_L \Psi_j = \xi_i^\dagger \bar{\sigma}^\mu \xi_j, \quad \bar{\Psi}_i \gamma^\mu P_R \Psi_j = \chi_i \sigma^\mu \chi_j^\dagger \quad (286)$$

etc.

Let us now see how the Standard Model quarks and leptons are described in this notation. The complete list of left-handed Weyl spinors can be given names corresponding to the chiral supermultiplets

in Table 1:

$$Q_i = \begin{pmatrix} u \\ d \end{pmatrix}, \begin{pmatrix} c \\ s \end{pmatrix}, \begin{pmatrix} t \\ b \end{pmatrix}, \quad (287)$$

$$\bar{u}_i = \bar{u}, \bar{c}, \bar{t}, \quad (288)$$

$$\bar{d}_i = \bar{d}, \bar{s}, \bar{b} \quad (289)$$

$$L_i = \begin{pmatrix} \nu_e \\ e \end{pmatrix}, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}, \quad (290)$$

$$\bar{e}_i = \bar{e}, \bar{\mu}, \bar{\tau}. \quad (291)$$

Here $i = 1, 2, 3$ is a family index. The bars on these fields are part of the names of the fields, and do *not* denote any kind of conjugation. Rather, the unbarred fields are the left-handed pieces of a Dirac spinor, while the barred fields are the names given to the conjugates of the right-handed piece of a Dirac spinor. For example, e is the same thing as e_L in Table 1, and \bar{e} is the same as e_R^\dagger . Together they

form a Dirac spinor:

$$\begin{pmatrix} e \\ \bar{e}^\dagger \end{pmatrix} \equiv \begin{pmatrix} e_L \\ e_R \end{pmatrix}, \quad (292)$$

and similarly for all of the other quark and charged lepton Dirac spinors. (The neutrinos of the Standard Model are not part of a Dirac spinor, at least in the approximation that they are massless.) The fields Q_i and L_i are weak isodoublets, which always go together when one is constructing interactions invariant under the full Standard Model gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$. Suppressing all color and weak isospin indices, the kinetic and gauge part of the Standard Model fermion Lagrangian density is then

$$\mathcal{L} = iQ^{\dagger i} \bar{\sigma}^\mu D_\mu Q_i + i\bar{u}_i^\dagger \bar{\sigma}^\mu D_\mu \bar{u}^i + i\bar{d}_i^\dagger \bar{\sigma}^\mu D_\mu \bar{d}^i + iL^{\dagger i} \bar{\sigma}^\mu D_\mu L_i + i\bar{e}_i^\dagger \bar{\sigma}^\mu D_\mu \bar{e}^i \quad (293)$$

with the family index i summed over, and D_μ the appropriate

Standard Model covariant derivative. For example,

$$D_\mu \begin{pmatrix} \nu_e \\ e \end{pmatrix} = [\partial_\mu - igW_\mu^a(\tau^a/2) - ig'Y_L B_\mu] \begin{pmatrix} \nu_e \\ e \end{pmatrix} \quad (294)$$

$$D_\mu \bar{e} = [\partial_\mu - ig'Y_{\bar{e}} B_\mu] \bar{e} \quad (295)$$

with τ^a ($a = 1, 2, 3$) equal to the Pauli matrices, $Y_L = -1/2$ and $Y_{\bar{e}} = +1$. The gauge eigenstate weak bosons are related to the mass eigenstates by

$$W_\mu^\pm = (W_\mu^1 \mp iW_\mu^2)/\sqrt{2}, \quad (296)$$

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_W & -\sin \theta_W \\ \sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}. \quad (297)$$

Similar expressions hold for the other quark and lepton gauge eigenstates, with $Y_Q = 1/6$, $Y_{\bar{u}} = -2/3$, and $Y_{\bar{d}} = 1/3$. The quarks also have a term in the covariant derivative corresponding to gluon

interactions proportional to g_3 (with $\alpha_S = g_3^2/4\pi$) with generators $T^a = \lambda^a/2$ for Q , and in the complex conjugate representation $T^a = -(\lambda^a)^*/2$ for \bar{u} and \bar{d} , where λ^a are the Gell-Mann matrices.

Constructing a Supersymmetric Lagrangian

As already describe, one begins with two kinds of “supermultiplets”. By supersymmetry, these should have an equal number of bosonic and fermionic degrees of freedom, both on-shell (*i.e.* after using equations of motion) and off-shell (*i.e.* before using equations of motion).

Chiral Supermultiplet

On-shell the propagating degrees of freedom are a complex scalar field ϕ (2 dof) and a left-handed two component Weyl spinor ψ (2 dof).

Off-shell, the ψ has two *complex* components, *i.e.* 4 dof, and

supersymmetry requires two more scalar dof. These reside in a complex “auxiliary” field, F .

	ϕ	ψ	F
on-shell ($n_B = n_F = 2$)	2	2	0
off-shell ($n_B = n_F = 4$)	2	4	2

Table 3: Counting of real degrees of freedom in the Wess-Zumino model.

An auxiliary field does not yield a propagating degree of freedom since its equation of motion is such that there are no derivatives. It is only present in order to close the supersymmetry transformation laws “off-shell”.

As we have learned, chiral supermultiplets are used for matter fields (e.g. quarks, leptons, higgs bosons, ...)

The Superpotential for chiral supermultiplets

For the case of chiral fields and their interactions, one writes down what is called a “superpotential”. Using an index i for different chiral fields, the most general form of the superpotential is

$$W = L^i \phi_i + \frac{1}{2} M^{ij} \phi_i \phi_j + \frac{1}{6} y^{ijk} \phi_i \phi_j \phi_k, \quad (298)$$

where the M^{ij} and y^{ijk} are totally symmetric in their indices.

Note the absence of ϕ^* 's above. W is not a scalar potential in the usual sense. It's not even real. Instead, it is an analytic function of the scalar fields treated as complex variables. This is a requirement for the Lagrangian form obtained from the above superpotential using the rules sketched below to be invariant under supersymmetry transformations.

The L^i parameters have dimensions of $[\text{mass}]^2$, the M^{ij} have dimension $[\text{mass}]$, and the y^{ijk} are dimensionless.

The L^i affect only the scalar potential (*i.e.* that for the spin-0 component fields) part of the Lagrangian. Such linear terms are only allowed when ϕ_i is a gauge singlet.

There are no such gauge singlet chiral supermultiplets in the MSSM with minimal field content, but this term can be present in the NMSSM.

Further, this type of term does play an important role in the discussion of spontaneous supersymmetry breaking.

For the moment, I will omit it.

You could ask why not go to terms with four or more chiral fields. The reason is that the coefficients would then have to have dimensions of [mass] to some inverse power, which would have the implication that the quantum field theory loop corrections would be non-renormalizable.

The next step is to compute

$$W^i = \frac{\delta W}{\delta \phi_i} = M^{ij} \phi_j + \frac{1}{2} y^{ijk} \phi_j \phi_k. \quad (299)$$

and

$$W^{ij} = \frac{\delta W^i}{\delta \phi_j} = M^{ij} + y^{ijk} \phi_k. \quad (300)$$

The Lagrangian density for chiral supermultiplets

A Lagrangian density invariant under supersymmetry transforms is then constructed as:

$$\mathcal{L} = -\partial^\mu \phi^{*i} \partial_\mu \phi_i + i \psi^{\dagger i} \bar{\sigma}^\mu \partial_\mu \psi_i - \frac{1}{2} (W^{ij} \psi_i \psi_j + W_{ij}^* \psi^{\dagger i} \psi^{\dagger j}) - W^i W_i^*. \quad (301)$$

It follows from Eq. (301) that the scalar potential for the theory is just given in terms of the superpotential by

$$V(\phi, \phi^*) = W^k W_k^*$$

$$\begin{aligned}
&= M_{ik}^* M^{kj} \phi^{*i} \phi_j + \frac{1}{2} M^{in} y_{jkn}^* \phi_i \phi^{*j} \phi^{*k} + \frac{1}{2} M_{in}^* y^{jkn} \phi^{*i} \phi_j \phi_k \\
&\quad + \frac{1}{4} y^{ijn} y_{kln}^* \phi_i \phi_j \phi^{*k} \phi^{*l}.
\end{aligned} \tag{302}$$

This scalar potential is automatically bounded from below; in fact, since it is a sum of squares of absolute values (of the W^k), it is always non-negative. If (dropping L^i) we substitute the general form for the superpotential Eq. (298) into Eq. (301), we obtain for the full Lagrangian density

$$\begin{aligned}
\mathcal{L} = & -\partial^\mu \phi^{*i} \partial_\mu \phi_i - V(\phi, \phi^*) + i\psi^{\dagger i} \bar{\sigma}^\mu \partial_\mu \psi_i - \frac{1}{2} M^{ij} \psi_i \psi_j - \frac{1}{2} M_{ij}^* \psi^{\dagger i} \psi^{\dagger j} \\
& - \frac{1}{2} y^{ijk} \phi_i \psi_j \psi_k - \frac{1}{2} y_{ijk}^* \phi^{*i} \psi^{\dagger j} \psi^{\dagger k}.
\end{aligned} \tag{303}$$

Now we can compare the masses of the fermions and scalars by looking at the linearized equations of motion:

$$\partial^\mu \partial_\mu \phi_i = M_{ik}^* M^{kj} \phi_j + \dots, \tag{304}$$

$$i\bar{\sigma}^\mu \partial_\mu \psi_i = M_{ij}^* \psi^{\dagger j} + \dots, \quad i\sigma^\mu \partial_\mu \psi^{\dagger i} = M^{ij} \psi_j + \dots \quad (305)$$

One can eliminate ψ in terms of ψ^\dagger and vice versa in Eq. (305), obtaining [after use of the identities Eqs. (273) and (274)]:

$$\partial^\mu \partial_\mu \psi_i = M_{ik}^* M^{kj} \psi_j + \dots, \quad \partial^\mu \partial_\mu \psi^{\dagger j} = \psi^{\dagger i} M_{ik}^* M^{kj} + \dots \quad (306)$$

Therefore, the fermions and the bosons satisfy the same wave equation with exactly the same squared-mass matrix with real non-negative eigenvalues, namely $(M^2)_i^j = M_{ik}^* M^{kj}$.

It follows that diagonalizing this matrix by redefining the fields with a unitary matrix gives a collection of chiral supermultiplets, each of which contains a mass-degenerate complex scalar and Weyl fermion, in agreement with the general argument in the Introduction.

Vector or Gauge Supermultiplet

The propagating degrees of freedom in the Wess Zumino gauge are

a massless gauge boson field A_μ^a and a two-component Weyl fermion gaugino λ^a . Again there is also an auxiliary field, D^a .

Here, a runs over the adjoint representation of the gauge group: $a = 1, \dots, 8$ for $SU(3)_c$ colored gluons and gluinos; $a = 1, 2, 3$ for $SU(2)_L$ weak isospin $W^{1,2,3}$ gauge bosons and associated winos; $a = 1$ for $U(1)_Y$ weak hypercharge B gauge boson and associated bino.

The on shell degrees of freedom for A_μ^a and λ_α^a comprise two bosonic and two fermionic helicity states (for each a), as required by SUSY.

However, off-shell λ_α^a consists of two complex, or four real, fermionic degrees of freedom, while A_μ^a has only three real bosonic degrees of freedom (one being removed by gauge transformations). This is why we need one real bosonic auxiliary field, D^a , in order for supersymmetry to be consistent off-shell. As for the F field of the chiral multiplet, D^a does not correspond to a propagating degree of

freedom and can be eliminated on-shell using its algebraic equation of motion. The counting of degrees of freedom appears in Table 4.

	A_μ	λ	D
on-shell ($n_B = n_F = 2$)	2	2	0
off-shell ($n_B = n_F = 4$)	3	4	1

Table 4: Counting of real degrees of freedom for each gauge supermultiplet.

Gauge transformations of the fields are given by:

$$\delta_{\text{gauge}} A_\mu^a = \partial_\mu \Lambda^a + g f^{abc} A_\mu^b \Lambda^c, \quad (307)$$

$$\delta_{\text{gauge}} \lambda^a = g f^{abc} \lambda^b \Lambda^c, \quad (308)$$

where Λ^a is an infinitesimal gauge transformation parameter, g is the gauge coupling, and f^{abc} are the totally antisymmetric structure constants that define the gauge group. The special case of an

Abelian group is obtained by just setting $f^{abc} = 0$; the corresponding gaugino is a gauge singlet in that case. The conventions are such that for QED, $A^\mu = (V, \vec{A})$ where V and \vec{A} are the usual electric potential and vector potential, with electric and magnetic fields given by $\vec{E} = -\vec{\nabla}V - \partial_0\vec{A}$ and $\vec{B} = \vec{\nabla} \times \vec{A}$.

Lagrangians for gauge supermultiplets

Therefore, the Lagrangian density for a gauge supermultiplet ought to be

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}F_{\mu\nu}^a F^{\mu\nu a} + i\lambda^{\dagger a}\bar{\sigma}^\mu D_\mu\lambda^a + \frac{1}{2}D^a D^a, \quad (309)$$

where

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf^{abc}A_\mu^b A_\nu^c \quad (310)$$

is the usual Yang-Mills field strength, and

$$D_\mu\lambda^a = \partial_\mu\lambda^a + gf^{abc}A_\mu^b\lambda^c \quad (311)$$

is the covariant derivative of the gaugino field.

If we had not included the auxiliary field D^a , then the supersymmetry algebra would hold only after using the equations of motion for λ^a and $\lambda^{\dagger a}$.

The auxiliary fields satisfy a trivial equation of motion $D^a = 0$, but this is modified if one couples the gauge supermultiplets to chiral supermultiplets, as we now do.

A “Derivation” of the Scalar Lagrangian

While we could just proceed by using the above rules, I want to provide some idea of how at least the scalar/chiral field Lagrangian is obtained.

One approach is to consider an extension of ordinary space and time in which Grassman coordinates $\theta_{1,2}$ and $\bar{\theta}_{1,2}$ are added to the coordinate system. These Grassman coordinate anticommute with one another and, as a result, $\theta_\alpha \theta_\alpha = 0$ for any fixed α . As a result, if

one Taylor expands any function $f(x, \theta, \bar{\theta})$ in powers of the θ_α , then the power series quickly terminates:

$$F(x, \theta, \bar{\theta}) = f(x) + \theta\phi(x) + \bar{\theta}\bar{\chi}(x) + \theta\theta m(x) + \bar{\theta}\bar{\theta}n(x) + \theta\sigma^\mu\bar{\theta}v_\mu(x) + \theta\theta\bar{\theta}\bar{\lambda}(x) + \bar{\theta}\bar{\theta}\theta\psi(x) + \theta\theta\bar{\theta}\bar{\theta}d(x). \quad (312)$$

Here, an expression like $\theta\theta$ corresponds to $\theta^\alpha\theta^\beta\epsilon_{\alpha\beta}$ where $\alpha, \beta = 1, 2$, and similar for $\theta\phi$ etc.

A chiral superfield Φ is one that is a function of only θ and $y^\mu = x^\mu + i\theta\sigma^\mu\bar{\theta}$. Its Taylor expansion then looks like (for clarity, I temporarily use A instead of Martin's notation of ϕ , which is easily confused with Φ)

$$\begin{aligned} \Phi &= A(y) + \sqrt{2}\theta\psi(y) + \theta\theta F(y) \\ &= A(x) + i\theta\sigma^\mu\bar{\theta}\partial_\mu A(x) + \frac{1}{4}\theta\theta\bar{\theta}\bar{\theta}\square A(x) + \sqrt{2}\theta\psi(x) \\ &\quad - \frac{i}{\sqrt{2}}\theta\theta\partial_\mu\psi(x)\sigma^\mu\bar{\theta} + \theta\theta F(x). \end{aligned} \quad (313)$$

The conjugate field Φ^\dagger then looks like

$$\Phi^\dagger = A^*(y^\dagger) + \sqrt{2}\bar{\theta}\bar{\psi}(y^\dagger) + \bar{\theta}\bar{\theta}F^*(y^\dagger). \quad (314)$$

Notice that a product of 2 chiral superfields is automatically another chiral superfield since terms like $\theta_\alpha\theta\theta$ automatically contain $\theta_\alpha\theta_\alpha = 0$.

For example,

$$\begin{aligned} \Phi_i\Phi_j = & A_i(y)A_j(y) + \sqrt{2}\theta[\psi_i(y)A_j(y) + A_i(y)\psi_j(y)] \\ & + \theta\theta[A_i(y)F_j(y) + F_i(y)A_j(y) - \psi_i(y)\psi_j(y)]. \end{aligned} \quad (315)$$

Now, generators of supersymmetry transformations look like

$$Q_\alpha = \frac{\partial}{\partial\theta^\alpha} - i\sigma_{\alpha\dot{\alpha}}^\mu\bar{\theta}^{\dot{\alpha}}\partial_\mu. \quad (316)$$

Thus, it is apparent that when applied to a chiral superfield, each of the components will be shifted. However, since the $\theta\theta$ term is

the maximum power in θ possible, the F component can only be shifted by a total derivative (from the 2nd term in Q_α acting on the $\sqrt{2}\theta[\dots]$ term in a general Φ). Thus, the $\theta\theta$ component of a product of superfields is a candidate for the Lagrangian density.

The other possible thing you can do with two chiral superfields is to form

$$\begin{aligned} \Phi_i^\dagger(y^\dagger)\Phi_j(y) = & \dots + \theta\theta\bar{\theta}\bar{\theta}[F_i^* F_j + \frac{1}{4}A_i^*\square A_j + \frac{1}{4}\square A_i^* A_j - \frac{1}{2} \\ & \partial_\mu A_i^* \partial^\mu A_j + \frac{i}{2}\partial_\mu \bar{\psi}_i \bar{\sigma}^\mu \psi_j - \frac{i}{2}\bar{\psi}_i \bar{\sigma}^\mu \partial_\mu \psi_j] \quad (317) \end{aligned}$$

where all the fields are functions of x after expanding out their dependence on y . Once again, under supersymmetry transformations the $\theta\theta\bar{\theta}\bar{\theta}$ term can only be changed by a total derivative and is thus a candidate for a Lagrangian density.

In this respect, you can see that if we only keep $i = j$ terms of

this type then this component will have the (diagonal) kinetic energy terms that we want for free A and ψ fields.

With the above background we are now able to write down a candidate Lagrangian density that is invariant (up to a total derivative) under supersymmetry transformations:

$$\begin{aligned}
 \mathcal{L} &= \Phi_i^\dagger \Phi_i |_{\theta\theta\bar{\theta}\bar{\theta}} \text{ component} + \left[\left(\frac{1}{2} M^{ij} \Phi_i \Phi_j \right. \right. \\
 &\quad \left. \left. + \frac{1}{6} y^{ijk} \Phi_i \Phi_j \Phi_k + L^i \phi_i \right) \Big|_{\theta\theta} \text{ component} + h.c. \right] \\
 &= i \partial_\mu \bar{\psi}_i \bar{\sigma}^\mu \psi_i + A_i^* \square A_i + F_i^* F_i + \left[M^{ij} \left(A_i F_j - \frac{1}{2} \psi_i \psi_j \right) \right. \\
 &\quad \left. + \frac{1}{2} y^{ijk} (A_i A_j F_k - \psi_i \psi_j A_k) + L^i F_i + h.c. \right] \quad (318)
 \end{aligned}$$

where we have used the symmetry of the M^{ij} and y^{ijk} under index interchange. At this point, the auxiliary fields F_i which have no

kinetic energy terms can be eliminated through their Euler equations:

$$\begin{aligned}\frac{\partial \mathcal{L}}{\partial F_k^*} &= F_k + L^{*k} + M^{*ik} A_i + \frac{1}{2} y^{*ijk} A_i^* A_j^* = 0 \\ \frac{\partial \mathcal{L}}{\partial F_k} &= F_k^* + L^k + M^{ik} A_i + \frac{1}{2} y^{ijk} A_i A_j = 0\end{aligned}\quad (319)$$

yielding a \mathcal{L} expressed solely in terms of the dynamical fields A_i and ψ_i :

$$\begin{aligned}\mathcal{L} &= i\partial_\mu \bar{\psi}_i \bar{\sigma}^\mu \psi_i + A_i^* \square A_i - \frac{1}{2} M^{ik} \psi_i \psi_k - \frac{1}{2} M^{*ik} \bar{\psi}_i \bar{\psi}_k \\ &\quad - \frac{1}{2} y^{ijk} \psi_i \psi_j A_k - \frac{1}{2} y^{*ijk} \bar{\psi}_i \bar{\psi}_j A_k^* - \mathcal{V}(A_i, A_j^*),\end{aligned}\quad (320)$$

where the scalar field potential \mathcal{V} takes the simple form

$$\mathcal{V} = F_k^* F_k, \quad (321)$$

where F, F^* are the solutions to the equations in Eq. (319).

Converting to the Martin notation of $A \rightarrow \phi$, we get the previously stated result, Eq. (303), for the scalar field \mathcal{L} after recognizing that the solutions to Eq. (319) are that $F_k = -W^k, F_k^* = -W^{*k}$ with the W 's as given in Martin.

The procedure in the case of the vector superfield is analogous, but more complicated. There, one defines a superfield $V = V^\dagger$ with the most general form (after going to the “Wess-Zumino” gauge):

$$V(x, \theta, \bar{\theta}) = -\theta\sigma^\mu\bar{\theta}A_\mu(x) + i\theta\theta\bar{\theta}\bar{\lambda}(x) - i\bar{\theta}\bar{\theta}\theta\lambda(x) + \frac{1}{2}\theta\theta\bar{\theta}\bar{\theta}D(x) \quad (322)$$

From V , one extracts a chiral field W^α and then one forms

$$\mathcal{L} = \frac{1}{4} \left(W^\alpha W_\alpha|_{\theta\theta} + \bar{W}_{\dot{\alpha}} \bar{W}^{\dot{\alpha}}|_{\bar{\theta}\bar{\theta}} \right). \quad (323)$$

Again, the $\theta\theta$ and $\bar{\theta}\bar{\theta}$ components are invariant under supersymmetry transformations and so this is a candidate Lagrangian. After some partial integration, we get

$$\int d^4x \mathcal{L} = \int d^4x \left(\frac{1}{2} D^2 - \frac{1}{4} F^{\mu\nu} F_{\mu\nu} - i\lambda\sigma^\mu \partial_\mu \bar{\lambda} \right). \quad (324)$$

In the absence of interactions between the vector and chiral supermultiplet one would then use the simple Euler-Lagrange equation for D to give $D = 0$.

Supersymmetric gauge interactions

Now we are ready to consider a general Lagrangian density for a supersymmetric theory with both chiral and gauge supermultiplets. Suppose that the chiral supermultiplets transform under the gauge group in a representation with hermitian matrices $(T^a)_i^j$ satisfying $[T^a, T^b] = if^{abc}T^c$. [For example, if the gauge group is $SU(2)$,

then $f^{abc} = \epsilon^{abc}$, and the T^a are $1/2$ times the Pauli matrices for a chiral supermultiplet transforming in the fundamental representation.] Since supersymmetry and gauge transformations commute, the scalar, fermion, and auxiliary fields must be in the same representation of the gauge group, so

$$\delta_{\text{gauge}} X_i = ig\Lambda^a (T^a X)_i \quad (325)$$

for $X_i = \phi_i, \psi_i, F_i$. To have a gauge-invariant Lagrangian, we now need to replace the ordinary derivatives in

$$\mathcal{L}_{\text{free}} = -\partial^\mu \phi^{*i} \partial_\mu \phi_i + i\psi^{\dagger i} \bar{\sigma}^\mu \partial_\mu \psi_i + F^{*i} F_i, \quad (326)$$

with covariant derivatives:

$$\partial_\mu \phi_i \rightarrow D_\mu \phi_i = \partial_\mu \phi_i - igA_\mu^a (T^a \phi)_i \quad (327)$$

$$\partial_\mu \phi^{*i} \rightarrow D_\mu \phi^{*i} = \partial_\mu \phi^{*i} + igA_\mu^a (\phi^* T^a)^i \quad (328)$$

$$\partial_\mu \psi_i \rightarrow D_\mu \psi_i = \partial_\mu \psi_i - ig A_\mu^a (T^a \psi)_i. \quad (329)$$

Naively, this simple procedure achieves the goal of coupling the vector bosons in the gauge supermultiplet to the scalars and fermions in the chiral supermultiplets. However, we also have to consider whether there are any other interactions allowed by gauge invariance and involving the gaugino and D^a fields, which might have to be included to make a supersymmetric Lagrangian. Since A_μ^a couples to ϕ_i and ψ_i , it makes sense that λ^a and D^a should as well.

In fact, there are three such possible interaction terms that are renormalizable (of field mass dimension ≤ 4), namely

$$(\phi^* T^a \psi) \lambda^a, \quad \lambda^{\dagger a} (\psi^\dagger T^a \phi), \quad \text{and} \quad (\phi^* T^a \phi) D^a. \quad (330)$$

Now one can add them, with unknown dimensionless coupling coefficients, to the Lagrangians for the chiral and gauge supermultiplets,

and demand that the whole mess be real and invariant under supersymmetry transformations, up to a total derivative.

To fix the coefficients above, we must consider details regarding supersymmetry transformations, and at the same time I will give some different justifications for the final form of the chiral superfield Lagrangian.

Supersymmetry transformations

Our starting point is the Lagrangian density for a collection of free chiral supermultiplets labeled by an index i , which runs over all gauge and flavor degrees of freedom.

Since we will want to construct an interacting theory with supersymmetry closing off-shell, each supermultiplet contains a complex scalar ϕ_i and a left-handed Weyl fermion ψ_i as physical degrees of freedom, plus a complex auxiliary field F_i , which does not propagate. The free part

of the Lagrangian is

$$\mathcal{L}_{\text{free}} = -\partial^\mu \phi^{*i} \partial_\mu \phi_i + i\psi^{\dagger i} \bar{\sigma}^\mu \partial_\mu \psi_i + F^{*i} F_i, \quad (331)$$

where we sum over repeated indices i (not to be confused with the suppressed spinor indices), with the convention that fields ϕ_i and ψ_i always carry lowered indices, while their conjugates always carry raised indices.

It is invariant under the supersymmetry transformations (note: ϵ is an infinitesimal, anticommuting parameter and transformations take spin 0 to spin 1/2 or vice-versa)

$$\delta\phi_i = \epsilon\psi_i, \quad \delta\phi^{*i} = \epsilon^\dagger\psi^{\dagger i}, \quad (332)$$

$$\delta(\psi_i)_\alpha = -i(\sigma^\mu\epsilon^\dagger)_\alpha \partial_\mu\phi_i + \epsilon_\alpha F_i, \quad \delta(\psi^{\dagger i})_{\dot{\alpha}} = i(\epsilon\sigma^\mu)_{\dot{\alpha}} \partial_\mu\phi^{*i} + \epsilon_{\dot{\alpha}}^\dagger F^{*i}, \quad (333)$$

$$\delta F_i = -i\epsilon^\dagger \bar{\sigma}^\mu \partial_\mu \psi_i, \quad \delta F^{*i} = i\partial_\mu \psi^{\dagger i} \bar{\sigma}^\mu \epsilon. \quad (334)$$

We will now find the most general set of renormalizable interactions for these fields that is consistent with supersymmetry. We do this working in the field theory before integrating out the auxiliary fields.

To begin, note that in order to be renormalizable by power counting, each term must have field content with total mass dimension ≤ 4 . So, the only candidate terms are:

$$\mathcal{L}_{\text{int}} = \left(-\frac{1}{2}W^{ij}\psi_i\psi_j + W^i F_i + x^{ij} F_i F_j \right) + \text{c.c.} - U, \quad (335)$$

where W^{ij} , W^i , x^{ij} , and U are polynomials in the scalar fields ϕ_i, ϕ^{*i} , with degrees 1, 2, 0, and 4, respectively (to obey $[\text{mass}]^{\leq 4}$). [Terms of the form $F^{*i} F_j$ are already included in $\mathcal{L}_{\text{free}}$ of Eq. (331), with the coefficient fixed by requiring invariance of $\mathcal{L}_{\text{free}}$ under the transformation rules (332)-(334).]

We must now require that \mathcal{L}_{int} is invariant under the supersymmetry transformations, since $\mathcal{L}_{\text{free}}$ was already invariant by itself.

This immediately requires that the candidate term $U(\phi_i, \phi^{*i})$ must vanish. If there were such a term, then under a supersymmetry transformation Eq. (332) it would transform into another function of the scalar fields only, multiplied by $\epsilon\psi_i$ or $\epsilon^\dagger\psi^\dagger i$, and with no spacetime derivatives or F_i , F^{*i} fields. It is easy to see from Eqs. (332)-(335) that nothing of this form can possibly be canceled by the supersymmetry transformation of any other term in the Lagrangian. Similarly, the dimensionless coupling x^{ij} must be zero, because its supersymmetry transformation likewise cannot possibly be canceled by any other term.

So, we are left with

$$\mathcal{L}_{\text{int}} = \left(-\frac{1}{2}W^{ij}\psi_i\psi_j + W^i F_i \right) + \text{c.c.} \quad (336)$$

as the only possibilities. At this point, we are not assuming that W^{ij} and W^i are related to each other in any way. However, soon we will

find out that they *are* related, which is why we have chosen to use the same letter for them. Notice that Eq. (265), *i.e.* $\psi_i\psi_j = \psi_j\psi_i$, tells us that W^{ij} is symmetric under $i \leftrightarrow j$.

It is easiest to divide the variation of \mathcal{L}_{int} into several parts, which must cancel separately. The four-spinor terms are:

$$\delta\mathcal{L}_{\text{int}}|_{4\text{-spinor}} = \left[-\frac{1}{2} \frac{\delta W^{ij}}{\delta\phi_k} (\epsilon\psi_k)(\psi_i\psi_j) - \frac{1}{2} \frac{\delta W^{ij}}{\delta\phi^{*k}} (\epsilon^\dagger\psi^\dagger{}^k)(\psi_i\psi_j) \right] + \text{c.c.} \quad (337)$$

The term proportional to $(\epsilon\psi_k)(\psi_i\psi_j)$ cannot cancel against any other term. Fortunately, however, the Fierz identity Eq. (269) implies

$$(\epsilon\psi_i)(\psi_j\psi_k) + (\epsilon\psi_j)(\psi_k\psi_i) + (\epsilon\psi_k)(\psi_i\psi_j) = 0, \quad (338)$$

so this contribution to $\delta\mathcal{L}_{\text{int}}$ vanishes identically *if and only if* $\delta W^{ij}/\delta\phi_k$ is totally symmetric under interchange of i, j, k .

There is no such identity available for the term proportional to

$(\epsilon^\dagger \psi^\dagger{}^k)(\psi_i \psi_j)$. Since that term cannot cancel with any other, requiring it to be absent just tells us that W^{ij} cannot contain ϕ^{*k} . In other words, W^{ij} is analytic (or holomorphic) in the complex fields ϕ_k .

Combining what we have learned so far (including $W^{ij} \propto [\text{mass}]^1$), we get the already stated result

$$W^{ij} = M^{ij} + y^{ijk} \phi_k \quad (339)$$

where M^{ij} is a symmetric mass matrix for the fermion fields, and y^{ijk} is a Yukawa coupling of a scalar ϕ_k and two fermions $\psi_i \psi_j$ that must be totally symmetric under interchange of i, j, k . This means that we get the result we already know:

$$W^{ij} = \frac{\delta^2}{\delta\phi_i \delta\phi_j} W \quad (340)$$

in terms of the superpotential form given earlier:

$$W = \frac{1}{2}M^{ij}\phi_i\phi_j + \frac{1}{6}y^{ijk}\phi_i\phi_j\phi_k, \quad (341)$$

To repeat, W is not a scalar potential in the ordinary sense; in fact, it is not even real. It is instead an analytic function of the scalar fields ϕ_i treated as complex variables.

Continuing the quest for total invariance under a supersymmetry transformation, one must next consider the parts of $\delta\mathcal{L}_{\text{int}}$ that contain a spacetime derivative — these come from the $\delta\psi$ and δF :

$$\delta\mathcal{L}_{\text{int}}|_{\partial} = (iW^{ij}\partial_{\mu}\phi_j\psi_i\sigma^{\mu}\epsilon^{\dagger} + iW^i\partial_{\mu}\psi_i\sigma^{\mu}\epsilon^{\dagger}) + \text{c.c.} \quad (342)$$

Here we have used the identity Eq. (267) on the second term, which

came from $(\delta F_i)W^i$. Now we can use Eq. (340) to observe that

$$W^{ij}\partial_\mu\phi_j = \partial_\mu\left(\frac{\delta W}{\delta\phi_i}\right). \quad (343)$$

Therefore, Eq. (342) will be a total derivative if

$$W^i = \frac{\delta W}{\delta\phi_i} = M^{ij}\phi_j + \frac{1}{2}y^{ijk}\phi_j\phi_k, \quad (344)$$

which explains why we chose its name as we did. The remaining terms in $\delta\mathcal{L}_{\text{int}}$ are all linear in F_i or F^{*i} , and it is easy to show that they cancel, given the results for W^i and W^{ij} that we have already found.

Back to inclusion of gauge interactions

Now, after introducing gauge interactions, it is presumably not surprising that invariance under supersymmetry transformations is

possible only if the supersymmetry transformation laws for the matter fields are modified to include gauge-covariant rather than ordinary derivatives. Also, it is necessary to include one strategically chosen extra term in δF_i , so in the presence of gauge interactions our supersymmetry transformations look like:

$$\delta\phi_i = \epsilon\psi_i \quad (345)$$

$$\delta\psi_{i\alpha} = -i(\sigma^\mu\epsilon^\dagger)_\alpha D_\mu\phi_i + \epsilon_\alpha F_i \quad (346)$$

$$\delta F_i = -i\epsilon^\dagger\bar{\sigma}^\mu D_\mu\psi_i + \sqrt{2}g(T^a\phi)_i \epsilon^\dagger\lambda^{\dagger a}. \quad (347)$$

As we said in class, the extra terms of Eq. (330) need to be present **with correct coefficients** in order to get invariance under supersymmetry transformations as defined above. After some algebra to determine the correct coefficients, the full Lagrangian density for a renormalizable supersymmetric theory is

$$\mathcal{L} = \mathcal{L}_{\text{chiral}} + \mathcal{L}_{\text{gauge}} - \sqrt{2}g(\phi^*T^a\psi)\lambda^a - \sqrt{2}g\lambda^{\dagger a}(\psi^\dagger T^a\phi) + g(\phi^*T^a\phi)D^a. \quad (348)$$

Here $\mathcal{L}_{\text{chiral}}$ means the chiral supermultiplet Lagrangian found earlier [e.g., Eq. (301) or (303)], but with ordinary derivatives replaced everywhere by gauge-covariant derivatives, and $\mathcal{L}_{\text{gauge}}$ was given in Eq. (357).

To prove that Eq. (356) is invariant under the supersymmetry transformations, one must in particular have

$$W^i(\delta F_i)_{2nd \text{ term Eq. (347)}} = 0, \quad (349)$$

which requires the identity

$$W^i(T^a \phi)_i = 0. \quad (350)$$

This is precisely the condition that must be satisfied anyway in order for the superpotential, and thus $\mathcal{L}_{\text{chiral}}$, to be gauge invariant, since the left side is proportional to $\delta_{\text{gauge}} W$.

Thus, the second line in Eq. (356) consists of interactions that must be present and whose strengths are fixed to be gauge couplings by the requirements of supersymmetry, even though they are not gauge interactions from the point of view of an ordinary field theory.

The first two terms are a direct coupling of gauginos to matter fields; this can be thought of as the “supersymmetrization” of the usual gauge boson couplings to matter fields $(\phi^* T^a \overleftrightarrow{\partial}^\mu \phi A_\mu^a)$.

The last term combines with the $D^a D^a / 2$ term in $\mathcal{L}_{\text{gauge}}$ to provide an equation of motion

$$D^a = -g(\phi^* T^a \phi). \quad (351)$$

Thus, like the auxiliary fields F_i and F^{*i} , the D^a are expressible purely algebraically in terms of the scalar fields. Replacing the auxiliary fields in Eq. (356) using Eq. (351), one finds that the

complete scalar potential is (recall that \mathcal{L} contains $-V$):

$$V(\phi, \phi^*) = F^{*i}F_i + \frac{1}{2} \sum_a D^a D^a = W_i^* W^i + \frac{1}{2} \sum_a g_a^2 (\phi^* T^a \phi)^2. \quad (352)$$

The two types of terms in this expression are called “ F -term” and “ D -term” contributions, respectively.

In the second term in Eq. (352), we have now written an explicit sum \sum_a to cover the case that the gauge group has several distinct factors with different gauge couplings g_a . [For instance, in the MSSM the three factors $SU(3)_C$, $SU(2)_L$ and $U(1)_Y$ have different gauge couplings g_3 , g and g' .]

Since $V(\phi, \phi^*)$ is a sum of squares, it is always greater than or equal to zero for every field configuration.

It is an interesting and unique feature of supersymmetric theories that the scalar potential is completely determined by the *other* interactions

in the theory. The F -terms are fixed by Yukawa couplings and fermion mass terms, and the D -terms are fixed by the gauge interactions.

Summary: How to build a supersymmetric model

In a renormalizable supersymmetric field theory, the interactions and masses of all particles are determined just by their gauge transformation properties and by the superpotential W . By construction, we found that W had to be an analytic function of the complex scalar fields ϕ_i , which are always defined to transform under supersymmetry into left-handed Weyl fermions.

As we saw, in an equivalent language, W is a function of chiral *superfields*. In the superfield notation, the derivation of all of our preceding results can be obtained in a fairly elegant manner using superfield methods, which have the advantage of making invariance under supersymmetry transformations manifest by defining the Lagrangian in terms of a specific procedure in a “superspace”

with fermionic as well as ordinary commuting coordinates.

In the superfield formulation, one writes instead of Eq. (298)

$$W = L^i \Phi_i + \frac{1}{2} M^{ij} \Phi_i \Phi_j + \frac{1}{6} y^{ijk} \Phi_i \Phi_j \Phi_k, \quad (353)$$

However, the familiar and accessible component field approach is most appropriate for making contact with phenomenology in a universe with supersymmetry breaking. The only (occasional) use we will make of superfield notation is the purely cosmetic one of following the common practice of specifying superpotentials like Eq. (353) rather than (298).

The specification of the superpotential is really a code for the terms that it implies in the Lagrangian, so the reader may feel free to think of the superpotential either as a function of the scalar fields ϕ_i or as the same function of the superfields Φ_i .

Given the supermultiplet content of the theory, the form of the superpotential is restricted by the requirement of gauge invariance [see Eq. (350)].

In any given theory, only a subset of the parameters L^i , M^{ij} , and y^{ijk} are allowed to be non-zero.

1. The parameter L^i is only allowed if Φ_i is a gauge singlet. (There are no such chiral supermultiplets in the MSSM with the minimal field content.)
2. The entries of the mass matrix M^{ij} can only be non-zero for i and j such that the supermultiplets Φ_i and Φ_j transform under the gauge group in representations that are conjugates of each other. (In the MSSM there is only one such term, as we will see.)
3. Likewise, the Yukawa couplings y^{ijk} can only be non-zero when Φ_i , Φ_j , and Φ_k transform in representations that can combine to form a singlet.

The interactions implied by the superpotential Eq. (353) (with $L^i = 0$) were listed in Eqs. (302), (303), repeated below,

$$\begin{aligned} \mathcal{L} = & -\partial^\mu \phi^{*i} \partial_\mu \phi_i - V(\phi, \phi^*) + i\psi^{\dagger i} \bar{\sigma}^\mu \partial_\mu \psi_i - \frac{1}{2} M^{ij} \psi_i \psi_j - \frac{1}{2} M_{ij}^* \psi^{\dagger i} \psi^{\dagger j} \\ & - \frac{1}{2} y^{ijk} \phi_i \psi_j \psi_k - \frac{1}{2} y_{ijk}^* \phi^{*i} \psi^{\dagger j} \psi^{\dagger k}. \end{aligned} \quad (354)$$

with

$$\begin{aligned} V(\phi, \phi^*) = W^k W_k^* = & M_{ik}^* M^{kj} \phi^{*i} \phi_j + \frac{1}{2} M^{in} y_{jkn}^* \phi_i \phi^{*j} \phi^{*k} + \frac{1}{2} M_{in}^* y^{jkn} \phi^{*i} \phi_j \phi_k \\ & + \frac{1}{4} y^{ijn} y_{klm}^* \phi_i \phi_j \phi^{*k} \phi^{*l}. \end{aligned} \quad (355)$$

and are shown⁴ in Figures 5 and 6.

Those in Figure 5 are all determined by the dimensionless parameters

⁴Here, the auxiliary fields have been eliminated using their equations of motion (“integrated out”). One could instead give Feynman rules that include the auxiliary fields, or directly in terms of superfields on superspace, although this is usually less useful in practical phenomenological applications.

y^{ijk} . The Yukawa interaction in Figure 5a corresponds to the next-to-last term in Eq. (354).

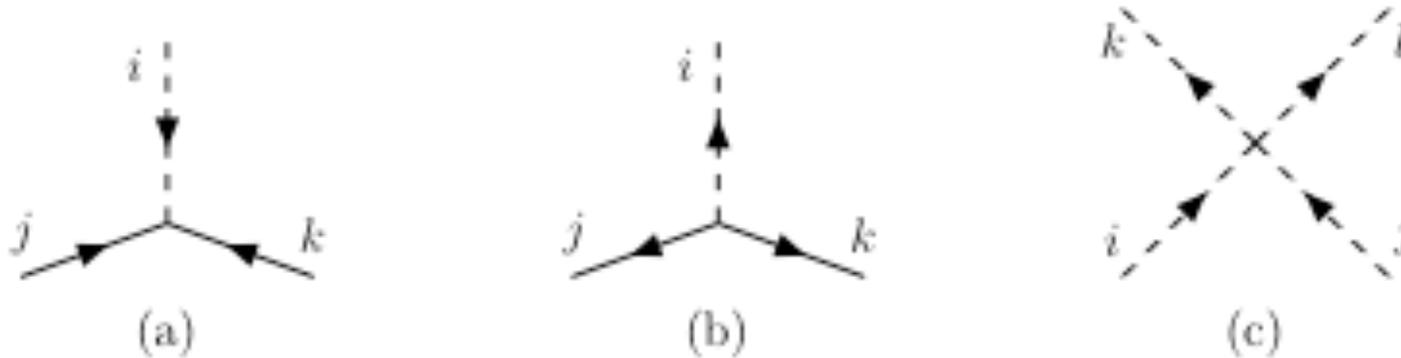


Figure 5: The dimensionless non-gauge interaction vertices in a supersymmetric theory: (a) scalar-fermion-fermion Yukawa interaction y^{ijk} , (b) the complex conjugate interaction y_{ijk} , and (c) quartic scalar interaction $y^{ijn}y_{klm}^*$.

For each particular Yukawa coupling of $\phi_i\psi_j\psi_k$ with strength y^{ijk} , there must be equal couplings of $\phi_j\psi_i\psi_k$ and $\phi_k\psi_i\psi_j$, since y^{ijk} is completely symmetric under interchange of any two of its indices.

The arrows on the fermion and scalar lines point in the direction for

propagation of ϕ and ψ and opposite the direction of propagation of ϕ^* and ψ^\dagger . Thus there is also a vertex corresponding to the one in Figure 5a but with all arrows reversed, corresponding to the complex conjugate [the last term in Eq. (354)]. It is shown in Figure 5b.

There is also a dimensionless coupling for $\phi_i\phi_j\phi^{*k}\phi^{*l}$, with strength $y^{ijn}y_{kl n}^*$, as required by supersymmetry [see the last term in Eq. (355)].

The relationship between the Yukawa interactions in Figures 5a,b and the scalar interaction of Figure 5c is exactly of the special type needed to cancel the quadratic divergences in quantum corrections to scalar masses, as discussed with regard to Eq. (252)].

Figure 6 shows the only interactions corresponding to renormalizable and supersymmetric vertices with coupling dimensions of [mass] and [mass]².

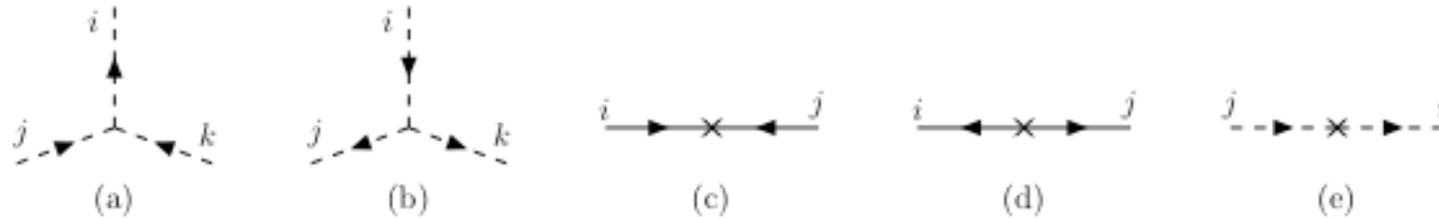


Figure 6: Supersymmetric dimensionful couplings: (a) (scalar)³ interaction vertex $M_{in}^* y^{jkn}$ and (b) the conjugate interaction $M^{in} y_{jkn}^*$, (c) fermion mass term M^{ij} and (d) conjugate fermion mass term M_{ij}^* , and (e) scalar squared-mass term $M_{ik}^* M^{kj}$.

First, there are (scalar)³ couplings in Figure 6a,b, which are entirely determined by the superpotential mass parameters M^{ij} and Yukawa couplings y^{ijk} , as indicated by the second and third terms in Eq. (355).

The propagators of the fermions and scalars in the theory are constructed in the usual way using the fermion mass M^{ij} and scalar squared mass $M_{ik}^* M^{kj}$. The fermion mass terms M^{ij} and M_{ij}^* each lead to a chirality-changing insertion in the fermion propagator; note

the directions of the arrows in Figure 6c,d.

There is no such arrow-reversal for a scalar propagator in a theory with exact supersymmetry; as depicted in Figure 6e, if one treats the scalar squared-mass term as an insertion in the propagator, the arrow direction is preserved.

For the gauge-related interactions, we return to the earlier equations:

$$\mathcal{L} = \mathcal{L}_{\text{chiral}} + \mathcal{L}_{\text{gauge}} - \sqrt{2}g(\phi^* T^a \psi)\lambda^a - \sqrt{2}g\lambda^{\dagger a}(\psi^\dagger T^a \phi) + g(\phi^* T^a \phi)D^a. \quad (356)$$

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4}F_{\mu\nu}^a F^{\mu\nu a} + i\lambda^{\dagger a}\bar{\sigma}^\mu D_\mu \lambda^a + \frac{1}{2}D^a D^a, \quad (357)$$

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf^{abc}A_\mu^b A_\nu^c \quad (358)$$

$$D_\mu \lambda^a = \partial_\mu \lambda^a + gf^{abc}A_\mu^b \lambda^c \quad (359)$$

$$\mathcal{L}_{\text{chiral}} = -D^\mu \phi^{*i} D_\mu \phi_i + i\psi^{\dagger i}\bar{\sigma}^\mu D_\mu \psi_i + F^{*i} F_i, \quad (360)$$

$$D_\mu \phi_i = \partial_\mu \phi_i - igA_\mu^a (T^a \phi)_i, \quad D_\mu \phi^{*i} = \partial_\mu \phi^{*i} + igA_\mu^a (\phi^* T^a)^i, \quad (361)$$

$$D_\mu \psi_i = \partial_\mu \psi_i - igA_\mu^a (T^a \psi)_i, \quad (362)$$

$$F_i \rightarrow -W^i, \quad F_i^* \rightarrow -W^{*i}, \quad D^a \rightarrow -g(\phi^* T^a \phi). \quad (363)$$

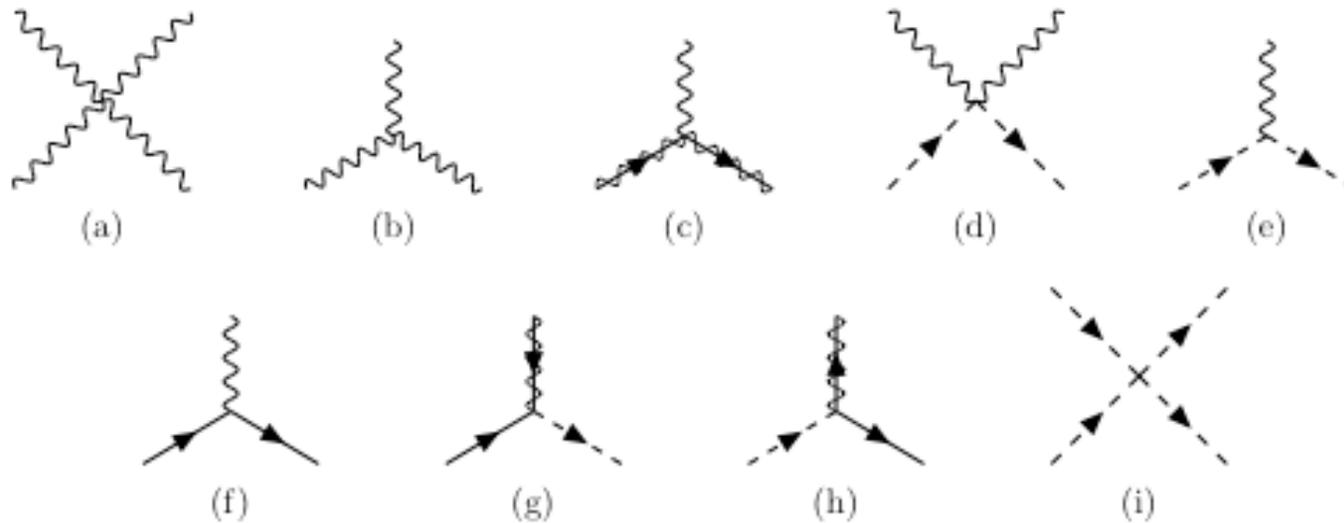


Figure 7: Supersymmetric gauge interaction vertices.

Figure 7 shows the gauge interactions in a supersymmetric theory. Figures 7a,b,c occur only when the gauge group is non-Abelian, for example for $SU(3)_C$ color and $SU(2)_L$ weak isospin in the MSSM.

Figures 7a and 7b are the interactions of gauge bosons, which derive from the first term in $\mathcal{L}_{\text{gauge}}$. In the MSSM these are exactly the same as the well-known QCD gluon and electroweak gauge boson

vertices of the Standard Model. (We do not show the interactions of ghost fields, which are necessary only for consistent loop amplitudes.)

Figures 7c,d,e,f are just the standard interactions between gauge bosons and fermion and scalar fields that must occur in any gauge theory because of the form of the covariant derivative; they come from Eqs. (359) and (361)-(362) inserted in the kinetic part of the Lagrangian. Figure 7c shows the coupling of a gaugino to a gauge boson [insert 2nd term of Eq. (359) into 2nd term of Eq. (357)]; the gaugino line in a Feynman diagram is traditionally drawn as a solid fermion line superimposed on a wavy line.

In Figure 7g we have the coupling of a gaugino to a chiral fermion and a complex scalar [the first of the extra terms in Eq. (356)]. One can think of this as the “supersymmetrization” of Figure 7e or 7f; any of these three vertices may be obtained from any other (up to a factor of $\sqrt{2}$) by replacing two of the particles by their supersymmetric partners.

There is also an interaction in Figure 7h which is just like Figure 7g but with all arrows reversed, corresponding to the complex conjugate term in the Lagrangian [the second term in the second line in Eq. (356)].

Finally, in Figure 7i we have a scalar quartic interaction vertex [the last term in Eq. (356) plus the $\frac{1}{2}D^a D^a$ term in Eq. (357) after using the eom, $D^a = -(\phi^* T^a \phi)$], which is also determined by the gauge coupling.

The results of this section can be used as a recipe for constructing the supersymmetric interactions for any model. In the case of the MSSM, we already know the gauge group, particle content and the gauge transformation properties, so it only remains to decide on the superpotential.

Soft supersymmetry breaking interactions

A realistic phenomenological model must contain supersymmetry

breaking. From a theoretical perspective, we expect that supersymmetry, if it exists at all, should be an exact symmetry that is broken spontaneously.

In other words, the underlying model should have a Lagrangian density that is invariant under supersymmetry, but a vacuum state that is not.

In this way, supersymmetry is hidden at low energies in a manner analogous to the fate of the electroweak symmetry in the ordinary Standard Model.

Many models of spontaneous symmetry breaking have indeed been proposed.

These always involve extending the MSSM to include new particles and interactions at very high mass scales, and **there is no consensus on exactly how this should be done.**

From a practical point of view, it is extremely useful to simply

parameterize our ignorance of these issues by just introducing extra terms that break supersymmetry explicitly in the effective MSSM Lagrangian.

As was argued earlier, the supersymmetry-breaking couplings should be soft (of positive mass dimension) in order to be able to naturally maintain a hierarchy between the electroweak scale and the Planck (or any other very large) mass scale. This means in particular that dimensionless supersymmetry-breaking couplings should be absent.

The possible soft supersymmetry-breaking terms in the Lagrangian of a general theory are

$$\mathcal{L}_{\text{soft}} = - \left(\frac{1}{2} M_a \lambda^a \lambda^a + \frac{1}{6} a^{ijk} \phi_i \phi_j \phi_k + \frac{1}{2} b^{ij} \phi_i \phi_j + t^i \phi_i \right) + \text{c.c.} - (m^2)_j^i \phi^{j*} \phi_i, \quad (364)$$

$$\mathcal{L}_{\text{maybe soft}} = - \frac{1}{2} c_i^{jk} \phi^{*i} \phi_j \phi_k + \text{c.c.} \quad (365)$$

They consist of gaugino masses M_a for each gauge group, scalar

squared-mass terms $(m^2)_i^j$ and b^{ij} , and (scalar)³ couplings a^{ijk} and c_i^{jk} , and “tadpole” couplings t^i .

The last of these can only occur if ϕ_i is a gauge singlet, and so is absent from the MSSM.

One might wonder why we have not included possible soft mass terms for the chiral supermultiplet fermions, like $\mathcal{L} = -\frac{1}{2}m^{ij}\psi_i\psi_j + \text{c.c.}$ Including such terms would be redundant; they can always be absorbed into a redefinition of the superpotential and the terms $(m^2)_j^i$ and c_i^{jk} .

It has been shown rigorously that a softly broken supersymmetric theory with $\mathcal{L}_{\text{soft}}$ as given by Eq. (364) is indeed free of quadratic divergences in quantum corrections to scalar masses, to all orders in perturbation theory.

The situation is slightly more subtle if one tries to include the non-analytic (scalar)³ couplings in $\mathcal{L}_{\text{maybe soft}}$. If any of the chiral

supermultiplets in the theory are singlets under all gauge symmetries, then non-zero c_i^{jk} terms can lead to quadratic divergences, despite the fact that they are formally soft.

Now, this constraint need not apply to the MSSM, which does not have any gauge-singlet chiral supermultiplets. Nevertheless, the possibility of c_i^{jk} terms is nearly always neglected. The real reason for this is that it is difficult to construct models of spontaneous supersymmetry breaking in which the c_i^{jk} are not negligibly small.

In the special case of a theory that has chiral supermultiplets that are singlets or in the adjoint representation of a simple factor of the gauge group, then there are also possible soft supersymmetry-breaking Dirac mass terms between the corresponding fermions ψ_a and the gauginos:

$$\mathcal{L} = -M_{\text{Dirac}}^a \lambda^a \psi_a + \text{c.c.} \quad (366)$$

This is not relevant for the MSSM with minimal field content, which does not have adjoint representation chiral supermultiplets.

Therefore, equation (364) is usually taken to be the general form of the soft supersymmetry-breaking Lagrangian.

The terms in $\mathcal{L}_{\text{soft}}$ clearly do break supersymmetry, because they involve only scalars and gauginos and not their respective superpartners.

In fact, the soft terms in $\mathcal{L}_{\text{soft}}$ are capable of giving masses to all of the scalars and gauginos in a theory, even if the gauge bosons and fermions in chiral supermultiplets are massless (or relatively light).

The gaugino masses M_a are always allowed by gauge symmetry.

The $(m^2)_j^i$ terms are allowed for i, j such that ϕ_i, ϕ^{j*} transform in complex conjugate representations of each other under all gauge symmetries; in particular this is true of course when $i = j$, so every scalar is eligible to get a mass in this way if supersymmetry is broken.

The remaining soft terms may or may not be allowed by the symmetries.

The a^{ijk} , b^{ij} , and t^i terms have the same form as the y^{ijk} , M^{ij} , and L^i terms in the superpotential [compare Eq. (364) to Eq. (298) or Eq. (353)], so they will each be allowed by gauge invariance if and only if a corresponding superpotential term is allowed.

The Feynman diagram interactions corresponding to the allowed soft terms in Eq. (364) are shown in Figure 8.

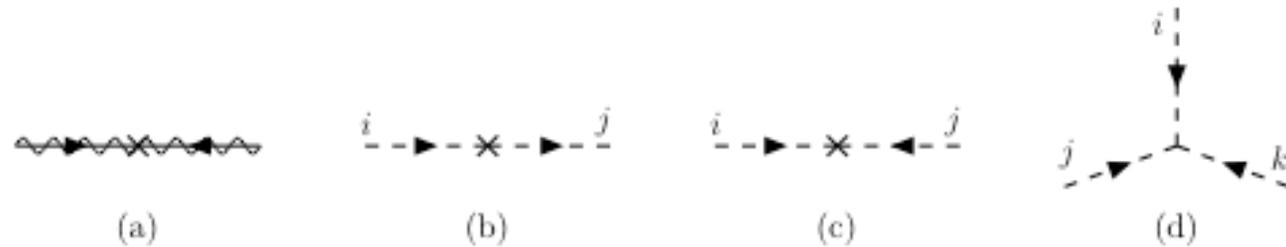


Figure 8: Soft supersymmetry-breaking terms: (a) Gaugino mass M_a ; (b) non-analytic scalar squared mass $(m^2)_j^i$; (c) analytic scalar squared mass b^{ij} ; and (d) scalar cubic coupling a^{ijk} .

For each of the interactions in Figures 8a,c,d there is another with all arrows reversed, corresponding to the complex conjugate term in the Lagrangian. We will apply these general results to the specific case of the MSSM in the next section.

The Minimal Supersymmetric Model

- The superpotential and supersymmetric interactions

The superpotential for the MSSM is

$$W_{\text{MSSM}} = \bar{u} \mathbf{y}_u Q H_u - \bar{d} \mathbf{y}_d Q H_d - \bar{e} \mathbf{y}_e L H_d + \mu H_u H_d. \quad (367)$$

The objects $H_u, H_d, Q, L, \bar{u}, \bar{d}, \bar{e}$ appearing here are chiral superfields corresponding to the chiral supermultiplets in Table 1. (Alternatively, they can be just thought of as the corresponding scalar fields, but we prefer not to put the tildes on $Q, L, \bar{u}, \bar{d}, \bar{e}$ in order to reduce clutter.)

The dimensionless Yukawa coupling parameters $\mathbf{y}_u, \mathbf{y}_d, \mathbf{y}_e$ are 3×3 matrices in family space.

All of the gauge [$SU(3)_C$ color and $SU(2)_L$ weak isospin] and family indices in Eq. (367) are suppressed.

The “ μ term”, as it is traditionally called, can be written out as $\mu(H_u)_\alpha(H_d)_\beta\epsilon^{\alpha\beta}$, where $\epsilon^{\alpha\beta}$ is used to tie together $SU(2)_L$ weak isospin indices $\alpha, \beta = 1, 2$ in a gauge-invariant way.

And, the term $\bar{u}_u \mathbf{y}_u Q H_u$ can be written out as $\bar{u}^{ia} (\mathbf{y}_u)_i^j Q_{j\alpha a} (H_u)_\beta \epsilon^{\alpha\beta}$, where $i = 1, 2, 3$ is a family index, and $a = 1, 2, 3$ is a color index which is lowered (raised) in the $\mathbf{3}$ ($\bar{\mathbf{3}}$) representation of $SU(3)_C$.

The μ term in Eq. (367) is the supersymmetric version of the Higgs boson mass in the Standard Model. It is unique, because terms $H_u^* H_u$ or $H_d^* H_d$ are forbidden in the superpotential, which must be analytic in the chiral superfields (or equivalently in the scalar fields) treated as complex variables.

We can also see from the form of Eq. (367) why both H_u and H_d are needed in order to give Yukawa couplings, and thus masses, to

all of the quarks and leptons. Since the superpotential must be analytic, the $\bar{u}QH_u$ Yukawa terms cannot be replaced by something like $\bar{u}QH_d^*$. Similarly, the $\bar{d}QH_d$ and $\bar{e}LH_d$ terms cannot be replaced by something like $\bar{d}QH_u^*$ and $\bar{e}LH_u^*$. The analogous Yukawa couplings would be allowed in a general non-supersymmetric two Higgs doublet model, but are forbidden by the structure of supersymmetry. So we need both H_u and H_d , even without invoking the argument based on anomaly cancellation mentioned earlier.

The Yukawa matrices determine the current masses and CKM mixing angles of the ordinary quarks and leptons, after the neutral scalar components of H_u and H_d get VEVs. Since the top quark, bottom quark and tau lepton are the heaviest fermions in the Standard Model, it is often useful to make an approximation that only the $(3, 3)$ family components of each of \mathbf{y}_u , \mathbf{y}_d and \mathbf{y}_e are important:

$$\mathbf{y}_u \approx \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & y_t \end{pmatrix}, \quad \mathbf{y}_d \approx \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & y_b \end{pmatrix}, \quad \mathbf{y}_e \approx \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & y_\tau \end{pmatrix}. \quad (368)$$

In this limit, only the third family and Higgs fields contribute to the MSSM superpotential.

It is instructive to write the superpotential in terms of the separate $SU(2)_L$ weak isospin components [$Q_3 = (tb)$, $L_3 = (\nu_\tau \tau)$, $H_u = (H_u^+ H_u^0)$, $H_d = (H_d^0 H_d^-)$, $\bar{u}_3 = \bar{t}$, $\bar{d}_3 = \bar{b}$, $\bar{e}_3 = \bar{\tau}$], so:

$$W_{\text{MSSM}} \approx y_t(\bar{t}tH_u^0 - \bar{t}bH_u^+) - y_b(\bar{b}tH_d^- - \bar{b}bH_d^0) - y_\tau(\bar{\tau}\nu_\tau H_d^- - \bar{\tau}\tau H_d^0) + \mu(H_u^+ H_d^- - H_u^0 H_d^0). \quad (369)$$

The minus signs inside the parentheses appear because of the antisymmetry of the $\epsilon^{\alpha\beta}$ symbol used to tie up the $SU(2)_L$ indices.

The other minus signs in Eq. (367) were chosen so that the terms $y_t\bar{t}tH_u^0$, $y_b\bar{b}bH_d^0$, and $y_\tau\bar{\tau}\tau H_d^0$, which will become the top, bottom and tau masses when H_u^0 and H_d^0 get VEVs, each have overall positive signs in Eq. (369).

Since the Yukawa interactions y^{ijk} in a general supersymmetric theory must be completely symmetric under interchange of i, j, k ,

we know that y_u , y_d and y_e imply not only Higgs-quark-quark and Higgs-lepton-lepton couplings as in the Standard Model, but also squark-Higgsino-quark and slepton-Higgsino-lepton interactions. To illustrate this, Figures 9a,b,c show some of the interactions involving the top-quark Yukawa coupling y_t .



Figure 9: The top-quark Yukawa coupling (a) and its “supersymmetrizations” (b), (c), all of strength y_t .

Figure 9a is the Standard Model-like coupling of the top quark to the neutral complex scalar Higgs boson, which follows from the first term in Eq. (369). For variety, we have used t_L and t_R^\dagger in place of

their synonyms t and \bar{t} .

In Figure 9b, we have the coupling of the left-handed top squark \tilde{t}_L to the neutral higgsino field \tilde{H}_u^0 and right-handed top quark.

In Figure 9c the right-handed top anti-squark field (known either as \tilde{t} or \tilde{t}_R^* depending on taste) couples to \tilde{H}_u^0 and t_L .

For each of the three interactions, there is another with $H_u^0 \rightarrow H_u^+$ and $t_L \rightarrow -b_L$ (with tildes where appropriate), corresponding to the second part of the first term in Eq. (369).

All of these interactions are required by supersymmetry to have the same strength y_t .

These couplings are dimensionless and can be modified by the introduction of soft supersymmetry breaking only through finite (and small) radiative corrections, so this equality of interaction strengths is also a prediction of softly broken supersymmetry.

If some SUSY-like signal is seen, it will be critical to test the equality of these interaction strengths.

A useful mnemonic is that each of Figures 9a,b,c can be obtained from any of the others by changing two of the particles into their superpartners.

There are also scalar quartic interactions with strength proportional to y_t^2 , as can be seen from Figure 5c or the last term in Eq. (302). Three of them are shown in Figure 10.

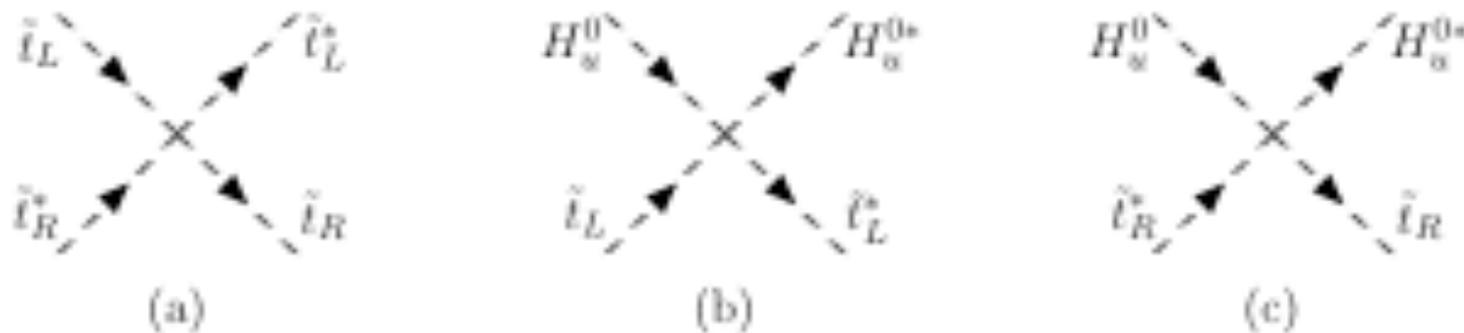


Figure 10: Some of the (scalar)⁴ interactions with strength proportional to y_t^2 .

Using Eq. (302) and Eq. (369), one can see that there are five more, which can be obtained by replacing $\tilde{t}_L \rightarrow \tilde{b}_L$ and/or $H_u^0 \rightarrow H_u^+$ in each vertex.

This illustrates the remarkable economy of supersymmetry; there are many interactions determined by only a single parameter.

In a similar way, the existence of all the other quark and lepton Yukawa couplings in the superpotential Eq. (367) leads not only to Higgs-quark-quark and Higgs-lepton-lepton Lagrangian terms as in the ordinary Standard Model, but also to squark-higgsino-quark and slepton-higgsino-lepton terms, and scalar quartic couplings

$[(\text{squark})^4, (\text{slepton})^4, (\text{squark})^2(\text{slepton})^2, (\text{squark})^2(\text{Higgs})^2, \text{ and } (\text{slepton})^2(\text{Higgs})^2]$.

If needed, these can all be obtained in terms of the Yukawa matrices \mathbf{y}_u , \mathbf{y}_d , and \mathbf{y}_e as outlined above.

However, the dimensionless interactions determined by the superpotentials

are usually not the most important ones of direct interest for phenomenology. This is because the Yukawa couplings are already known to be very small, except for those of the third family (top, bottom, tau).

Instead, production and decay processes for superpartners in the MSSM are typically dominated by the supersymmetric interactions of gauge-coupling strength.

The couplings of the Standard Model gauge bosons (photon, W^\pm , Z^0 and gluons) to the MSSM particles are determined completely by the gauge invariance of the kinetic terms in the Lagrangian.

The gauginos also couple to (squark, quark) and (slepton, lepton) and (Higgs, higgsino) pairs as illustrated in the general case in Figure 7g,h and the first two terms in the extra pieces in Eq. (356).

For instance, each of the squark-quark-gluino couplings is given by $\sqrt{2}g_3(\tilde{q}T^a q\tilde{g} + \text{c.c.})$ where $T^a = \lambda^a/2$ ($a = 1 \dots 8$) are the matrix

generators for $SU(3)_C$. The Feynman diagram for this interaction is shown in Figure 11a.

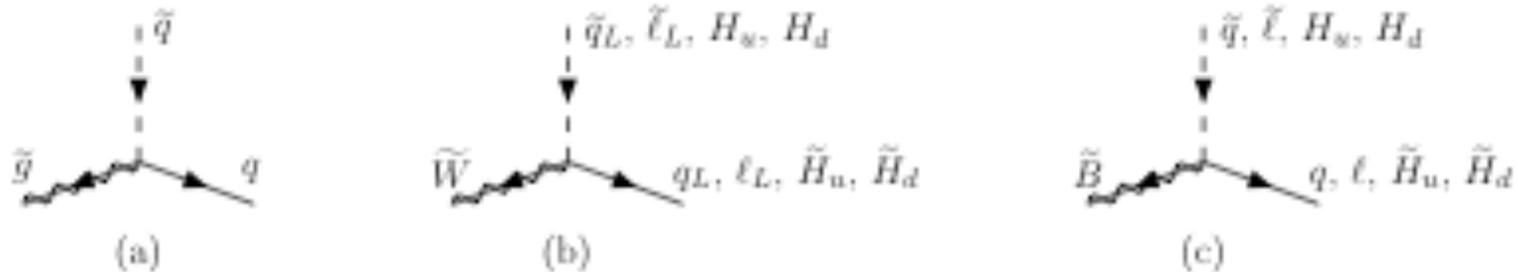


Figure 11: Couplings of the gluino, wino, and bino to MSSM (scalar, fermion) pairs.

In Figures 11b,c we show in a similar way the couplings of (squark, quark), (lepton, slepton) and (Higgs, higgsino) pairs to the winos and bino, with strengths proportional to the electroweak gauge couplings g and g' respectively.

For each of these diagrams, there is another with all arrows reversed.

Note that the winos only couple to the left-handed squarks and sleptons, and the (lepton, slepton) and (Higgs, higgsino) pairs of

course do not couple to the gluino.

The bino coupling to each (scalar, fermion) pair is also proportional to the weak hypercharge Y as given in Table 1.

The interactions shown in Figure 11 provide, for example, for decays $\tilde{q} \rightarrow q\tilde{g}$ and $\tilde{q} \rightarrow \tilde{W}q'$ and $\tilde{q} \rightarrow \tilde{B}q$ when the final states are kinematically allowed to be on-shell.

However, a complication is that the \tilde{W} and \tilde{B} states are not mass eigenstates, because of splitting and mixing due to electroweak symmetry breaking, as we will explore later.

There are also various scalar quartic interactions in the MSSM that are uniquely determined by gauge invariance and supersymmetry, according to the last term in Eq. (352), as illustrated in Figure 7i.

Among them are (Higgs)⁴ terms proportional to g^2 and g'^2 in the scalar potential. These are the direct generalization of the last term

in the Standard Model Higgs potential,

$$V = m_H^2 |H|^2 + \lambda |H|^4, \quad (370)$$

to the case of the MSSM. We will have occasion to identify them explicitly when we discuss the minimization of the MSSM Higgs potential.

The dimensionful couplings in the **supersymmetric** part of the MSSM Lagrangian are all dependent on μ . Using the general result of Eq. (303), μ provides for higgsino fermion mass terms

$$-\mathcal{L}_{\text{higgsino mass}} = \mu(\tilde{H}_u^+ \tilde{H}_d^- - \tilde{H}_u^0 \tilde{H}_d^0) + \text{c.c.}, \quad (371)$$

as well as Higgs squared-mass terms in the scalar potential

$$-\mathcal{L}_{\text{supersymmetric Higgs mass}} = |\mu|^2 (|H_u^0|^2 + |H_u^+|^2 + |H_d^0|^2 + |H_d^-|^2). \quad (372)$$

Since Eq. (372) is non-negative with a minimum at $H_u^0 = H_d^0 = 0$, we cannot understand electroweak symmetry breaking without including a negative supersymmetry-breaking squared-mass soft term for the Higgs scalars.

An explicit treatment of the Higgs scalar potential will therefore have to wait until we have introduced the soft terms for the MSSM.

However, we can already see a puzzle: we expect that μ should be roughly of order 10^2 or 10^3 GeV, in order to allow a Higgs VEV of order 174 GeV without too much miraculous cancellation between $|\mu|^2$ and the negative soft squared-mass terms that we have not written down yet. But why should $|\mu|^2$ be so small compared to, say, M_{P}^2 , and in particular why should it be roughly of the same order as m_{soft}^2 ?

The scalar potential of the MSSM seems to depend on two types of dimensionful parameters that are conceptually quite distinct, namely

the supersymmetry-respecting mass μ and the supersymmetry-breaking soft mass terms. Yet the observed value for the electroweak breaking scale suggests that without miraculous cancellations, both of these apparently unrelated mass scales should be within an order of magnitude or so of 100 GeV.

This puzzle is called “the μ problem”.

Several different solutions to the μ problem have been proposed, involving extensions of the MSSM of varying intricacy. They all work in roughly the same way; the μ term is required or assumed to be absent at tree-level before symmetry breaking, and then it arises from the VEV(s) of some new field(s). These VEVs are in turn determined by minimizing a potential that depends on soft supersymmetry-breaking terms.

In this way, the value of the effective parameter μ is no longer conceptually distinct from the mechanism of supersymmetry breaking;

if we can explain why $m_{\text{soft}} \ll M_{\text{P}}$, we will also be able to understand why μ is of the same order. Later, we will examine the NMSSM solution to the μ problem.

From the point of view of the MSSM, however, we can just treat μ as an independent parameter.

The μ -term and the Yukawa couplings in the superpotential Eq. (367) combine to yield (scalar)³ couplings [see the second and third terms on the right-hand side of Eq. (302)] of the form

$$\begin{aligned} \mathcal{L}_{\text{supersymmetric (scalar)}^3} = & \mu^* (\tilde{u} \mathbf{y}_u \tilde{u} H_d^{0*} + \tilde{d} \mathbf{y}_d \tilde{d} H_u^{0*} + \tilde{e} \mathbf{y}_e \tilde{e} H_u^{0*} \\ & + \tilde{u} \mathbf{y}_u \tilde{d} H_d^{-*} + \tilde{d} \mathbf{y}_d \tilde{u} H_u^{+*} + \tilde{e} \mathbf{y}_e \tilde{\nu} H_u^{+*}) + \text{c.c.} \end{aligned} \quad (373)$$

Figure 12 shows some of these couplings,

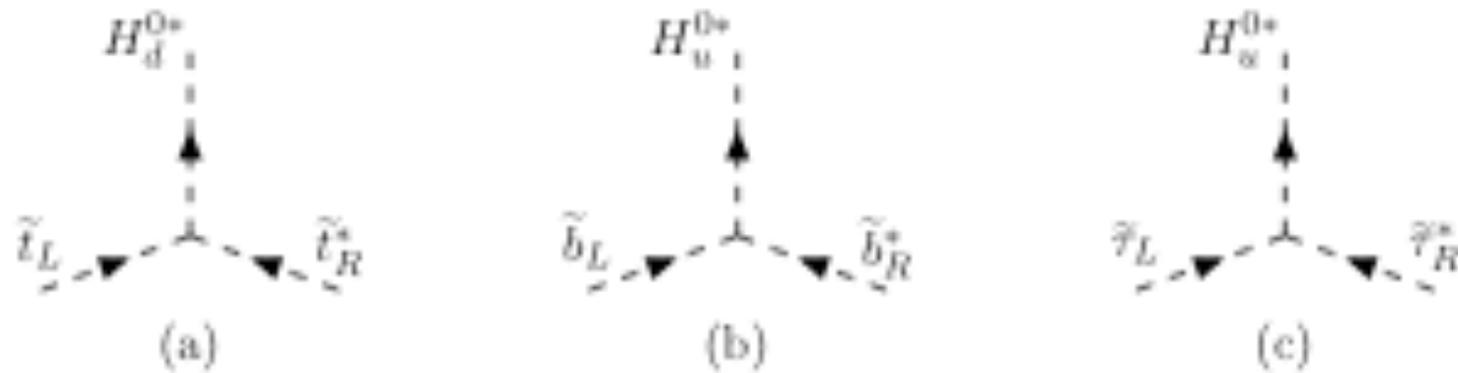


Figure 12: Some of the supersymmetric (scalar)³ couplings proportional to $\mu^* y_t$, $\mu^* y_b$, and $\mu^* y_\tau$. When H_u^0 and H_d^0 get VEVs, these contribute to (a) \tilde{t}_L, \tilde{t}_R mixing, (b) \tilde{b}_L, \tilde{b}_R mixing, and (c) $\tilde{\tau}_L, \tilde{\tau}_R$ mixing.

proportional to $\mu^* y_t$, $\mu^* y_b$, and $\mu^* y_\tau$ respectively. These play an important role in determining the mixing of top squarks, bottom squarks, and tau sleptons.

R-parity (also known as matter parity) and its consequences

The superpotential Eq. (367) is minimal in the sense that it is sufficient to produce a phenomenologically viable model.

However, there are other terms that one can write that are gauge-invariant and analytic in the chiral superfields, but are not included in the MSSM because they violate either baryon number (B) or total lepton number (L).

The most general gauge-invariant and renormalizable superpotential would include not only Eq. (367), but also the terms

$$W_{\Delta L=1} = \frac{1}{2} \lambda^{ijk} L_i L_j \bar{e}_k + \lambda'^{ijk} L_i Q_j \bar{d}_k + \mu'^i L_i H_u \quad (374)$$

$$W_{\Delta B=1} = \frac{1}{2} \lambda''^{ijk} \bar{u}_i \bar{d}_j \bar{d}_k \quad (375)$$

where family indices $i = 1, 2, 3$ have been restored. The chiral supermultiplets carry baryon number assignments $B = +1/3$ for Q_i ; $B = -1/3$ for \bar{u}_i, \bar{d}_i ; and $B = 0$ for all others. The total lepton number assignments are $L = +1$ for L_i , $L = -1$ for \bar{e}_i , and $L = 0$ for all others.

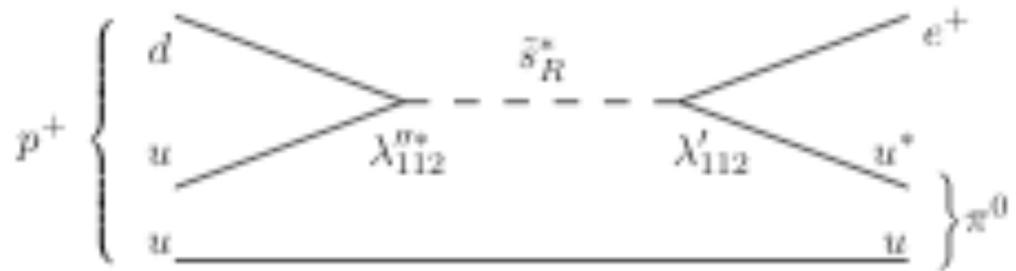
Therefore, the terms in Eq. (374) violate total lepton number by 1 unit (as well as the individual lepton flavors) and those in Eq. (375) violate baryon number by 1 unit.

The possible existence of such terms might seem rather disturbing, since corresponding B - and L -violating processes have not been seen experimentally.

The most obvious experimental constraint comes from the non-observation of proton decay, which would violate both B and L by 1 unit. If both λ' and λ'' couplings were present and unsuppressed, then the lifetime of the proton would be extremely short. For example, Feynman diagrams like the one in Figure 13⁵ would lead to

⁵In this diagram and others below, the arrows on propagators are often omitted for simplicity, and external fermion label refer to physical particle states rather than 2-component fermion fields.

Figure 13: Squarks would mediate disastrously rapid proton decay if R -parity were violated by both $\Delta B = 1$ and $\Delta L = 1$ interactions. This example shows $p \rightarrow e^+ \pi^0$ mediated by a strange (or bottom) squark.



$p^+ \rightarrow e^+ \pi^0$ (shown) or $e^+ K^0$ or $\mu^+ \pi^0$ or $\mu^+ K^0$ or $\nu \pi^+$ or νK^+ etc. depending on which components of λ' and λ'' are largest.⁶ As a rough estimate based on dimensional analysis, for example,

$$\Gamma_{p \rightarrow e^+ \pi^0} \sim m_{\text{proton}}^5 \sum_{i=2,3} |\lambda'^{11i} \lambda''^{11i}|^2 / m_{\tilde{d}_i}^4, \quad (376)$$

which would be a tiny fraction of a second if the couplings were of order unity and the squarks have masses of order 1 TeV. In contrast,

⁶The coupling λ'' must be antisymmetric in its last two flavor indices, since the color indices are combined antisymmetrically. That is why the squark in Figure 13 can be \tilde{s} or \tilde{b} , but not \tilde{d} , for u, d quarks in the proton.

the decay time of the proton into lepton+meson final states is known experimentally to be in excess of 10^{32} years. Therefore, at least one of λ'^{ijk} or λ''^{11k} for each of $i = 1, 2$; $j = 1, 2$; $k = 2, 3$ must be extremely small. Many other processes also give strong constraints on the violation of lepton and baryon numbers.

One could simply try to take **B** and **L** conservation as a postulate in the MSSM. However, this is clearly a step backward from the situation in the Standard Model, where the conservation of these quantum numbers is *not* assumed, but is rather a pleasantly “accidental” consequence of the fact that there are no possible renormalizable Lagrangian terms that violate **B** or **L**.

Furthermore, there is a quite general obstacle to treating **B** and **L** as fundamental symmetries of Nature, since they are known to be necessarily violated by non-perturbative electroweak effects (even though those effects are calculably negligible for experiments at ordinary energies).

Therefore, in the MSSM one adds a new symmetry, which has the effect of eliminating the possibility of B and L violating terms in the **renormalizable** superpotential, while allowing the good terms in Eq. (367). This new symmetry is called “ R -parity” or equivalently “matter parity”.

Matter parity is a multiplicatively conserved quantum number defined as

$$P_M = (-1)^{3(B-L)} \quad (377)$$

for each particle in the theory.

It is easy to check that the quark and lepton supermultiplets all have $P_M = -1$, while the Higgs supermultiplets H_u and H_d have $P_M = +1$.

The gauge bosons and gauginos of course do not carry baryon number or lepton number, so they are assigned matter parity $P_M = +1$.

The symmetry principle to be enforced is that a candidate term in the

Lagrangian (or in the superpotential) is allowed only if the product of P_M for all of the fields in it is $+1$.

It is easy to see that each of the terms in Eqs. (374) and (375) is thus forbidden, while the good and necessary terms in Eq. (367) are allowed.

This discrete symmetry commutes with supersymmetry, as all members of a given supermultiplet have the same matter parity.

The advantage of matter parity is that it can in principle be an *exact* and fundamental symmetry, which B and L themselves cannot, since they are known to be violated by non-perturbative electroweak effects.

Even with exact matter parity conservation in the MSSM, one expects that baryon number and total lepton number violation can occur in tiny amounts, due to **non-renormalizable** terms in the Lagrangian. However $B - L$ remains exactly conserved.

To repeat, the MSSM does not have **renormalizable** (*i.e.* dimensionless coupling constant) interactions that violate **B** or **L** after adopting the standard assumption of matter parity conservation for the superpotential (which is where the dimensionless couplings reside, the soft-supersymmetry-breaking interactions all having by definition dimensionful couplings).

It is often useful to recast matter parity in terms of R -parity, defined for each particle as

$$P_R = (-1)^{3(B-L)+2s} \quad (378)$$

where s is the spin of the particle.

Now, matter parity conservation and R -parity conservation are precisely equivalent, since the product of $(-1)^{2s}$ for the particles involved in any interaction vertex in a theory that conserves angular momentum is always equal to $+1$.

However, particles within the same supermultiplet do not have the

same R -parity.

In general, symmetries with the property that fields within the same supermultiplet have different transformations are called R symmetries; they do not commute with supersymmetry.

Continuous $U(1)$ R symmetries are often encountered in the model-building literature; they should not be confused with R -parity, which is a discrete Z_2 symmetry.

In fact, the matter parity version of R -parity makes clear that there is really nothing intrinsically “ R ” about it; in other words it secretly does commute with supersymmetry, so its name is somewhat suboptimal.

Nevertheless, the R -parity assignment is very useful for phenomenology because all of the Standard Model particles and the Higgs bosons have even R -parity ($P_R = +1$), while all of the squarks, sleptons, gauginos, and higgsinos have odd R -parity ($P_R = -1$).

The R -parity odd particles are known as “supersymmetric particles”

or “sparticles” for short, and they are distinguished by a tilde (see Tables 1 and 2).

If R -parity is exactly conserved, then there can be no mixing between the sparticles and the $P_R = +1$ particles. Furthermore, every interaction vertex in the theory contains an even number of $P_R = -1$ sparticles.

This has three extremely important phenomenological consequences:

- The lightest sparticle with $P_R = -1$, *i.e.* the “lightest supersymmetric particle” or LSP, must be absolutely stable. If the LSP is electrically neutral, it interacts only weakly with ordinary matter, and so can make an attractive candidate for the non-baryonic dark matter that seems to be required by cosmology.
- Each sparticle other than the LSP must eventually decay into a state that contains an odd number of LSPs (usually just one).
- In collider experiments, sparticles can only be produced in even

numbers (usually two-at-a-time).

We *define* the MSSM to conserve R -parity or equivalently matter parity.

While this decision seems to be well-motivated phenomenologically by proton decay constraints and the hope that the LSP will provide a good dark matter candidate, it might appear somewhat artificial from a theoretical point of view. After all, the MSSM would not suffer any internal inconsistency if we did not impose matter parity conservation.

Furthermore, it is fair to ask why matter parity should be exactly conserved, given that the discrete symmetries in the Standard Model (ordinary parity P , charge conjugation C , time reversal T , etc.) are all known to be inexact symmetries.

Fortunately, it *is* sensible to formulate matter parity as a discrete symmetry that is exactly conserved.

In general, exactly conserved, or “gauged” discrete symmetries can exist provided that they satisfy certain anomaly cancellation conditions (much like continuous gauged symmetries).

One particularly attractive way this could occur is if $B - L$ is a continuous gauge symmetry that is spontaneously broken at some very high energy scale. A continuous $U(1)_{B-L}$ forbids the renormalizable terms that violate B and L , but this gauge symmetry must be spontaneously broken, since there is no corresponding massless vector boson.

However, if gauged $U(1)_{B-L}$ is only broken by scalar VEVs (or other order parameters) that carry even integer values of $3(B - L)$, then P_M will automatically survive as an exactly conserved discrete remnant subgroup. A variety of extensions of the MSSM in which exact R -parity conservation is guaranteed in just this way have been proposed.

It may also be possible to have gauged discrete symmetries that do not owe their exact conservation to an underlying continuous gauged symmetry, but rather to some other structure such as can occur in string theory. It is also possible that R -parity is broken, or is replaced by some alternative discrete symmetry.

Soft supersymmetry breaking in the MSSM

To complete the description of the MSSM, we need to specify the soft supersymmetry breaking terms. Earlier, we learned how to write down the most general set of such terms in any supersymmetric theory. Applying this recipe to the MSSM, we have:

$$\begin{aligned}
 \mathcal{L}_{\text{soft}}^{\text{MSSM}} = & -\frac{1}{2} \left(M_3 \tilde{g}\tilde{g} + M_2 \tilde{W}\tilde{W} + M_1 \tilde{B}\tilde{B} + \text{c.c.} \right) \\
 & - \left(\tilde{u} \mathbf{a}_u \tilde{Q} H_u - \tilde{d} \mathbf{a}_d \tilde{Q} H_d - \tilde{e} \mathbf{a}_e \tilde{L} H_d + \text{c.c.} \right) \\
 & - \tilde{Q}^\dagger \mathbf{m}_Q^2 \tilde{Q} - \tilde{L}^\dagger \mathbf{m}_L^2 \tilde{L} - \tilde{u} \mathbf{m}_u^2 \tilde{u}^\dagger - \tilde{d} \mathbf{m}_d^2 \tilde{d}^\dagger - \tilde{e} \mathbf{m}_e^2 \tilde{e}^\dagger \\
 & - m_{H_u}^2 H_u^* H_u - m_{H_d}^2 H_d^* H_d - (b H_u H_d + \text{c.c.}) . \tag{379}
 \end{aligned}$$

You should keep in mind the notation that, for example, $H_u^* H_u$ really stands for an isospin singlet contraction $\epsilon_{ij} H_u^i H_u^j$, and similarly for other such constructs.

1. In the first line of Eq. (379), M_3 , M_2 , and M_1 are the gluino, wino, and bino mass terms.⁷
2. The second line in Eq. (379) contains the (scalar)³ couplings [of the type a^{ijk} in Eq. (364)]. Each of \mathbf{a}_u , \mathbf{a}_d , \mathbf{a}_e is a complex 3×3 matrix in family space, with dimensions of [mass]. They are in one-to-one correspondence with the Yukawa couplings of the superpotential.
3. The third line of Eq. (379) consists of squark and slepton mass terms of the $(m^2)_i^j$ type in Eq. (364). Each of \mathbf{m}_Q^2 , \mathbf{m}_u^2 , \mathbf{m}_d^2 , \mathbf{m}_L^2 , \mathbf{m}_e^2 is a 3×3 matrix in family space that can have complex entries, but they must be hermitian so that the Lagrangian is real. (To

⁷Here, and from now on, we suppress the adjoint representation gauge indices on the wino and gluino fields, and the gauge indices on all of the chiral supermultiplet fields.

avoid clutter, we do not put tildes on the Q in \mathbf{m}_Q^2 , etc.)

4. Finally, in the last line of Eq. (379) we have supersymmetry-breaking contributions to the Higgs potential; $m_{H_u}^2$ and $m_{H_d}^2$ are squared-mass terms of the $(m^2)_i^j$ type, while b is the only squared-mass term of the type b^{ij} in Eq. (364) that can occur in the MSSM.⁸

As argued in earlier, we expect

$$M_1, M_2, M_3, \mathbf{a}_u, \mathbf{a}_d, \mathbf{a}_e \sim m_{\text{soft}}, \quad (380)$$

$$\mathbf{m}_Q^2, \mathbf{m}_L^2, \mathbf{m}_u^2, \mathbf{m}_d^2, \mathbf{m}_e^2, m_{H_u}^2, m_{H_d}^2, b \sim m_{\text{soft}}^2, \quad (381)$$

with characteristic mass scale m_{soft} that is not much larger than 1000 GeV.

The expression Eq. (379) is the most general soft supersymmetry-breaking Lagrangian of the form Eq. (364) that is compatible with

⁸The parameter called b here is often seen elsewhere as $B\mu$ or m_{12}^2 or m_3^2 .

gauge invariance and matter parity conservation in the MSSM.

Unlike the supersymmetry-preserving part of the Lagrangian, the above $\mathcal{L}_{\text{soft}}^{\text{MSSM}}$ introduces many new parameters that were not present in the ordinary Standard Model.

A careful count reveals that there are 105 masses, phases and mixing angles in the MSSM Lagrangian that cannot be rotated away by redefining the phases and flavor basis for the quark and lepton supermultiplets, and that have no counterpart in the ordinary Standard Model.

Thus, in principle, supersymmetry breaking (not supersymmetry itself) appears to introduce a tremendous arbitrariness in the Lagrangian.

To avoid phenomenological problems these soft terms must be properly organized. There are many organizing principles that we may return to later — the best such principles are realized in the context of specific models of soft-supersymmetry-breaking.

For example, there are dangerous flavor-changing and CP-violating effects in the MSSM. These can be evaded if one assumes (or can explain!) that supersymmetry breaking is suitably “universal”.

Consider an idealized limit in which the squark and slepton squared-mass matrices are flavor-blind, each proportional to the 3×3 identity matrix in family space:

$$\mathbf{m}_Q^2 = m_Q^2 \mathbf{1}, \quad \mathbf{m}_u^2 = m_u^2 \mathbf{1}, \quad \mathbf{m}_d^2 = m_d^2 \mathbf{1}, \quad \mathbf{m}_L^2 = m_L^2 \mathbf{1}, \quad \mathbf{m}_e^2 = m_e^2 \mathbf{1}. \quad (382)$$

Then all squark and slepton mixing angles are rendered trivial, because squarks and sleptons with the same electroweak quantum numbers will be degenerate in mass and can be rotated into each other at will.

Supersymmetric contributions to flavor-changing neutral current processes will therefore be very small in such an idealized limit, up to mixing induced by \mathbf{a}_u , \mathbf{a}_d , \mathbf{a}_e .

Making the further assumption that the (scalar)³ couplings are each

proportional to the corresponding Yukawa coupling matrix,

$$\mathbf{a}_u = A_{u0} \mathbf{y}_u, \quad \mathbf{a}_d = A_{d0} \mathbf{y}_d, \quad \mathbf{a}_e = A_{e0} \mathbf{y}_e, \quad (383)$$

will ensure that only the squarks and sleptons of the third family can have large (scalar)³ couplings.

Finally, one can avoid disastrously large CP-violating effects by assuming that the soft parameters do not introduce new complex phases.

This is automatic for $m_{H_u}^2$ and $m_{H_d}^2$, and for m_Q^2 , $m_{\bar{u}}^2$, etc. if Eq. (382) is assumed; if they were not real numbers, the Lagrangian would not be real.

One can also fix μ in the superpotential and b in Eq. (379) to be real, by appropriate phase rotations of fermion and scalar components of the H_u and H_d supermultiplets.

If one then assumes that

$$\arg(M_1), \arg(M_2), \arg(M_3), \arg(A_{u0}), \arg(A_{d0}), \arg(A_{e0}) = 0 \text{ or } \pi, \quad (384)$$

then the only CP-violating phase in the theory will be the usual CKM phase found in the ordinary Yukawa couplings.

Together, the conditions Eqs. (382)-(384) make up a rather weak version of what is often called the hypothesis of *soft supersymmetry-breaking universality*.

The MSSM with these flavor- and CP-preserving relations imposed has far fewer parameters than the most general case.

Besides the usual Standard Model gauge and Yukawa coupling parameters, there are 3 independent real gaugino masses, only 5 real squark and slepton squared mass parameters, 3 real scalar cubic coupling parameters, and 4 Higgs mass parameters (one of which can be traded for the known electroweak breaking scale).

Typically, universality might hold in a model at some high scale. Renormalization group evolution will then create deviations at a low scale, but these will be loop suppressed and logarithmically dependent upon the ratio of the high scale to the TeV scale.

Gauge Coupling Unification

The idea of a simple model at some high scale gains considerable support from “gauge coupling unification”.

We will not go into detail here. A treatment of this topic is typically presented in 230C. I summarize the results.

The 1-loop RG equations for the Standard Model gauge couplings g_1, g_2, g_3 are

$$\beta_{g_a} \equiv \frac{d}{dt} g_a = \frac{1}{16\pi^2} b_a g_a^3, \quad (b_1, b_2, b_3) = \begin{cases} (41/10, -19/6, -7) & \text{Standard Model} \\ (33/5, 1, -3) & \text{MSSM} \end{cases} \quad (385)$$

where $t = \ln(Q/Q_0)$, with Q the RG scale.

The MSSM coefficients are larger because of the extra MSSM particles in loops.

The normalization for g_1 here is chosen to agree with the canonical covariant derivative for grand unification of the gauge group $SU(3)_C \times SU(2)_L \times U(1)_Y$ into $SU(5)$ or $SO(10)$.

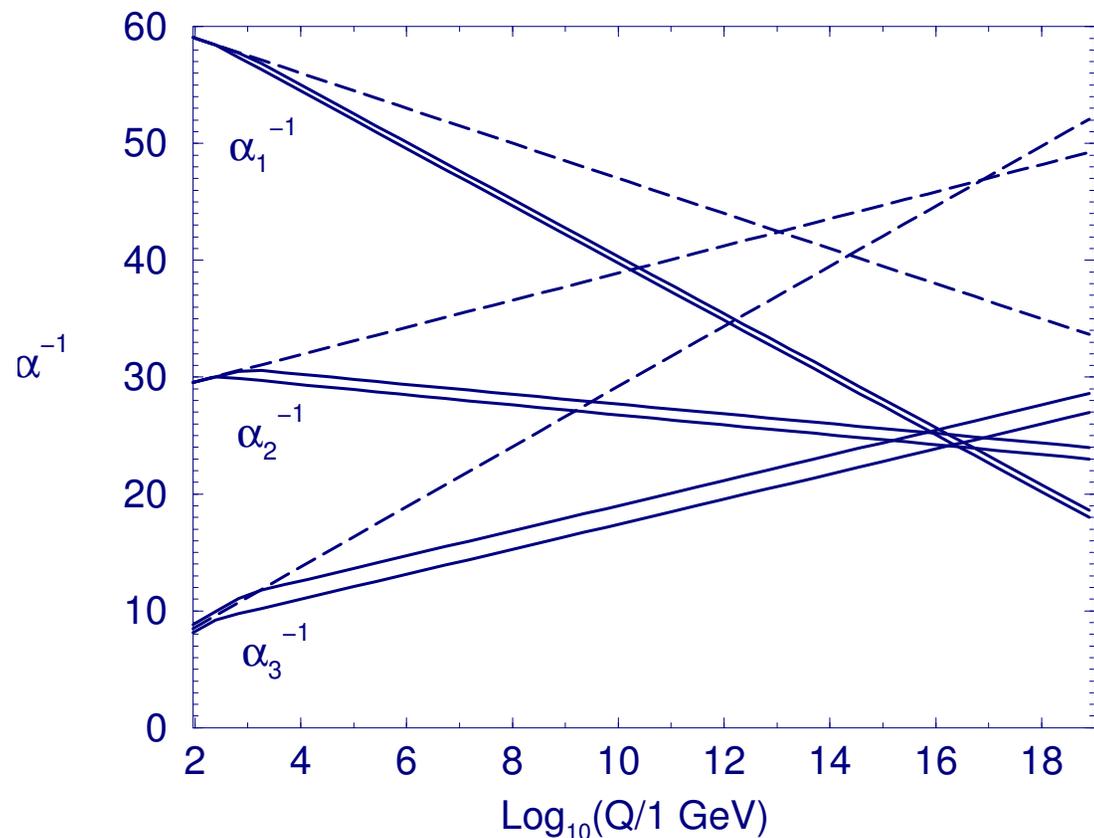
Thus in terms of the conventional electroweak gauge couplings g and g' with $e = g \sin \theta_W = g' \cos \theta_W$, one has $g_2 = g$ and $g_1 = \sqrt{5/3}g'$.

The quantities $\alpha_a = g_a^2/4\pi$ have the nice property that their reciprocals run linearly with RG scale at one-loop order:

$$\frac{d}{dt}\alpha_a^{-1} = -\frac{b_a}{2\pi} \quad (a = 1, 2, 3) \quad (386)$$

Figure 14 compares the RG evolution of the α_a^{-1} , including two-loop effects, in the Standard Model (dashed lines) and the MSSM (solid lines).

Figure 14: RG evolution of the inverse gauge couplings $\alpha_a^{-1}(Q)$ in the Standard Model (dashed lines) and the MSSM (solid lines). In the MSSM case, the sparticle mass thresholds are varied between 250 GeV and 1 TeV, and $\alpha_3(m_Z)$ between 0.113 and 0.123. Two-loop effects are included.



Unlike the Standard Model, the MSSM includes just the right particle

content to ensure that the gauge couplings can unify, at a scale $M_U \sim 2 \times 10^{16}$ GeV.

While the apparent unification of gauge couplings at M_U might be just an accident, it may also be taken as a strong hint in favor of a grand unified theory (GUT) or superstring models, both of which can naturally accommodate gauge coupling unification below M_P .

Furthermore, if this hint is taken seriously, then we can reasonably expect to be able to apply a similar RG analysis to the other MSSM couplings and soft masses as well.

The unification shown in the plot assumes that the switchover from SM to MSSM occurs at or near the TeV scale. This is another very important motivation for thinking sparticles have masses of order a TeV.

Gaugino Mass Unification

The one-loop RG equations for the three gaugino mass parameters in the MSSM are determined by the same quantities b_a^{MSSM} that appear in the gauge coupling RG Eqs. (385):

$$\beta_{M_a} \equiv \frac{d}{dt} M_a = \frac{1}{8\pi^2} b_a g_a^2 M_a \quad (b_a = 33/5, 1, -3) \quad (387)$$

for $a = 1, 2, 3$.

It follows that the three ratios M_a/g_a^2 are each constant (RG scale independent) up to small two-loop corrections.

Since the gauge couplings are observed to unify at $Q = M_U = 2 \times 10^{16}$ GeV, it is a popular assumption that the gaugino masses also unify⁹

⁹In GUT models, it is automatic that the gauge couplings and gaugino masses are unified at all scales $Q \geq M_U$, because in the unified theory the gauginos all live in the same representation of the unified gauge group. In many superstring models, this can also be a good approximation.

near that scale, with a value called $m_{1/2}$. If so, then it follows that

$$\frac{M_1}{g_1^2} = \frac{M_2}{g_2^2} = \frac{M_3}{g_3^2} = \frac{m_{1/2}}{g_U^2} \quad (388)$$

at any RG scale, up to small (and known) two-loop effects and possibly much larger (and not so known) threshold effects near M_U .

Here g_U is the unified gauge coupling at $Q = M_U$. The hypothesis of Eq. (388) is particularly powerful because the gaugino mass parameters feed strongly into the RG equations for all of the other soft terms,

General Picture of MSSM soft-SUSY-breaking

For various reasons, we expect that the MSSM soft terms arise indirectly or radiatively, rather than from tree-level renormalizable couplings to the supersymmetry-breaking order parameters. Supersymmetry breaking evidently occurs in a “hidden sector” of particles that have

no (or only very small) direct couplings to the “visible sector” chiral supermultiplets of the MSSM.

However, the two sectors do share some interactions that are responsible for mediating supersymmetry breaking from the hidden sector to the visible sector, resulting in the MSSM soft terms. (See Figure 15.)

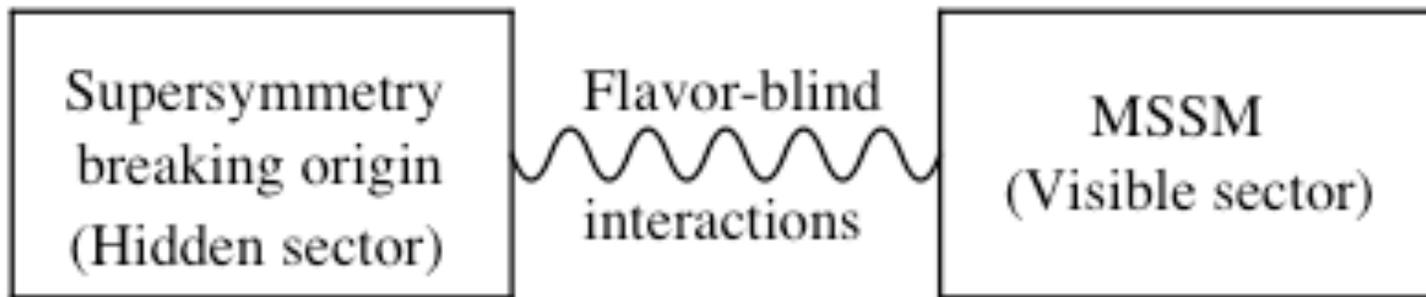


Figure 15: The presumed schematic structure for supersymmetry breaking.

An important feature of such a picture is that if the mediating interactions between the SUSY-breaking sector and the MSSM

are flavor-blind, then the soft terms appearing in the MSSM will automatically obey conditions like Eqs. (382), (383) and (384).

To give you some idea of possible model implications for SUSY-breaking, I summarize what the high-scale boundary conditions at M_U would look like in a couple of scenarios.

Planck-scale-mediated supersymmetry breaking

There possibility is to assume that communication of the MSSM sector with a hidden sector takes place at scales of order M_P .

For the parameters appearing in Eq. (379) one finds:

$$M_3 = M_2 = M_1 = m_{1/2}, \quad (389)$$

$$\mathbf{m}_Q^2 = \mathbf{m}_u^2 = \mathbf{m}_d^2 = \mathbf{m}_L^2 = \mathbf{m}_e^2 = m_0^2 \mathbf{1}, \quad m_{H_u}^2 = m_{H_d}^2 = m_0^2, \quad (390)$$

$$\mathbf{a}_u = A_0 \mathbf{y}_u, \quad \mathbf{a}_d = A_0 \mathbf{y}_d, \quad \mathbf{a}_e = A_0 \mathbf{y}_e, \quad (391)$$

$$b = B_0 \mu, \quad (392)$$

at a renormalization scale $Q \approx M_P$.

It is a matter of some controversy whether the assumptions going into this parameterization are well-motivated on purely theoretical grounds,¹⁰ but from a phenomenological perspective they are clearly very nice.

This framework successfully evades the most dangerous types of flavor changing and CP violation. In particular, Eqs. (390) and (391) are just stronger versions of Eqs. (382) and (383), respectively. If $m_{1/2}$, A_0 and B_0 all have the same complex phase, then Eq. (384) will also be satisfied.

Equations (389)-(392) also have the virtue of being highly predictive. [Of course, Eq. (392) is content-free unless one can relate B_0 to the other parameters in some non-trivial way.] As discussed earlier, they should be applied as RG boundary conditions at the scale M_P .

The RG evolution of the soft parameters down to the electroweak

¹⁰The familiar flavor blindness of gravity expressed in Einstein's equivalence principle need not imply Eqs. (389)-(391).

scale will then allow us to predict the entire MSSM spectrum in terms of just five parameters $m_{1/2}$, m_0^2 , A_0 , B_0 , and μ (plus the already-measured gauge and Yukawa couplings of the MSSM).

A popular approximation is to start this RG running from the unification scale $M_U \approx 2 \times 10^{16}$ GeV instead of M_P . The reason for this is more practical than principled; the apparent unification of gauge couplings gives us a strong hint that we know something about how the RG equations behave up to M_U , but unfortunately gives us little guidance about what to expect at scales between M_U and M_P .

The errors made in neglecting these effects are proportional to a loop suppression factor times $\ln(M_P/M_U)$. These corrections hopefully can be partly absorbed into a redefinition of m_0^2 , $m_{1/2}$, A_0 and B_0 at M_U , but in many cases can lead to other important effects.

The framework described in the above few paragraphs has been the subject of the bulk of phenomenological studies of supersymmetry. It

is sometimes referred to as the *minimal supergravity* (MSUGRA) or *supergravity-inspired* scenario for the soft terms.

Particular models of gravity-mediated supersymmetry breaking can be even more predictive, relating some of the parameters $m_{1/2}$, m_0^2 , A_0 and B_0 to each other and to the mass of the gravitino $m_{3/2}$. For example, three popular kinds of models for the soft terms are:

- Dilaton-dominated: $m_0^2 = m_{3/2}^2$, $m_{1/2} = -A_0 = \sqrt{3}m_{3/2}$.
- Polonyi: $m_0^2 = m_{3/2}^2$, $A_0 = (3 - \sqrt{3})m_{3/2}$, $m_{1/2} = \mathcal{O}(m_{3/2})$.
- “No-scale”: $m_{1/2} \gg m_0, A_0, m_{3/2}$.
- Gauge-Mediated Supersymmetry Breaking Models

In gauge-mediated supersymmetry breaking (GMSB) models, the

ordinary gauge interactions, rather than gravity, are responsible for the appearance of soft supersymmetry breaking in the MSSM.

The basic idea is to introduce some new chiral supermultiplets, called messengers, that couple to the ultimate source of supersymmetry breaking, and also couple indirectly to the (s)quarks and (s)leptons and Higgs(inos) of the MSSM through the ordinary $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge boson and gaugino interactions.

There is still gravitational communication between the MSSM and the source of supersymmetry breaking, of course, but that effect is now relatively unimportant compared to the gauge interaction effects.

In contrast to Planck-scale mediation, GMSB can be understood entirely in terms of loop effects in a renormalizable framework. In the simplest such model, the messenger fields are a set of left-handed chiral supermultiplets $q, \bar{q}, \ell, \bar{\ell}$ transforming under $SU(3)_C \times$

$SU(2)_L \times U(1)_Y$ as

$$q \sim (\mathbf{3}, \mathbf{1}, -\frac{1}{3}), \quad \bar{q} \sim (\bar{\mathbf{3}}, \mathbf{1}, \frac{1}{3}), \quad \ell \sim (\mathbf{1}, \mathbf{2}, \frac{1}{2}), \quad \bar{\ell} \sim (\mathbf{1}, \mathbf{2}, -\frac{1}{2}). \quad (393)$$

These supermultiplets contain messenger quarks $\psi_q, \psi_{\bar{q}}$ and scalar quarks q, \bar{q} and messenger leptons $\psi_\ell, \psi_{\bar{\ell}}$ and scalar leptons $\ell, \bar{\ell}$. All of these particles must get very large masses so as not to have been discovered already.

Assume they do so by coupling to a gauge-singlet chiral supermultiplet S through a superpotential:

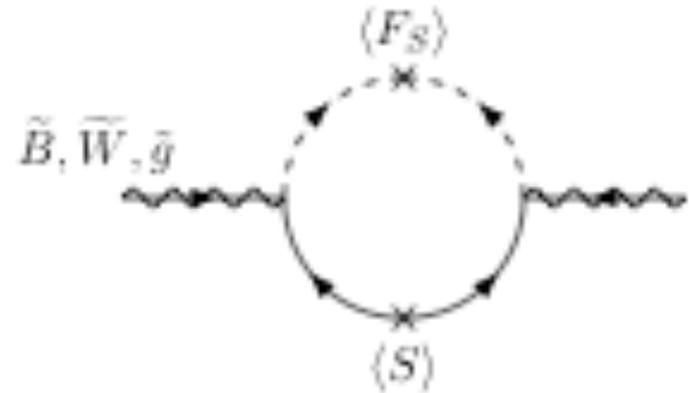
$$W_{\text{mess}} = y_2 S \ell \bar{\ell} + y_3 S q \bar{q}. \quad (394)$$

The scalar component of S and its auxiliary (F -term) component are each supposed to acquire VEVs, denoted $\langle S \rangle$ and $\langle F_S \rangle$ respectively.

The supersymmetry violation apparent in this messenger spectrum for

$\langle F_S \rangle \neq 0$ is communicated to the MSSM sparticles through radiative corrections. The MSSM gauginos obtain masses from the 1-loop Feynman diagram shown in Figure 16.

Figure 16: Contributions to the MSSM gaugino masses in gauge-mediated supersymmetry breaking models come from one-loop graphs involving virtual messenger particles.



The scalar and fermion lines in the loop are messenger fields.

The interaction vertices in Figure 16 are of gauge coupling strength even though they do not involve gauge bosons; compare Figure 7g. In this way, gauge-mediation provides that q, \bar{q} messenger loops give masses to the gluino and the bino, and $\ell, \bar{\ell}$ messenger loops give masses to the wino and bino fields.

Computing the 1-loop diagrams, one finds that the resulting MSSM

gaugino masses are given by

$$M_a = \frac{\alpha_a}{4\pi} \Lambda, \quad (a = 1, 2, 3), \quad (395)$$

in the conventional normalization for α_a , where we have introduced a mass parameter

$$\Lambda \equiv \langle F_S \rangle / \langle S \rangle. \quad (396)$$

(Note that if $\langle F_S \rangle$ were 0, then $\Lambda = 0$ and the messenger scalars would be degenerate with their fermionic superpartners and there would be no contribution to the MSSM gaugino masses.)

In contrast, the corresponding MSSM gauge bosons cannot get a corresponding mass shift, since they are protected by gauge invariance.

So supersymmetry breaking has been successfully communicated to the MSSM (“visible sector”).

To a good approximation, Eq. (395) holds for the running gaugino masses at an RG scale Q_0 corresponding to the average characteristic mass of the heavy messenger particles, roughly of order $M_{\text{mess}} \sim y_I \langle S \rangle$ for $I = 2, 3$. The running mass parameters can then be RG-evolved down to the electroweak scale to predict the physical masses to be measured by future experiments.

The scalars of the MSSM do not get any radiative corrections to their masses at one-loop order. The leading contribution to their masses comes from the two-loop graphs shown in Figure 17, with the messenger fermions (heavy solid lines) and messenger scalars (heavy dashed lines) and ordinary gauge bosons and gauginos running around the loops.

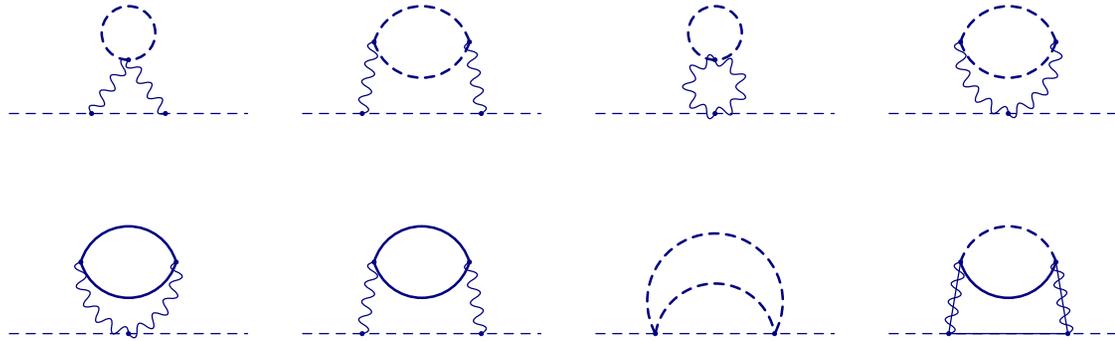


Figure 17: MSSM scalar squared masses in gauge-mediated supersymmetry breaking models arise in leading order from these two-loop Feynman graphs. The heavy dashed lines are messenger scalars, the solid lines are messenger fermions, the wavy lines are ordinary Standard Model gauge bosons, and the solid lines with wavy lines superimposed are the MSSM gauginos.

By computing these graphs, one finds that each MSSM scalar ϕ_i gets a squared mass given by:

$$m_{\phi_i}^2 = 2\Lambda^2 \left[\left(\frac{\alpha_3}{4\pi}\right)^2 C_3(i) + \left(\frac{\alpha_2}{4\pi}\right)^2 C_2(i) + \left(\frac{\alpha_1}{4\pi}\right)^2 C_1(i) \right], \quad (397)$$

where $C_a(i)$ are the quadratic Casimir group theory invariants for the

superfield, defined in terms of the Lie algebra generators T^a by

$$(T^a T^a)_i^j = C_a(i) \delta_i^j \quad (398)$$

with gauge couplings g_a . Explicitly, for the MSSM supermultiplets:

$$C_3(i) = \begin{cases} 4/3 & \text{for } \Phi_i = Q, \bar{u}, \bar{d}, \\ 0 & \text{for } \Phi_i = L, \bar{e}, H_u, H_d, \end{cases} \quad (399)$$

$$C_2(i) = \begin{cases} 3/4 & \text{for } \Phi_i = Q, L, H_u, H_d, \\ 0 & \text{for } \Phi_i = \bar{u}, \bar{d}, \bar{e}, \end{cases} \quad (400)$$

$$C_1(i) = 3Y_i^2/5 \quad \text{for each } \Phi_i \text{ with weak hypercharge } Y_i. \quad (401)$$

The squared masses in Eq. (397) are positive (fortunately!).

The terms \mathbf{a}_u , \mathbf{a}_d , \mathbf{a}_e arise first at two-loop order, and are suppressed by an extra factor of $\alpha_a/4\pi$ compared to the gaugino masses. So, to

a very good approximation one has, at the messenger scale,

$$\mathbf{a}_u = \mathbf{a}_d = \mathbf{a}_e = 0, \quad (402)$$

a significantly stronger condition than Eq. (383).

Again, Eqs. (397) and (402) should be applied at an RG scale equal to the average mass of the messenger fields running in the loops.

However, evolving the RG equations down to the electroweak scale generates non-zero \mathbf{a}_u , \mathbf{a}_d , and \mathbf{a}_e proportional to the corresponding Yukawa matrices and the non-zero gaugino masses. These will only be large for the third-family squarks and sleptons, in the approximation of Eq. (368).

The parameter b may also be taken to vanish near the messenger scale, but this is quite model-dependent, and in any case b will be non-zero when it is RG-evolved to the electroweak scale.

In practice, b can be fixed in terms of the other parameters by the requirement of correct electroweak symmetry breaking, as discussed when we consider the Higgs sector.

Messengers with masses far below the GUT scale will affect the running of gauge couplings and might therefore be expected to ruin the apparent unification shown in Figure 14.

However, if the messengers come in complete multiplets of the $SU(5)$ global symmetry¹¹ that contains the Standard Model gauge group, and are not very different in mass, then approximate unification of gauge couplings will still occur when they are extrapolated up to the same scale M_U (but with a larger unified value for the gauge couplings at that scale).

For this reason, a popular class of models is one in which gauge

¹¹This $SU(5)$ may or may not be promoted to a local gauge symmetry at the GUT scale. For our present purposes, it is used only as a classification scheme, since the global $SU(5)$ symmetry is only approximate in the effective theory at the (much lower) messenger mass scale where gauge mediation takes place.

coupling unification is easily implemented and is obtained by taking the messengers to consist of N_5 copies of the $\mathbf{5} + \bar{\mathbf{5}}$ of $SU(5)$, resulting in

$$M_a = \frac{\alpha_a}{4\pi} \Lambda N_5, \quad (403)$$

$$m_{\phi_i}^2 = 2\Lambda^2 N_5 \sum_{a=1}^3 C_a(i) \left(\frac{\alpha_a}{4\pi} \right)^2, \quad (404)$$

since now there are N_5 copies of the minimal messenger sector particles running around the loops.

For example, the minimal model in Eq. (393) corresponds to $N_5 = 1$.

A single copy of $\mathbf{10} + \bar{\mathbf{10}}$ of $SU(5)$ has Dynkin indices $\sum_I n_a(I) = 3$, and so can be substituted for 3 copies of $\mathbf{5} + \bar{\mathbf{5}}$.

(Other combinations of messenger multiplets can also preserve the apparent unification of gauge couplings.)

Note that the gaugino masses scale like N_5 , while the scalar masses scale like $\sqrt{N_5}$. This means that sleptons and squarks will tend to be lighter relative to the gauginos for larger values of N_5 in non-minimal models.

However, if N_5 is too large, then the running gauge couplings will diverge before they can unify at M_U . For messenger masses of order 10^6 GeV or less, for example, one needs $N_5 \leq 4$.

Extra-dimensional and anomaly-mediated supersymmetry breaking

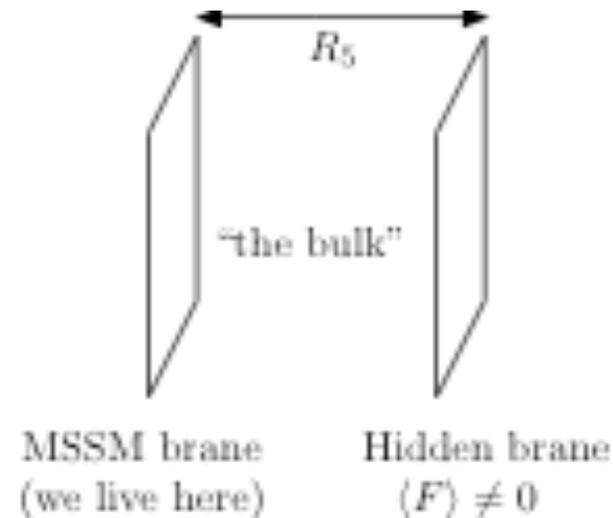
It is also possible to take the partitioning of the MSSM and supersymmetry breaking sectors shown in Fig. 15 seriously as geography. This can be accomplished by assuming that there are extra spatial dimensions of the Kaluza-Klein or warped type, so that a physical distance separates the visible and hidden¹² sectors. This general idea opens up numerous possibilities, which are hard to classify in

¹²The name “sequestered” is often used instead of “hidden” in this context.

a detailed way. For example, string theory suggests six such extra dimensions, with a staggeringly huge number of possible solutions.

Many of the more recently popular models used to explore this extra-dimensional mediated supersymmetry breaking (the acronym XMSB is tempting) use just one single hidden extra dimension with the MSSM chiral supermultiplets confined to one 4-dimensional spacetime brane and the supersymmetry-breaking sector confined to a parallel brane a distance R_5 away, separated by a 5-dimensional bulk, as in Fig. 18.

Figure 18: The separation of the supersymmetry-breaking sector from the MSSM sector could take place along a hidden spatial dimension, as in the simple example shown here. The branes are 4-dimensional parallel spacetime hypersurfaces in a 5-dimensional spacetime.



In this picture, there is an anomalous violation of superconformal (scale) invariance manifested in the running of the couplings. This causes supersymmetry breaking to show up in the MSSM by virtue of the non-zero beta functions and anomalous dimensions of the MSSM brane couplings and fields. The resulting soft terms are (using F_ϕ to denote the VEV of the hidden brane F term from now on):

$$M_a = F_\phi \beta_{g_a} / g_a, \quad (405)$$

$$(m^2)_j^i = \frac{1}{2} |F_\phi|^2 \frac{d}{dt} \gamma_j^i = \frac{1}{2} |F_\phi|^2 \left[\beta_{g_a} \frac{\partial}{\partial g_a} + \beta_{y^{kmn}} \frac{\partial}{\partial y^{kmn}} + \beta_{y_{kmn}^*} \frac{\partial}{\partial y_{kmn}^*} \right] \gamma_j^i \quad (406)$$

$$a^{ijk} = -F_\phi \beta_{y^{ijk}}, \quad (407)$$

where the anomalous dimensions are the γ_j^i . Gaugino masses arise at one-loop order, but scalar squared masses arise at two-loop order.

Also, these results are approximately flavor-blind for the first two families, because the non-trivial flavor structure derives only from the MSSM Yukawa couplings.

There are several unique features of the AMSB scenario.

- First, there is no need to specify at which renormalization scale Eqs. (405)-(407) should be applied as boundary conditions.

This is because they hold at every renormalization scale, exactly, to all orders in perturbation theory. In other words, Eqs. (405)-(407) are not just boundary conditions for the renormalization group equations of the soft parameters, but solutions as well.

(These AMSB renormalization group trajectories can also be found from this renormalization group invariance property alone, without reference to the supergravity derivation.)

- In fact, even if there are heavy supermultiplets in the theory that have to be decoupled, the boundary conditions hold both above and below the arbitrary decoupling scale.

This remarkable insensitivity to ultraviolet physics in AMSB ensures the absence of flavor violation in the low-energy MSSM soft terms.

- Another interesting prediction is that the gravitino mass $m_{3/2}$ in these models is actually much larger than the scale m_{soft} of the

MSSM soft terms, since the latter are loop-suppressed compared to $m_{3/2}$.

There is only one unknown parameter, F_ϕ , among the MSSM soft terms in AMSB. Unfortunately, this exemplary falsifiability is marred by the fact that it is already falsified. The dominant contributions to the first-family squark and slepton squared masses are:

$$m_{\tilde{q}}^2 = \frac{|F_\phi|^2}{(16\pi^2)^2} (8g_3^4 + \dots), \quad (408)$$

$$m_{\tilde{e}_L}^2 = -\frac{|F_\phi|^2}{(16\pi^2)^2} \left(\frac{3}{2}g_2^4 + \frac{99}{50}g_1^4 \right) \quad (409)$$

$$m_{\tilde{e}_R}^2 = -\frac{|F_\phi|^2}{(16\pi^2)^2} \frac{198}{25}g_1^4 \quad (410)$$

The squarks have large positive squared masses, but the sleptons have negative squared masses, so the AMSB model in its simplest

form is not viable. These signs come directly from those of the beta functions of the strong and electroweak gauge interactions, as can be seen from the right side of Eq. (406).

The characteristic ultraviolet insensitivity to physics at high mass scales also makes it somewhat non-trivial to modify the theory to escape this tachyonic slepton problem by deviating from the AMSB trajectory.

There can be large deviations from AMSB provided by supergravity, but then in general the flavor-blindness is also forfeit.

One way to modify AMSB is to introduce additional supermultiplets that contain supersymmetry-breaking mass splittings that are large compared to their average mass. Another way is to combine AMSB with gaugino mediation.

Finally, there is a perhaps less motivated approach in which a common parameter m_0^2 is added to all of the scalar squared masses at some

scale, and chosen large enough to allow the sleptons to have positive squared masses above LEP bounds. This allows the phenomenology to be studied in a framework conveniently parameterized by just:

$$F_\phi, m_0^2, \tan \beta, \arg(\mu), \quad (411)$$

with $|\mu|$ and b determined by requiring correct electroweak symmetry breaking as described in the next section. (Some sources use $m_{3/2}$ or M_{aux} to denote F_ϕ .)

The MSSM gaugino masses at the leading non-trivial order are unaffected by the *ad hoc* addition of m_0^2 :

$$M_1 = \frac{F_\phi}{16\pi^2} \frac{33}{5} g_1^2 \quad (412)$$

$$M_2 = \frac{F_\phi}{16\pi^2} g_2^2 \quad (413)$$

$$M_3 = -\frac{F_\phi}{16\pi^2} 3g_3^2 \quad (414)$$

This implies that $|M_2| \ll |M_1| \ll |M_3|$, so the lightest neutralino is actually mostly wino, with a lightest chargino that is only of order 200 MeV heavier, depending on the values of μ and $\tan\beta$.

The decay $\tilde{C}_1^\pm \rightarrow \tilde{N}_1 \pi^\pm$ produces a very soft pion, implying unique and difficult signatures in colliders.

The mass spectrum of the MSSM

- Electroweak symmetry breaking and the Higgs bosons

In the MSSM, the description of electroweak symmetry breaking is slightly complicated by the fact that there are two complex Higgs doublets $H_u = (H_u^+, H_u^0)$ and $H_d = (H_d^0, H_d^-)$ rather than just one as in the ordinary Standard Model.

The classical scalar potential for the Higgs scalar fields in the MSSM is given by

$$\begin{aligned} V = & (|\mu|^2 + m_{H_u}^2)(|H_u^0|^2 + |H_u^+|^2) + (|\mu|^2 + m_{H_d}^2)(|H_d^0|^2 + |H_d^-|^2) \\ & + [b(H_u^+ H_d^- - H_u^0 H_d^0) + \text{c.c.}] \\ & + \frac{1}{8}(g^2 + g'^2)(|H_u^0|^2 + |H_u^+|^2 - |H_d^0|^2 - |H_d^-|^2)^2 \\ & + \frac{1}{2}g^2 |H_u^+ H_d^{0*} + H_u^0 H_d^{-*}|^2. \end{aligned} \tag{415}$$

We note the following:

- The terms proportional to $|\mu|^2$ come from F -terms [see Eq. (372)].
- The terms proportional to g^2 and g'^2 are the D -term contributions, obtained from the general formula Eq. (352) after some rearranging.
- Finally, the terms proportional to $m_{H_u}^2$, $m_{H_d}^2$ and b are just a rewriting of the last three terms of Eq. (379) using the identity

$$|H_u^i{}^* H_d^i|^2 + |\epsilon_{ij} H_u^i H_d^j|^2 = (H_u^i{}^* H_u^i)(H_d^j{}^* H_d^j) \quad (416)$$

The full scalar potential of the theory also includes many terms involving the squark and slepton fields that we can ignore here, since they do not get VEVs because they have large positive squared masses.

We now have to demand that the minimum of this potential should break electroweak symmetry down to electromagnetism $SU(2)_L \times U(1)_Y \rightarrow U(1)_{\text{EM}}$, in accord with experiment.

We can use the freedom to make gauge transformations to simplify this analysis.

- First, the freedom to make $SU(2)_L$ gauge transformations allows us to rotate away a possible VEV for **one** of the weak isospin components of **one** of the scalar fields, so without loss of generality we can take $H_u^+ = 0$ at the minimum of the potential.
- Then, we can examine the condition for a minimum of the potential satisfying

$$\left. \frac{\partial V}{\partial H_u^+} \right|_{H_u^+=0} = bH_d^- + \frac{1}{2}g^2 H_d^{0*} H_d^- H_u^{0*} = 0. \quad (417)$$

For generic parameter choices this will not vanish unless $H_d^- = 0$. This is good, because it means that at the minimum of the potential electromagnetism is necessarily unbroken, due to the fact that the charged components of the Higgs scalars cannot get VEVs.

- After setting $H_u^+ = H_d^- = 0$, we are left to consider the scalar

potential involving only the neutral Higgs fields:

$$V = (|\mu|^2 + m_{H_u}^2)|H_u^0|^2 + (|\mu|^2 + m_{H_d}^2)|H_d^0|^2 - (b H_u^0 H_d^0 + \text{c.c.}) + \frac{1}{8}(g^2 + g'^2)(|H_u^0|^2 - |H_d^0|^2)^2. \quad (418)$$

- The only term in this potential that depends on the phases of the fields is the b -term.

Therefore, a redefinition of the phase of H_u or H_d can absorb any phase in b , so we can take b to be real and positive.

- Then it is clear that a minimum of the potential V requires that $H_u^0 H_d^0$ is also real and positive, so $\langle H_u^0 \rangle$ and $\langle H_d^0 \rangle$ must have cancelling phases.
- We can therefore use a $U(1)_Y$ gauge transformation to make them both be real and positive without loss of generality, since H_u and H_d have opposite weak hypercharges ($\pm 1/2$).

- It follows that CP cannot be spontaneously broken by the Higgs scalar potential, since the VEVs and b can be simultaneously chosen real, as a convention.
- This means that the Higgs scalar mass eigenstates can be assigned well-defined eigenvalues of CP, at least at tree-level. (CP-violating phases in other couplings can induce loop-suppressed CP violation in the Higgs sector, but do not change the fact that b , $\langle H_u^0 \rangle$, and $\langle H_d \rangle$ can always be chosen real and positive.)

In order for the MSSM scalar potential to be viable, we must first make sure that the potential is bounded from below for arbitrarily large values of the scalar fields, so that V will really have a minimum. (Recall that scalar potentials in purely supersymmetric theories are automatically non-negative and so clearly bounded from below. But, now that we have introduced supersymmetry breaking, we must be careful.)

The scalar quartic interactions in V will stabilize the potential for

almost all arbitrarily large values of H_u^0 and H_d^0 .

However, for the special directions in field space $|H_u^0| = |H_d^0|$, the quartic contributions to V [the second line in Eq. (418)] are identically zero.

Such directions in field space are called D -flat directions, because along them the part of the scalar potential coming from D -terms vanishes.

In order for the potential to be bounded from below, we need the quadratic part of the scalar potential to be positive along the D -flat directions. This requirement amounts to

$$2b < 2|\mu|^2 + m_{H_u}^2 + m_{H_d}^2. \quad (419)$$

Note that the b -term always favors electroweak symmetry breaking.

Requiring that one linear combination of H_u^0 and H_d^0 has a negative

squared mass near $H_u^0 = H_d^0 = 0$ (*i.e.* requiring that the determinant of the mass-squared matrix be negative) gives

$$b^2 > (|\mu|^2 + m_{H_u}^2)(|\mu|^2 + m_{H_d}^2). \quad (420)$$

If this inequality is not satisfied, then $H_u^0 = H_d^0 = 0$ will be a stable minimum of the potential (or there will be no stable minimum at all), and electroweak symmetry breaking will not occur.

Interestingly, if $m_{H_u}^2 = m_{H_d}^2$ then the constraints Eqs. (419) and (420) cannot both be satisfied.

In models derived from the minimal supergravity or gauge-mediated boundary conditions, $m_{H_u}^2 = m_{H_d}^2$ is supposed to hold at tree level at the input scale, but the contribution to the RG equation for $m_{H_u}^2$ proportional to the square of the large top-quark Yukawa coupling y_t naturally pushes $m_{H_u}^2$ to negative or small values $m_{H_u}^2 < m_{H_d}^2$ at the electroweak scale.

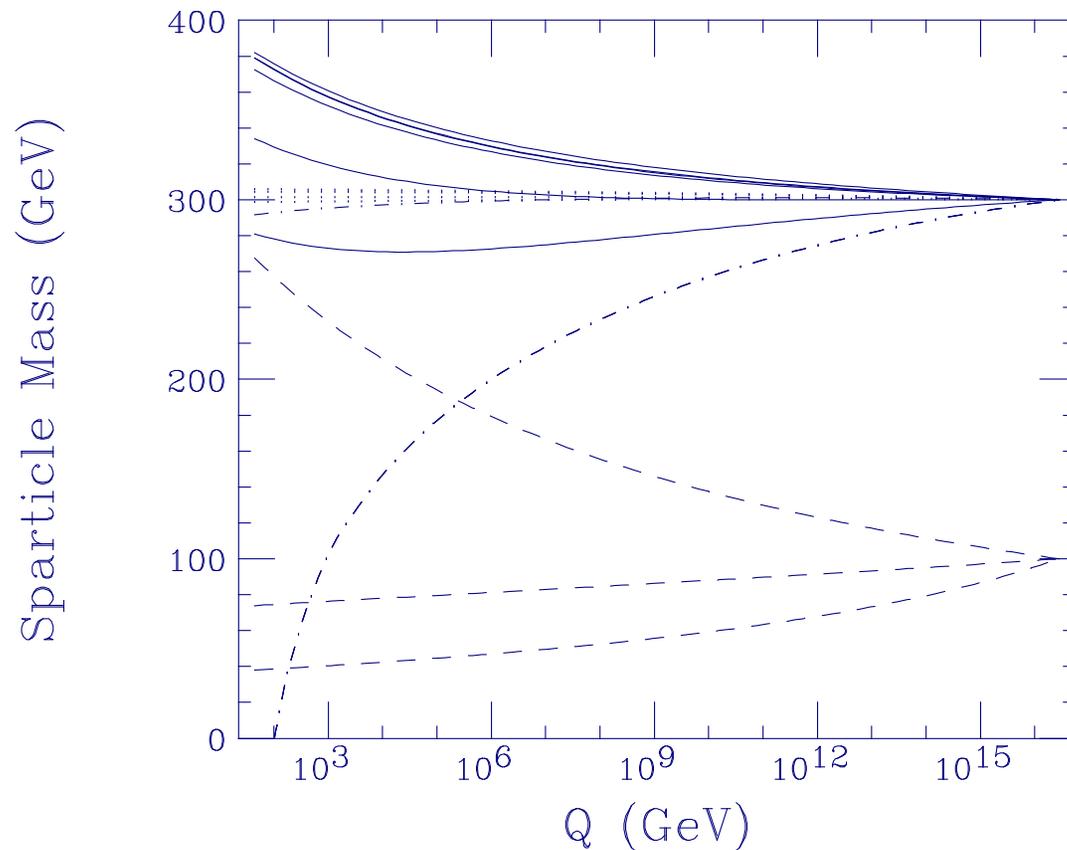


Figure 19: Illustration of RG evolution of soft parameters showing how $m_{H_u}^2$ is driven negative in evolving from GUT scale to m_Z scale. Some other things to note: gaugino masses can unify if $M_3 \sim 3M_2 \sim 6M_1$ at scale m_Z ; squark masses increase as scale decreases, but slepton masses don't change a lot.

Unless this effect is significant, the parameter space in which the electroweak symmetry is broken would be quite small. So, in these models electroweak symmetry breaking is actually driven by quantum corrections; this mechanism is therefore known as *radiative electroweak symmetry breaking*.

Note that although a negative value for $|\mu|^2 + m_{H_u}^2$ will help Eq. (420) to be satisfied, it is not strictly necessary.

Furthermore, even if $m_{H_u}^2 < 0$, there may be no electroweak symmetry breaking if $|\mu|$ is too large or if b is too small.

Still, the large negative contributions to $m_{H_u}^2$ from the RG equation are an important factor in ensuring that electroweak symmetry breaking can occur in models with simple GUT-scale boundary conditions for the soft terms.

The realization that this works most naturally with a large top-quark Yukawa coupling provides additional motivation for these models.

Having established the conditions necessary for H_u^0 and H_d^0 to get non-zero VEVs, we can now require that they are compatible with the observed phenomenology of electroweak symmetry breaking, $SU(2)_L \times U(1)_Y \rightarrow U(1)_{\text{EM}}$. Let us write

$$v_u = \langle H_u^0 \rangle, \quad v_d = \langle H_d^0 \rangle. \quad (421)$$

These VEVs are related to the known mass of the Z^0 boson and the electroweak gauge couplings:

$$v_u^2 + v_d^2 = v^2 = 2m_Z^2 / (g^2 + g'^2) \approx (174 \text{ GeV})^2. \quad (422)$$

The ratio of the VEVs is traditionally written as

$$\tan \beta \equiv v_u / v_d. \quad (423)$$

The value of $\tan \beta$ is not fixed by present experiments, but it depends on the Lagrangian parameters of the MSSM in a calculable way.

Since $v_u = v \sin \beta$ and $v_d = v \cos \beta$ were taken to be real and positive by convention, we have $0 < \beta < \pi/2$, a requirement that will be sharpened below.

Now one can write down the conditions $\partial V / \partial H_u^0 = \partial V / \partial H_d^0 = 0$ under which the potential Eq. (418) will have a minimum satisfying Eqs. (422) and (423):

$$m_{H_u}^2 + |\mu|^2 - b \cot \beta - (m_Z^2/2) \cos(2\beta) = 0, \quad (424)$$

$$m_{H_d}^2 + |\mu|^2 - b \tan \beta + (m_Z^2/2) \cos(2\beta) = 0. \quad (425)$$

It is easy to check that these equations indeed satisfy the necessary conditions Eqs. (419) and (420). They allow us to eliminate two of the Lagrangian parameters b and $|\mu|$ in favor of $\tan \beta$, but do not determine the phase of μ .

Taking $|\mu|^2$, b , $m_{H_u}^2$ and $m_{H_d}^2$ as input parameters, and m_Z^2 and $\tan\beta$ as output parameters obtained by solving these two equations, one obtains:

$$\sin(2\beta) = \frac{2b}{m_{H_u}^2 + m_{H_d}^2 + 2|\mu|^2}, \quad (426)$$

$$m_Z^2 = \frac{|m_{H_d}^2 - m_{H_u}^2|}{\sqrt{1 - \sin^2(2\beta)}} - m_{H_u}^2 - m_{H_d}^2 - 2|\mu|^2. \quad (427)$$

(Note that $\sin(2\beta)$ is always positive. If $m_{H_u}^2 < m_{H_d}^2$, as is usually assumed, then $\cos(2\beta)$ is negative; otherwise it is positive.)

As an aside, Eqs. (426) and (427) highlight the “ μ problem” already mentioned earlier.

- Without miraculous cancellations, all of the input parameters ought to be within an order of magnitude or two of m_Z^2 .

- However, in the MSSM, μ is a supersymmetry-respecting parameter appearing in the superpotential, while b , $m_{H_u}^2$, $m_{H_d}^2$ are supersymmetry breaking parameters.
- This has led to a widespread belief that the MSSM must be extended at very high energies to include a mechanism that relates the effective value of μ to the supersymmetry-breaking mechanism in some way.
- Even if the value of μ is set by soft supersymmetry breaking, the cancellation needed by Eq. (427) is often very substantial (\Rightarrow **finetuning**) when evaluated in specific model frameworks, after constraints from direct searches for the Higgs bosons and superpartners are taken into account.
- For example, expanding for large $\tan \beta$, Eq. (427) becomes

$$m_Z^2 = -2(m_{H_u}^2 + |\mu|^2) + \frac{2}{\tan^2 \beta}(m_{H_d}^2 - m_{H_u}^2) + \mathcal{O}(1/\tan^4 \beta). \quad (428)$$

Typical viable solutions for the MSSM have $-m_{H_u}^2$ and $|\mu|^2$ each much larger than m_Z^2 , so that significant cancellation is needed.

- In particular, large top squark squared masses, needed to avoid having the Higgs boson mass turn out too small [see Eq. (441) below] compared to the direct search limits from LEP, will feed into $m_{H_u}^2$.

The cancellation needed in the minimal model may therefore be at the several per cent level. It is impossible to objectively characterize whether this should be considered worrisome, but it could be taken as a weak hint in favor of non-minimal models.

Radiative corrections to the Higgs masses

The discussion above is based on the tree-level potential, and involves running renormalized Lagrangian parameters, which depend on the choice of renormalization scale.

In practice, one must include radiative corrections at one-loop order,

at least, in order to get numerically stable results.

To do this, one can compute the loop corrections ΔV to the effective potential $V_{\text{eff}}(v_u, v_d) = V + \Delta V$ as a function of the VEVs. The impact of this is that the equations governing the VEVs of the full effective potential are obtained by simply replacing

$$m_{H_u}^2 \rightarrow m_{H_u}^2 + \frac{1}{2v_u} \frac{\partial(\Delta V)}{\partial v_u}, \quad m_{H_d}^2 \rightarrow m_{H_d}^2 + \frac{1}{2v_d} \frac{\partial(\Delta V)}{\partial v_d} \quad (429)$$

in Eqs. (526)-(427), treating v_u and v_d as real variables in the differentiation.

The result for ΔV has now been obtained through two-loop order in the MSSM.

The most important corrections come from the one-loop diagrams involving the top squarks and top quark, and experience shows that the validity of the tree-level approximation and the convergence of

perturbation theory are therefore improved by choosing a renormalization scale roughly of order the average of the top squark masses.

Mass eigenstates

The Higgs scalar fields in the MSSM consist of two complex $SU(2)_L$ -doublet, or eight real, scalar degrees of freedom.

When the electroweak symmetry is broken, three of them are the would-be Nambu-Goldstone bosons G^0 , G^\pm , which become the longitudinal modes of the Z^0 and W^\pm massive vector bosons.

The remaining five Higgs scalar mass eigenstates consist of two CP-even neutral scalars h^0 and H^0 , one CP-odd neutral scalar A^0 , and a charge $+1$ scalar H^+ and its conjugate charge -1 scalar H^- . (Here we define $G^- = G^{+*}$ and $H^- = H^{+*}$. Also, by convention, h^0 is lighter than H^0 .)

The gauge-eigenstate fields can be expressed in terms of the mass

eigenstate fields as:

$$\begin{pmatrix} H_u^0 \\ H_d^0 \end{pmatrix} = \begin{pmatrix} v_u \\ v_d \end{pmatrix} + \frac{1}{\sqrt{2}} R_\alpha \begin{pmatrix} h^0 \\ H^0 \end{pmatrix} + \frac{i}{\sqrt{2}} R_{\beta_0} \begin{pmatrix} G^0 \\ A^0 \end{pmatrix} \quad (430)$$

$$\begin{pmatrix} H_u^+ \\ H_d^{-*} \end{pmatrix} = R_{\beta_\pm} \begin{pmatrix} G^+ \\ H^+ \end{pmatrix} \quad (431)$$

where the orthogonal rotation matrices

$$R_\alpha = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}, \quad (432)$$

$$R_{\beta_0} = \begin{pmatrix} \sin \beta_0 & \cos \beta_0 \\ -\cos \beta_0 & \sin \beta_0 \end{pmatrix}, \quad R_{\beta_\pm} = \begin{pmatrix} \sin \beta_\pm & \cos \beta_\pm \\ -\cos \beta_\pm & \sin \beta_\pm \end{pmatrix}, \quad (433)$$

are chosen so that the quadratic part of the potential has diagonal squared-masses:

$$\begin{aligned} V = & \frac{1}{2} m_{h^0}^2 (h^0)^2 + \frac{1}{2} m_{H^0}^2 (H^0)^2 + \frac{1}{2} m_{G^0}^2 (G^0)^2 + \frac{1}{2} m_{A^0}^2 (A^0)^2 \\ & + m_{G^\pm}^2 |G^\pm|^2 + m_{H^\pm}^2 |H^\pm|^2 + \dots, \end{aligned} \quad (434)$$

Then, provided that v_u, v_d minimize the tree-level potential,¹³ one finds that $\beta_0 = \beta_{\pm} = \beta$, and $m_{G^0}^2 = m_{G^{\pm}}^2 = 0$, and

$$m_{A^0}^2 = 2b/\sin(2\beta) = 2|\mu|^2 + m_{H_u}^2 + m_{H_d}^2 \quad (435)$$

$$m_{h^0, H^0}^2 = \frac{1}{2} \left(m_{A^0}^2 + m_Z^2 \mp \sqrt{(m_{A^0}^2 - m_Z^2)^2 + 4m_Z^2 m_{A^0}^2 \sin^2(2\beta)} \right), \quad (436)$$

$$m_{H^{\pm}}^2 = m_{A^0}^2 + m_W^2. \quad (437)$$

The mixing angle α is determined, at tree-level, by

$$\frac{\sin 2\alpha}{\sin 2\beta} = - \left(\frac{m_{H^0}^2 + m_{h^0}^2}{m_{H^0}^2 - m_{h^0}^2} \right), \quad \frac{\tan 2\alpha}{\tan 2\beta} = \left(\frac{m_{A^0}^2 + m_Z^2}{m_{A^0}^2 - m_Z^2} \right), \quad (438)$$

and is traditionally chosen to be negative; it follows that $-\pi/2 < \alpha < 0$ (provided $m_{A^0} > m_Z$). The Feynman rules for couplings of the mass eigenstate Higgs scalars to the Standard Model quarks and

¹³It is often more useful to expand around VEVs v_u, v_d that do not minimize the tree-level potential, for example to minimize the loop-corrected effective potential instead. In that case, β, β_0 , and β_{\pm} are all slightly different.

leptons and the electroweak vector bosons, as well as to the various sparticles, have been worked out in detail (Gunion-Haber, and HHG).

The masses of A^0 , H^0 and H^\pm can in principle be arbitrarily large since they all grow with $b/\sin(2\beta)$. In contrast, the mass of h^0 is bounded above. From Eq. (436), one finds at tree-level:

$$m_{h^0} < m_Z |\cos(2\beta)| \quad (439)$$

This corresponds to a shallow direction in the scalar potential, along the direction $(H_u^0 - v_u, H_d^0 - v_d) \propto (\cos \alpha, -\sin \alpha)$.

The existence of this shallow direction can be traced to the fact that the quartic Higgs couplings are given by the square of the electroweak gauge couplings, via the D -term.

A contour map of the potential, for a typical case with $\tan \beta \approx -\cot \alpha \approx 10$, is shown in Figure 20.

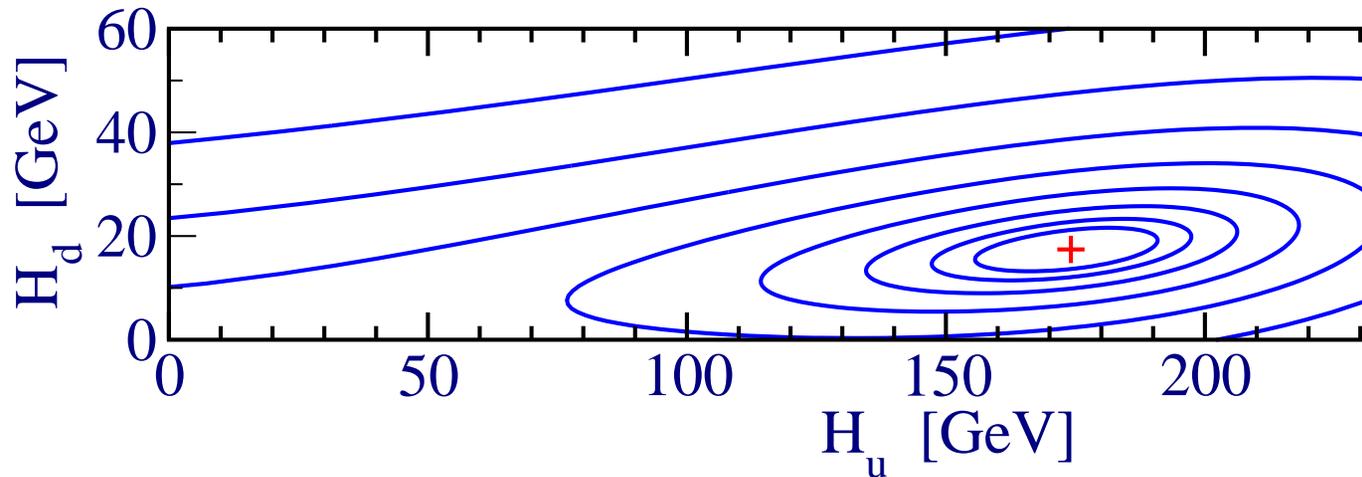


Figure 20: A contour map of the Higgs potential, for a typical case with $\tan \beta \approx -\cot \alpha \approx 10$. The minimum of the potential is marked by $+$, and the contours are equally spaced equipotentials. Oscillations along the shallow direction, with $H_u^0/H_d^0 \approx 10$, correspond to the mass eigenstate h^0 , while the orthogonal steeper direction corresponds to the mass eigenstate H^0 .

If the tree-level inequality (439) were robust, the lightest Higgs boson of the MSSM would have been discovered at LEP2. However, the tree-level formula for the squared mass of h^0 is subject to quantum corrections that are relatively drastic.

correction to Eq. (436):

$$\Delta(m_{h^0}^2) = \frac{3}{4\pi^2} \cos^2 \alpha y_t^2 m_t^2 \ln \left(m_{\tilde{t}_1} m_{\tilde{t}_2} / m_t^2 \right). \quad (440)$$

This shows that m_{h^0} can exceed the LEP bounds.

An alternative way to understand the size of the radiative correction to the h^0 mass is to consider an effective theory in which the heavy top squarks and top quark have been integrated out.

The quartic Higgs couplings in the low-energy effective theory get large positive contributions from the the one-loop diagrams of Fig. 22. This increases the steepness of the Higgs potential, and can be used to obtain the same result for the enhanced h^0 mass.



Figure 22: Integrating out the top quark and top squarks yields large positive contributions to the quartic Higgs coupling in the low-energy effective theory, especially from these one-loop diagrams.

An interesting case, often referred to as the “decoupling limit”, occurs when $m_{A^0} \gg m_Z$.

- Then m_{h^0} can saturate the upper bounds just mentioned, with $m_{h^0}^2 \approx m_Z^2 \cos^2(2\beta) + \text{loop corrections}$.
- The particles A^0 , H^0 , and H^\pm will be much heavier and nearly degenerate, forming an isospin doublet that decouples from sufficiently low-energy experiments.
- The angle α is very nearly $\beta - \pi/2$, and h^0 has the same couplings to quarks and leptons and electroweak gauge bosons as would

the physical Higgs boson of the ordinary Standard Model without supersymmetry.

- Indeed, model-building experiences have shown that it is not uncommon for h^0 to behave in a way nearly indistinguishable from a Standard Model-like Higgs boson, even if m_{A^0} is not too huge.
- However, it should be kept in mind that the couplings of h^0 might turn out to deviate significantly from those of a Standard Model Higgs boson.

Top-squark mixing (which we may discuss later) can result in a further large positive contribution to $m_{h^0}^2$. At one-loop order, and working in the decoupling limit for simplicity, Eq. (440) generalizes to:

$$\begin{aligned}
 m_{h^0}^2 &= m_Z^2 \cos^2(2\beta) \\
 &+ \frac{3}{4\pi^2} \sin^2\beta y_t^2 \left[m_t^2 \ln(m_{\tilde{t}_1} m_{\tilde{t}_2} / m_t^2) + c_t^2 s_t^2 (m_{\tilde{t}_2}^2 - m_{\tilde{t}_1}^2) \ln(m_{\tilde{t}_2}^2 / m_{\tilde{t}_1}^2) \right. \\
 &\left. + c_t^4 s_t^4 \left\{ (m_{\tilde{t}_2}^2 - m_{\tilde{t}_1}^2)^2 - \frac{1}{2} (m_{\tilde{t}_2}^4 - m_{\tilde{t}_1}^4) \ln(m_{\tilde{t}_2}^2 / m_{\tilde{t}_1}^2) \right\} / m_t^2 \right]. \quad (441)
 \end{aligned}$$

Here $c_{\tilde{t}}$ and $s_{\tilde{t}}$ are the cosine and sine of a top squark mixing angle $\theta_{\tilde{t}}$, defined more specifically later on when we discuss the squark sector.

For fixed top-squark masses, the maximum possible h^0 mass occurs for rather large top squark mixing, $c_{\tilde{t}}^2 s_{\tilde{t}}^2 = m_t^2 / [m_{\tilde{t}_2}^2 + m_{\tilde{t}_1}^2 - 2(m_{\tilde{t}_2}^2 - m_{\tilde{t}_1}^2) / \ln(m_{\tilde{t}_2}^2 / m_{\tilde{t}_1}^2)]$ or $1/4$, whichever is less.

It follows that the quantity in square brackets in Eq. (441) is always less than $m_t^2 [\ln(m_{\tilde{t}_2}^2 / m_t^2) + 3]$.

The LEP constraints on the MSSM Higgs sector make the case of large top-squark mixing noteworthy.

Including these and other important corrections one can obtain only a weaker, but still very interesting, bound

$$m_{h^0} \lesssim 135 \text{ GeV} \quad (442)$$

in the MSSM. This assumes that all of the sparticles that can contribute to $m_{h^0}^2$ in loops have masses that do not exceed 1 TeV.

By adding extra supermultiplets to the MSSM, this bound can be made even weaker.

However, assuming that none of the MSSM sparticles have masses exceeding 1 TeV and that all of the couplings in the theory remain perturbative up to the unification scale, one still has

$$m_{h^0} \lesssim 150 \text{ GeV}. \quad (443)$$

This bound is also weakened if, for example, the top squarks are heavier than 1 TeV (but recall $m_{\tilde{t}} < m_{\tilde{u}, \tilde{d}, \dots}$ and coupling unification requires all $\lesssim 1 \text{ TeV}$), but the upper bound rises only logarithmically with the soft masses, as can be seen from Eq. (440).

Thus it is a fairly robust prediction of supersymmetry at the electroweak scale that at least one of the Higgs scalar bosons must be light.

(However, if one is willing to extend the MSSM in a completely general way above the electroweak scale, none of these bounds need apply.)

For a given set of model parameters, it is always important to take into account the complete set of one-loop corrections and even the dominant two-loop effects in order to get reasonably accurate predictions for the Higgs masses and mixings.

In the MSSM, the masses and CKM mixing angles of the quarks and leptons are determined not only by the Yukawa couplings of the superpotential but also the parameter $\tan \beta$.

This is because the top, charm and up quark mass matrix is proportional to $v_u = v \sin \beta$ and the bottom, strange, and down quarks and the charge leptons get masses proportional to $v_d = v \cos \beta$.

At tree-level,

$$m_t = y_t v \sin \beta, \quad m_b = y_b v \cos \beta, \quad m_\tau = y_\tau v \cos \beta. \quad (444)$$

These relations hold for the running masses rather than the physical pole masses, which are significantly larger for t, b .

Including those corrections, one can relate the Yukawa couplings to $\tan \beta$ and the known fermion masses and CKM mixing angles.

It is now clear why we have not neglected y_b and y_τ , even though $m_b, m_\tau \ll m_t$. To a first approximation, $y_b/y_t = (m_b/m_t) \tan \beta$ and $y_\tau/y_t = (m_\tau/m_t) \tan \beta$, so that y_b and y_τ cannot be neglected if $\tan \beta$ is much larger than 1.

In fact, there are good theoretical motivations for considering models with large $\tan \beta$. For example, models based on the GUT gauge group $SO(10)$ can unify the running top, bottom and tau Yukawa

couplings at the unification scale; this requires $\tan\beta$ to be very roughly of order m_t/m_b .

Further notes:

- If one tries to make $\sin\beta$ too small, y_t will be nonperturbatively large.

Requiring that y_t does not blow up above the electroweak scale, one finds that $\tan\beta \gtrsim 1.2$ or so, depending on the mass of the top quark, the QCD coupling, and other details.

- In principle, there is also a constraint on $\cos\beta$ if one requires that y_b and y_τ do not become nonperturbatively large.

This gives a rough upper bound of $\tan\beta \lesssim 65$. However, this is complicated somewhat by the fact that the bottom quark mass gets significant one-loop non-QCD corrections in the large $\tan\beta$ limit.

- One can obtain a stronger upper bound on $\tan\beta$ in some models

- where $m_{H_u}^2 = m_{H_d}^2$ at the GUT or other high energy input scale, by requiring that y_b does not significantly exceed y_t .¹⁵
- The parameter $\tan \beta$ also directly impacts the masses and mixings of the MSSM sparticles, as we will see below.

- Neutralinos and charginos

The higgsinos and electroweak gauginos mix with each other because of the effects of electroweak symmetry breaking.

The neutral higgsinos (\tilde{H}_u^0 and \tilde{H}_d^0) and the neutral gauginos (\tilde{B} , \tilde{W}^0) combine to form four mass eigenstates called *neutralinos*.

The charged higgsinos (\tilde{H}_u^+ and \tilde{H}_d^-) and winos (\tilde{W}^+ and \tilde{W}^-) mix to form two mass eigenstates with charge ± 1 called *charginos*.

¹⁵If y_b were substantially larger than y_t , then the RG evolution equations for the soft-SUSY-breaking masses $m_{H_u}^2, m_{H_d}^2$ that we did not discuss, would imply $m_{H_d}^2 < m_{H_u}^2$ at the electroweak scale. In this case, the minimum of the potential would have $\langle H_d^0 \rangle > \langle H_u^0 \rangle$, which would be a contradiction with the supposition that $\tan \beta$ is large.

We will denote¹⁶ the neutralino and chargino mass eigenstates by \tilde{N}_i ($i = 1, 2, 3, 4$) and \tilde{C}_i^\pm ($i = 1, 2$).

By convention, these are labeled in ascending order, so that $m_{\tilde{N}_1} < m_{\tilde{N}_2} < m_{\tilde{N}_3} < m_{\tilde{N}_4}$ and $m_{\tilde{C}_1} < m_{\tilde{C}_2}$.

The lightest neutralino, \tilde{N}_1 , is usually assumed to be the LSP, unless there is a lighter gravitino or unless R -parity is not conserved, because it is the only MSSM particle that can make a good dark matter candidate.

We will now describe the mass spectrum and mixing of the neutralinos and charginos in the MSSM.

In the gauge-eigenstate basis $\psi^0 = (\tilde{B}, \tilde{W}^0, \tilde{H}_d^0, \tilde{H}_u^0)$, the neutralino

¹⁶Other common notations use $\tilde{\chi}_i^0$ or \tilde{Z}_i for neutralinos, and $\tilde{\chi}_i^\pm$ or \tilde{W}_i^\pm for charginos.

mass part of the Lagrangian is

$$\mathcal{L}_{\text{neutralino mass}} = -\frac{1}{2}(\psi^0)^T \mathbf{M}_{\tilde{N}} \psi^0 + \text{c.c.}, \quad (445)$$

where

$$\mathbf{M}_{\tilde{N}} = \begin{pmatrix} M_1 & 0 & -g'v_d/\sqrt{2} & g'v_u/\sqrt{2} \\ 0 & M_2 & gv_d/\sqrt{2} & -gv_u/\sqrt{2} \\ -g'v_d/\sqrt{2} & gv_d/\sqrt{2} & 0 & -\mu \\ g'v_u/\sqrt{2} & -gv_u/\sqrt{2} & -\mu & 0 \end{pmatrix}. \quad (446)$$

The entries M_1 and M_2 in this matrix come directly from the MSSM soft Lagrangian [see Eq. (379)], while the entries $-\mu$ are the supersymmetric higgsino mass terms [see Eq. (468)]. The terms proportional to g, g' are the result of Higgs-higgsino-gaugino couplings [see Eq. (356) and Figure 7g,h], with the Higgs scalars replaced by their VEVs [Eqs. (422), (423)].

This can also be written as

$$\mathbf{M}_{\tilde{N}} = \begin{pmatrix} M_1 & 0 & -c_\beta s_W m_Z & s_\beta s_W m_Z \\ 0 & M_2 & c_\beta c_W m_Z & -s_\beta c_W m_Z \\ -c_\beta s_W m_Z & c_\beta c_W m_Z & 0 & -\mu \\ s_\beta s_W m_Z & -s_\beta c_W m_Z & -\mu & 0 \end{pmatrix}. \quad (447)$$

Here we have introduced abbreviations $s_\beta = \sin \beta$, $c_\beta = \cos \beta$, $s_W = \sin \theta_W$, and $c_W = \cos \theta_W$.

Some technical details that were needed above.

1. First, the above mass matrix is, by convention, being written in terms of the 2-component spinors corresponding to the various states.

The c.c. part of Eq. (445) provides the full $\bar{\Psi}\Psi = \xi\xi + \xi^\dagger\xi^\dagger$ that would be appropriate in 4-component notation in the case of a Majorana fermion (recall $\Psi_M = \begin{pmatrix} \xi \\ \xi^\dagger \end{pmatrix}$, see Eq. (278)).

2. Second, since we are dealing with Majorana particles, the correct normalization of a mass term is that given earlier:

$$\mathcal{L}_{\text{Majorana}} = \frac{i}{2} \bar{\Psi}_M \gamma^\mu \partial_\mu \Psi_M - \frac{1}{2} M \bar{\Psi}_M \Psi_M \quad (448)$$

which in two-component notation reads

$$\mathcal{L}_{\text{Majorana}} = i \xi^\dagger \bar{\sigma}^\mu \partial_\mu \xi - \frac{1}{2} M (\xi \xi + \xi^\dagger \xi^\dagger) \quad (449)$$

3. This is why we wrote the soft-SUSY-breaking terms, *e.g.* for gauginos, in the form $\mathcal{L} \ni -\frac{1}{2} M_a (\lambda^a \lambda^a + \lambda^{a\dagger} \lambda^{a\dagger})$, where the λ^a were the 2-component objects.

This, hopefully, makes the normalization of the M_1 , M_2 and μ entries obvious.

4. It is the off-diagonal entries that are a bit tricky. These come from the

$$\mathcal{L} \ni -\sqrt{2} g \phi^* T^a \psi \lambda^a + h.c. \quad (450)$$

part of Eq. (356), where g is the coupling for whatever group we focus on, as follows.

In the above, the λ^a and ψ are (again) the two-component spinors.

5. Let us consider the ψ that goes with H_u , denoted by ψ_{H_u} , and the $SU(2)_L$ group with $g = g_2$. Writing out the above (without c.c. term) for $a = 3$ (as relevant for the neutral guy for which the 2-component spinor is λ^3) using $T^3 = \frac{1}{2}\tau^3$ gives us (standard spinor contractions are implied for example in writing $\psi_{H_u^0}\lambda^3$)

$$\begin{aligned}
 \mathcal{L} &\ni -\sqrt{2}g_2 (H_u^+ \quad H_u^0)^* \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} \psi_{H_u^+} \\ \psi_{H_u^0} \end{pmatrix} \lambda^3 \\
 &\ni \frac{1}{\sqrt{2}}g_2 v_u \psi_{H_u^0} \lambda^3 \\
 &= \frac{g}{\sqrt{2}} v_u \tilde{H}_u^0 \tilde{W}^0, \tag{451}
 \end{aligned}$$

once we identify g_2 as the usual $SU(2)_L$ coupling g and we convert

to Martin's notation of $\psi_{H_u^0} = \tilde{H}_u^0$ and $\lambda^3 = \tilde{W}^0$.

Matching to the basic defining form of $\mathcal{L} \ni -\frac{1}{2}(\psi^0)^T \mathbf{M}_{\tilde{N}} \psi^0$ and distributing to the two off-diagonal entries that contribute to $\tilde{H}_u^0 \tilde{W}^0$ gives us the indicated row-2, col-4 and row-4, col-2 entries of $-gv_u/\sqrt{2}$.

6. As another example, let us consider the ψ_{H_d} and the $U(1)$ contribution for which $g \rightarrow g'$, T^a is replaced by Y (with $Y = -\frac{1}{2}$ for the H_d stuff) and the 2-component gaugino field in question is λ' . Writing things out gives

$$\begin{aligned}
 \mathcal{L} &\ni -\sqrt{2}g' (H_d^0 \quad H_d^-)^* \left(-\frac{1}{2}\right) \mathbf{1}_{2 \times 2} \begin{pmatrix} \psi_{H_d^0} \\ \psi_{H_d^-} \end{pmatrix} \lambda' \\
 &\ni +\frac{1}{\sqrt{2}}g'v_d\psi_{H_d^0}\lambda' \\
 &= +\frac{g'}{\sqrt{2}}v_d\tilde{H}_d^0\tilde{B},
 \end{aligned} \tag{452}$$

Matching to the basic defining form of $\mathcal{L} \ni -\frac{1}{2}(\psi^0)^T \mathbf{M}_{\tilde{N}} \psi^0$ and distributing to the two off-diagonal entries that contribute to $\tilde{H}_d^0 \tilde{B}$ gives us the indicated row-1, col-3 and row-3, col-1 entries of $-g'v_d/\sqrt{2}$.

7. And so forth.

The mass matrix $\mathbf{M}_{\tilde{N}}$ can be diagonalized by a unitary matrix \mathbf{N} to obtain mass eigenstates:

$$\tilde{N}_i = \mathbf{N}_{ij} \psi_j^0, \quad (453)$$

so that

$$\mathbf{N}^* \mathbf{M}_{\tilde{N}} \mathbf{N}^{-1} = \begin{pmatrix} m_{\tilde{N}_1} & 0 & 0 & 0 \\ 0 & m_{\tilde{N}_2} & 0 & 0 \\ 0 & 0 & m_{\tilde{N}_3} & 0 \\ 0 & 0 & 0 & m_{\tilde{N}_4} \end{pmatrix} \quad (454)$$

has real positive entries on the diagonal. These are the magnitudes

of the eigenvalues of $\mathbf{M}_{\tilde{N}}$, or equivalently the square roots of the eigenvalues of $\mathbf{M}_{\tilde{N}}^\dagger \mathbf{M}_{\tilde{N}}$. The indices (i, j) on \mathbf{N}_{ij} are (mass, gauge) eigenstate labels.

The mass eigenvalues and the mixing matrix \mathbf{N}_{ij} can be given in closed form in terms of the parameters M_1 , M_2 , μ and $\tan\beta$, by solving quartic equations, but the results are very complicated and not illuminating except in certain limits.

In general, the parameters M_1 , M_2 , and μ in the equations above can have arbitrary complex phases.

A redefinition of the phases of \tilde{B} and \tilde{W} always allows us to choose a convention in which M_1 and M_2 are both real and positive.

The phase of μ within that convention is then really a physical parameter and cannot be rotated away. [We have already used up the freedom to redefine the phases of the Higgs fields, since we have picked b and $\langle H_u^0 \rangle$ and $\langle H_d^0 \rangle$ to be real and positive, to guarantee

that the off-diagonal entries in Eq. (619) proportional to m_Z are real.]

However, if μ is not real, then there can be potentially disastrous CP-violating effects in low-energy physics, including electric dipole moments for both the electron and the neutron.

Therefore, it is usual [although not strictly mandatory, because of the possibility of nontrivial cancellations involving the phases of the (scalar)³ couplings and the gluino mass] to assume that μ is real in the same set of phase conventions that make M_1 , M_2 , b , $\langle H_u^0 \rangle$ and $\langle H_d^0 \rangle$ real and positive. The sign of μ is still undetermined by this constraint.

In models where the gaugino masses are unified at the GUT scale, one has the nice prediction

$$M_1 \approx \frac{5}{3} \tan^2 \theta_W M_2 \approx 0.5 M_2 \quad (455)$$

at the electroweak scale, as was mentioned earlier.

If so, then the neutralino masses and mixing angles depend on only three unknown parameters.

This assumption is sufficiently theoretically compelling that it has been made in many phenomenological studies; nevertheless it should be recognized as an assumption, to be tested someday by experiment.

There is a not-unlikely limit in which electroweak symmetry breaking effects can be viewed as a small perturbation on the neutralino mass matrix. If (see Gunion+Haber paper for details)

$$m_Z \ll |\mu \pm M_1|, |\mu \pm M_2|, \quad (456)$$

then the neutralino mass eigenstates are very nearly a “bino-like” $\tilde{N}_1 \approx \tilde{B}$; a “wino-like” $\tilde{N}_2 \approx \tilde{W}^0$; and “higgsino-like” $\tilde{N}_3, \tilde{N}_4 \approx (\tilde{H}_u^0 \pm \tilde{H}_d^0)/\sqrt{2}$, with mass eigenvalues:

$$m_{\tilde{N}_1} = M_1 - \frac{m_Z^2 s_W^2 (M_1 + \mu \sin 2\beta)}{\mu^2 - M_1^2} + \dots \quad (457)$$

$$m_{\tilde{N}_2} = M_2 - \frac{m_W^2(M_2 + \mu \sin 2\beta)}{\mu^2 - M_2^2} + \dots \quad (458)$$

$$m_{\tilde{N}_3}, m_{\tilde{N}_4} = |\mu| + \frac{m_Z^2(I - \sin 2\beta)(\mu + M_1 c_W^2 + M_2 s_W^2)}{2(\mu + M_1)(\mu + M_2)} + \dots, \quad (459)$$

$$|\mu| + \frac{m_Z^2(I + \sin 2\beta)(\mu - M_1 c_W^2 - M_2 s_W^2)}{2(\mu - M_1)(\mu - M_2)} + \dots \quad (460)$$

where we have taken M_1 and M_2 real and positive by convention, and assumed μ is real with sign $I = \pm 1$.

The subscript labels of the mass eigenstates may need to be rearranged depending on the numerical values of the parameters; in particular the above labeling of \tilde{N}_1 and \tilde{N}_2 assumes $M_1 < M_2 \ll |\mu|$.

This limit, leading to a bino-like neutralino LSP, often emerges from minimal supergravity boundary conditions on the soft parameters, which tend to require it in order to get correct electroweak symmetry breaking.

It will later be useful to know some entries in the diagonalizing matrix. In the limit that the above masses are obtained, one finds

$$N = \begin{pmatrix} 1 & N_{12} & \frac{-m_Z s_W (M_1 c_\beta + \mu s_\beta)}{M_1^2 - \mu^2} & \frac{m_Z s_W (M_1 s_\beta + \mu c_\beta)}{M_1^2 - \mu^2} \\ N_{21} & 1 & \frac{m_Z c_W (M_2 c_\beta + \mu s_\beta)}{M_2^2 - \mu^2} & \frac{-m_Z c_W (M_2 s_\beta + \mu c_\beta)}{M_2^2 - \mu^2} \\ \frac{-m_Z s_W (s_\beta - c_\beta)}{\sqrt{2}(\mu + M_1)} & \frac{m_Z c_W (s_\beta - c_\beta)}{\sqrt{2}(\mu + M_2)} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-m_Z s_W (s_\beta + c_\beta)}{\sqrt{2}(\mu - M_1)} & \frac{m_Z c_W (s_\beta + c_\beta)}{\sqrt{2}(\mu - M_2)} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{2}} \end{pmatrix} \quad (461)$$

where

$$\begin{aligned} N_{12} &= \frac{m_Z^2 \sin 2\theta_W (M_1 + \mu \sin 2\beta)}{2(M_1 - M_2)(\mu^2 - M_1^2)} \\ N_{21} &= \frac{m_Z^2 \sin 2\theta_W (M_2 + \mu \sin 2\beta)}{(M_2 - M_1)(\mu^2 - M_2^2)} \end{aligned} \quad (462)$$

where the entries are labelled in the same order as the masses are given — reordering might be necessary if the $\tilde{\chi}_1^0$ is not the LSP.

The chargino spectrum can be analyzed in a similar way. In the gauge-eigenstate basis $(\psi^\pm)^T = (\widetilde{W}^+, \widetilde{H}_u^+, \widetilde{W}^-, \widetilde{H}_d^-)$, which combines the positively and negatively charged states into a single vector, the chargino mass terms in the Lagrangian are

$$\mathcal{L}_{\text{chargino mass}} = -\frac{1}{2}(\psi^\pm)^T \mathbf{M}_{\widetilde{C}} \psi^\pm + \text{c.c.} \quad (463)$$

where, in 2×2 off-diagonal (so that negatively charged states are connected to positively charged states) block form,

$$\mathbf{M}_{\widetilde{C}} = \begin{pmatrix} \mathbf{0} & \mathbf{X}^T \\ \mathbf{X} & \mathbf{0} \end{pmatrix}, \quad (464)$$

with

$$\mathbf{X} = \begin{pmatrix} M_2 & gv_u \\ gv_d & \mu \end{pmatrix} = \begin{pmatrix} M_2 & \sqrt{2}s_\beta m_W \\ \sqrt{2}c_\beta m_W & \mu \end{pmatrix}. \quad (465)$$

Writing this out at the \mathbf{X} level gives

$$\mathcal{L}_{\text{chargino mass}} = -\frac{1}{2} \left[(\widetilde{W}^+ \quad \widetilde{H}_u^+) \mathbf{X}^T \begin{pmatrix} \widetilde{W}^- \\ \widetilde{H}_d^- \end{pmatrix} + (\widetilde{W}^- \quad \widetilde{H}_d^-) \mathbf{X} \begin{pmatrix} \widetilde{W}^+ \\ \widetilde{H}_u^+ \end{pmatrix} \right] + \text{c.c.} \quad (466)$$

1. Note that for the diagonal terms of \mathbf{X} , there is a doubling because of having both \mathbf{X}^T and \mathbf{X} terms above. So, for example, the terms proportional to μ in the above are

$$\mathcal{L}_{\text{chargino mass}} \ni -\frac{1}{2}\mu \left[\tilde{H}_u^+ \tilde{H}_d^- + \tilde{H}_d^- \tilde{H}_u^+ \right] \quad (467)$$

which matches (after using the standard 2-component spinor identity $\tilde{H}_u^+ \tilde{H}_d^- = \tilde{H}_d^- \tilde{H}_u^+$) the μ term appearing in the charged-field part of the equation given earlier as part of the supersymmetry invariant Lagrangian:

$$- \mathcal{L}_{\text{higgsino mass}} = \mu(\tilde{H}_u^+ \tilde{H}_d^- - \tilde{H}_u^0 \tilde{H}_d^0) + \text{c.c.} \quad (468)$$

2. The M_2 terms above are

$$\mathcal{L}_{\text{chargino mass}} \ni -\frac{1}{2}M_2 \left[\tilde{W}^+ \tilde{W}^- + \tilde{W}^- \tilde{W}^+ \right] \quad (469)$$

which matches the soft-SUSY-breaking term written previously as

$$\mathcal{L} \ni -\frac{1}{2}M_2(\lambda^1\lambda^1 + \lambda^2\lambda^2) = -M_2\lambda^+\lambda^- \quad (470)$$

where $\lambda^\pm = \frac{1}{\sqrt{2}}(\lambda^1 \mp i\lambda^2)$ are denoted by \widetilde{W}^\pm by Martin, the spartners of $W^\pm = \frac{1}{\sqrt{2}}(W^1 \mp iW^2)$.

3. The off-diagonal entries comes from reducing the

$$\mathcal{L} \ni -\sqrt{2}g_2\phi^*T^a\psi\lambda^a \quad (471)$$

part of the supersymmetric Lagrangian for $a = (1 \pm i2)/\sqrt{2}$ with $g_2 \rightarrow g$.

The mass eigenstates are related to the gauge eigenstates by two unitary 2×2 matrices \mathbf{U} and \mathbf{V} according to

$$\begin{pmatrix} \widetilde{C}_1^+ \\ \widetilde{C}_2^+ \end{pmatrix} = \mathbf{V} \begin{pmatrix} \widetilde{W}^+ \\ \widetilde{H}_u^+ \end{pmatrix}, \quad \begin{pmatrix} \widetilde{C}_1^- \\ \widetilde{C}_2^- \end{pmatrix} = \mathbf{U} \begin{pmatrix} \widetilde{W}^- \\ \widetilde{H}_d^- \end{pmatrix}. \quad (472)$$

Note that since $\mathbf{X}^T \neq \mathbf{X}$ the mixing matrix for the positively charged left-handed fermions is different from that for the negatively charged left-handed fermions. They are chosen so that, for example, the term

$$\begin{aligned} (\widetilde{W}^- \quad \widetilde{H}_d^-) \mathbf{X} \begin{pmatrix} \widetilde{W}^+ \\ \widetilde{H}_u^+ \end{pmatrix} &= (\widetilde{C}_1^- \quad \widetilde{C}_2^-) \mathbf{U}^T \mathbf{X} \mathbf{V}^{-1} \begin{pmatrix} \widetilde{C}_1^+ \\ \widetilde{C}_2^+ \end{pmatrix} \\ &= (\widetilde{C}_1^- \quad \widetilde{C}_2^-) \mathbf{U}^* \mathbf{X} \mathbf{V}^{-1} \begin{pmatrix} \widetilde{C}_1^+ \\ \widetilde{C}_2^+ \end{pmatrix} \end{aligned} \quad (473)$$

(the latter follows since $\mathbf{U}^T = \mathbf{U}^{\dagger*} = \mathbf{U}^*$) reduces this mass term to a diagonal form

$$\mathbf{U}^* \mathbf{X} \mathbf{V}^{-1} = \begin{pmatrix} m_{\widetilde{C}_1} & 0 \\ 0 & m_{\widetilde{C}_2} \end{pmatrix}, \quad (474)$$

with positive real entries $m_{\widetilde{C}_i}$. Because these are only 2×2 matrices, it is not hard to solve for the masses explicitly:

$$\begin{aligned} m_{\widetilde{C}_1}^2, m_{\widetilde{C}_2}^2 &= \frac{1}{2} \left[|M_2|^2 + |\mu|^2 + 2m_W^2 \right. \\ &\quad \left. \mp \sqrt{(|M_2|^2 + |\mu|^2 + 2m_W^2)^2 - 4|\mu M_2 - m_W^2 \sin 2\beta|^2} \right]. \end{aligned} \quad (475)$$

These are the (doubly degenerate) eigenvalues of the 4×4 matrix $\mathbf{M}_{\tilde{C}}^\dagger \mathbf{M}_{\tilde{C}}$, or equivalently the eigenvalues of $\mathbf{X}^\dagger \mathbf{X}$. In particular, Eq. (474) also implies (recall $\mathbf{V}^{-1} = \mathbf{V}^\dagger$)

$$\mathbf{V} \mathbf{X}^\dagger \mathbf{U}^T = \begin{pmatrix} m_{\tilde{C}_1} & 0 \\ 0 & m_{\tilde{C}_2} \end{pmatrix}, \quad (476)$$

so that multiplying Eq. (474) times Eq. (476) gives the 2nd relation below with a similar derivation for the first relation below.

$$\mathbf{V} \mathbf{X}^\dagger \mathbf{X} \mathbf{V}^{-1} = \mathbf{U}^* \mathbf{X} \mathbf{X}^\dagger \mathbf{U}^T = \begin{pmatrix} m_{\tilde{C}_1}^2 & 0 \\ 0 & m_{\tilde{C}_2}^2 \end{pmatrix}. \quad (477)$$

(But, they are *not* the squares of the eigenvalues of \mathbf{X} .) In the limit of Eq. (456) with real M_2 and μ , the chargino mass eigenstates consist of a wino-like \tilde{C}_1^\pm and a higgsino-like \tilde{C}_2^\pm , with masses

$$m_{\tilde{C}_1} = M_2 - \frac{m_W^2 (M_2 + \mu \sin 2\beta)}{\mu^2 - M_2^2} + \dots \quad (478)$$

$$m_{\tilde{C}_2} = |\mu| + \frac{m_W^2 I(\mu + M_2 \sin 2\beta)}{\mu^2 - M_2^2} + \dots \quad (479)$$

Here again the labeling assumes $M_2 < |\mu|$, and I is the sign of μ .

Amusingly, \tilde{C}_1 is nearly degenerate with the neutralino \tilde{N}_2 in the approximation shown, but that is not an exact result.

Their higgsino-like colleagues \tilde{N}_3 , \tilde{N}_4 and \tilde{C}_2 have masses of order $|\mu|$.

The case of $M_1 \approx 0.5M_2 \ll |\mu|$ is not uncommonly found in viable models following from universal boundary conditions, and it has been elevated to the status of a benchmark framework in many phenomenological studies.

However it cannot be overemphasized that such expectations are not mandatory.

The Feynman rules involving neutralinos and charginos may be

inferred in terms of \mathbf{N} , \mathbf{U} and \mathbf{V} from the MSSM Lagrangian as discussed above; they are collected in the Haber-Kane Physics Report and in papers by Gunion and Haber. Feynman rules based on two-component spinor notation have also recently been given in the Dreiner, Haber, Martin review.

In practice, the masses and mixing angles for the neutralinos and charginos are best computed numerically. Note that the discussion above yields the tree-level masses. Loop corrections to these masses can be significant, and have been found systematically at one-loop order.

- The gluino

The gluino is a color octet fermion, so it cannot mix with any other particle in the MSSM, even if R -parity is violated. In this regard, it is unique among all of the MSSM sparticles. In models with minimal supergravity or gauge-mediated boundary conditions, the gluino mass

parameter M_3 is related to the bino and wino mass parameters M_1 and M_2 by

$$M_3 = \frac{g_3^2}{g_2^2} M_2 = \frac{\alpha_s}{\alpha} \sin^2 \theta_W M_2 = \frac{3\alpha_s}{5\alpha} \cos^2 \theta_W M_1 \quad (480)$$

at any RG scale, up to small two-loop corrections. This implies a rough prediction

$$M_3 : M_2 : M_1 \approx 6 : 2 : 1 \quad (481)$$

near the TeV scale. It is therefore reasonable to suspect that the gluino is considerably heavier than the lighter neutralinos and charginos (even in many models where the gaugino mass unification condition is not imposed).

Also, it will be useful to recall (when comparing to squark expectations to come) that

$$M_3 \sim \frac{g_3^2}{g_U^2} m_{1/2} \sim 3m_{1/2}. \quad (482)$$

For more precise estimates, one must take into account the fact that M_3 is really a running mass parameter with an implicit dependence on the RG scale Q .

Because the gluino is a strongly interacting particle, M_3 runs rather quickly with Q .

A more useful quantity physically is the RG scale-independent mass $m_{\tilde{g}}$ at which the renormalized gluino propagator has a pole. Including one-loop corrections to the gluino propagator due to gluon exchange and quark-squark loops, one finds that the pole mass is given in terms of the running mass in the $\overline{\text{DR}}$ scheme by

$$m_{\tilde{g}} = M_3(Q) \left(1 + \frac{\alpha_s}{4\pi} [15 + 6 \ln(Q/M_3) + \sum A_{\tilde{q}}] \right) \quad (483)$$

where

$$A_{\tilde{q}} = \int_0^1 dx \, x \ln [x m_{\tilde{q}}^2 / M_3^2 + (1-x) m_q^2 / M_3^2 - x(1-x) - i\epsilon]. \quad (484)$$

The sum in Eq. (483) is over all 12 squark-quark supermultiplets, and we have neglected small effects due to squark mixing.

The correction terms proportional to α_s in Eq. (483) can be quite significant, because the gluino is strongly interacting, with a large group theory factor [the 15 in Eq. (483)] due to its color octet nature, and because it couples to all of the squark-quark pairs.

The leading two-loop corrections to the gluino pole mass have also been found, and typically increase the prediction by another 1 or 2%.

- The squarks and sleptons

In principle, any scalars with the same electric charge, R -parity, and color quantum numbers can mix with each other. This means that with completely arbitrary soft terms, the mass eigenstates of the squarks and sleptons of the MSSM should be obtained by diagonalizing three 6×6 squared-mass matrices for up-type squarks

$(\tilde{u}_L, \tilde{c}_L, \tilde{t}_L, \tilde{u}_R, \tilde{c}_R, \tilde{t}_R)$, down-type squarks $(\tilde{d}_L, \tilde{s}_L, \tilde{b}_L, \tilde{d}_R, \tilde{s}_R, \tilde{b}_R)$, and charged sleptons $(\tilde{e}_L, \tilde{\mu}_L, \tilde{\tau}_L, \tilde{e}_R, \tilde{\mu}_R, \tilde{\tau}_R)$, and one 3×3 matrix for sneutrinos $(\tilde{\nu}_e, \tilde{\nu}_\mu, \tilde{\nu}_\tau)$.

Fortunately, the general hypothesis of flavor-blind soft parameters Eqs. (382) and (383) predicts that most of these mixing angles are very small.

The third-family squarks and sleptons can have very different masses compared to their first- and second-family counterparts, because of the effects of large Yukawa (y_t, y_b, y_τ) and soft (a_t, a_b, a_τ) couplings in the RG equations. Furthermore, they can have substantial mixing in pairs $(\tilde{t}_L, \tilde{t}_R)$, $(\tilde{b}_L, \tilde{b}_R)$ and $(\tilde{\tau}_L, \tilde{\tau}_R)$.

In contrast, the first- and second-family squarks and sleptons have negligible Yukawa couplings, so they end up in 7 very nearly degenerate, unmixed pairs $(\tilde{e}_R, \tilde{\mu}_R)$, $(\tilde{\nu}_e, \tilde{\nu}_\mu)$, $(\tilde{e}_L, \tilde{\mu}_L)$, $(\tilde{u}_R, \tilde{c}_R)$, $(\tilde{d}_R, \tilde{s}_R)$, $(\tilde{u}_L, \tilde{c}_L)$, $(\tilde{d}_L, \tilde{s}_L)$.

This avoids the problem of disastrously large virtual sparticle contribution to flavor-changing processes.

Let us first consider the spectrum of first- and second-family squarks and sleptons. In many models, including both minimal supergravity [Eq. (390)] and gauge-mediated [Eq. (397)] boundary conditions, their running squared masses can be conveniently parameterized, to a good approximation, as:

$$m_{Q_1}^2 = m_{Q_2}^2 = m_0^2 + K_3 + K_2 + \frac{1}{36}K_1, \quad (485)$$

$$m_{u_1}^2 = m_{u_2}^2 = m_0^2 + K_3 + \frac{4}{9}K_1, \quad (486)$$

$$m_{d_1}^2 = m_{d_2}^2 = m_0^2 + K_3 + \frac{1}{9}K_1, \quad (487)$$

$$m_{L_1}^2 = m_{L_2}^2 = m_0^2 + K_2 + \frac{1}{4}K_1, \quad (488)$$

$$m_{e_1}^2 = m_{e_2}^2 = m_0^2 + K_1. \quad (489)$$

A key point is that the same K_3 , K_2 and K_1 appear everywhere in Eqs. (485)-(489), since all of the chiral supermultiplets couple to the same gauginos with the same gauge couplings. The different coefficients in front of K_1 just correspond to the various values of weak hypercharge squared for each scalar.

In **minimal supergravity models**, m_0^2 is the same common scalar squared mass appearing in Eq. (390). It can be very small, as in the “no-scale” limit, but it could also be the dominant source of the scalar masses. The contributions K_3 , K_2 and K_1 are due to the RG running proportional to the gaugino masses. Explicitly, they are found at one loop order to take the form:

$$K_a(Q) = \begin{Bmatrix} 3/5 \\ 3/4 \\ 4/3 \end{Bmatrix} \times \frac{1}{2\pi^2} \int_{\ln Q}^{\ln Q_0} dt g_a^2(t) |M_a(t)|^2 \quad (a = 1, 2, 3). \quad (490)$$

Here Q_0 is the input RG scale at which the minimal supergravity boundary condition Eq. (390) is applied, and Q should be taken to be evaluated near the squark and slepton mass under consideration, presumably less than about 1 TeV.

The running parameters $g_a(Q)$ and $M_a(Q)$ obey Eqs. (385) and (388). If the input scale is approximated by the apparent scale of gauge coupling unification $Q_0 = M_U \approx 2 \times 10^{16}$ GeV, one finds that numerically

$$K_1 \approx 0.15m_{1/2}^2, \quad K_2 \approx 0.5m_{1/2}^2, \quad K_3 \approx (4.5 \text{ to } 6.5)m_{1/2}^2. \quad (491)$$

for Q near the electroweak scale. Here $m_{1/2}$ is the common gaugino mass parameter at the unification scale. If m_0 is small, and using $m_{\tilde{g}} \sim M_3 \sim 3m_{1/2}$, we see that squark and gluino masses would typically be similar in magnitude.

Note that $K_3 \gg K_2 \gg K_1$; this is a direct consequence of the relative sizes of the gauge couplings g_3 , g_2 , and g_1 .

The large uncertainty in K_3 is due in part to the experimental uncertainty in the QCD coupling constant, and in part to the uncertainty in where to choose Q , since K_3 runs rather quickly below 1 TeV.

If the gauge couplings and gaugino masses are unified between M_U and M_P , as would occur in a GUT model, then the effect of RG running for $M_U < Q < M_P$ can be absorbed into a redefinition of m_0^2 . Otherwise, it adds a further uncertainty roughly proportional to $\ln(M_P/M_U)$, compared to the larger contributions in Eq. (490), which go roughly like $\ln(M_U/1 \text{ TeV})$.

In **gauge-mediated models**, the same parameterization Eqs. (485)-(489) holds, but m_0^2 is always 0.

At the input scale Q_0 , each MSSM scalar gets contributions to its squared mass that depend only on its gauge interactions. It is not hard to see that in general these contribute in exactly the same

pattern as K_1 , K_2 , and K_3 in Eq. (485)-(489).

The subsequent evolution of the scalar squared masses down to the electroweak scale again just yields more contributions to the K_1 , K_2 , and K_3 parameters.

It is somewhat more difficult to give meaningful numerical estimates for these parameters in gauge-mediated models than in the minimal supergravity models without knowing the messenger mass scale(s) and the multiplicities of the messenger fields.

However, in the gauge-mediated case one quite generally expects that the numerical values of the ratios K_3/K_2 , K_3/K_1 and K_2/K_1 should be even larger than in Eq. (491). There are two reasons for this.

1. First, the running squark squared masses start off larger than slepton squared masses already at the input scale in gauge-mediated models, rather than having a common value m_0^2 .

2. Furthermore, in the gauge-mediated case, the input scale Q_0 is typically much lower than M_P or M_U , so that the RG evolution gives relatively more weight to RG scales closer to the electroweak scale, where the hierarchies $g_3 > g_2 > g_1$ and $M_3 > M_2 > M_1$ are already in effect.

In general, one therefore expects that the squarks should be considerably heavier than the sleptons, with the effect being more pronounced in gauge-mediated supersymmetry breaking models than in minimal supergravity models.

For any specific choice of model, this effect can be easily quantified with a numerical RG computation.

The hierarchy $m_{\text{squark}} > m_{\text{slepton}}$ tends to hold fairly generally because the RG contributions to squark masses from the gluino are always present and usually quite large, since QCD has a larger gauge coupling than the electroweak interactions.

Regardless of the type of model, there is also a “hyperfine” splitting in the squark and slepton mass spectrum produced by electroweak symmetry breaking.

Each squark and slepton ϕ will get a contribution Δ_ϕ to its squared mass, coming from the $SU(2)_L$ and $U(1)_Y$ D -term quartic interactions [see the last term in Eq. (352)] of the form (squark) 2 (Higgs) 2 and (slepton) 2 (Higgs) 2 , when the neutral Higgs scalars H_u^0 and H_d^0 get VEVs. They are model-independent for a given value of $\tan\beta$:

$$\Delta_\phi = (T_{3\phi}g^2 - Y_\phi g'^2)(v_d^2 - v_u^2) = (T_{3\phi} - Q_\phi \sin^2 \theta_W) \cos(2\beta) m_Z^2, \quad (492)$$

where $T_{3\phi}$, Y_ϕ , and Q_ϕ are the third component of weak isospin, the weak hypercharge, and the electric charge of the left-handed chiral supermultiplet to which ϕ belongs.

For example,

$$\Delta_{\tilde{u}_L} = \left(\frac{1}{2} - \frac{2}{3} \sin^2 \theta_W\right) \cos(2\beta) m_Z^2$$

$$\begin{aligned}
\Delta_{\tilde{d}_L} &= \left(-\frac{1}{2} + \frac{1}{3} \sin^2 \theta_W\right) \cos(2\beta) m_Z^2 \\
\Delta_{\tilde{u}_R} &= \left(\frac{2}{3} \sin^2 \theta_W\right) \cos(2\beta) m_Z^2.
\end{aligned} \tag{493}$$

These D -term contributions are typically smaller than the m_0^2 and K_1, K_2, K_3 contributions, but should not be neglected. They split apart the components of the $SU(2)_L$ -doublet sleptons and squarks.

Including them, the first-family squark and slepton masses are now given by:

$$m_{\tilde{d}_L}^2 = m_0^2 + K_3 + K_2 + \frac{1}{36}K_1 + \Delta_{\tilde{d}_L}, \tag{494}$$

$$m_{\tilde{u}_L}^2 = m_0^2 + K_3 + K_2 + \frac{1}{36}K_1 + \Delta_{\tilde{u}_L}, \tag{495}$$

$$m_{\tilde{u}_R}^2 = m_0^2 + K_3 + \frac{4}{9}K_1 + \Delta_{\tilde{u}_R}, \tag{496}$$

$$m_{\tilde{d}_R}^2 = m_0^2 + K_3 + \frac{1}{9}K_1 + \Delta_{\tilde{d}_R}, \tag{497}$$

$$m_{\tilde{e}_L}^2 = m_0^2 + K_2 + \frac{1}{4}K_1 + \Delta_{\tilde{e}_L}, \quad (498)$$

$$m_{\tilde{\nu}}^2 = m_0^2 + K_2 + \frac{1}{4}K_1 + \Delta_{\tilde{\nu}}, \quad (499)$$

$$m_{\tilde{e}_R}^2 = m_0^2 + K_1 + \Delta_{\tilde{e}_R}, \quad (500)$$

with identical formulas for the second-family squarks and sleptons.

The mass splittings for the left-handed squarks and sleptons are governed by model-independent sum rules

$$m_{\tilde{e}_L}^2 - m_{\tilde{\nu}_e}^2 = m_{\tilde{d}_L}^2 - m_{\tilde{u}_L}^2 = g^2(v_u^2 - v_d^2)/2 = -\cos(2\beta) m_W^2. \quad (501)$$

In the allowed range $\tan\beta > 1$, it follows that $m_{\tilde{e}_L} > m_{\tilde{\nu}_e}$ and $m_{\tilde{d}_L} > m_{\tilde{u}_L}$, with the magnitude of the splittings constrained by electroweak symmetry breaking.

Let us next consider the masses of the top squarks, for which there are several non-negligible contributions.

1. First, there are squared-mass terms for $\tilde{t}_L^* \tilde{t}_L$ and $\tilde{t}_R^* \tilde{t}_R$ that are just equal to $m_{Q_3}^2 + \Delta_{\tilde{u}_L}$ and $m_{\tilde{u}_3}^2 + \Delta_{\tilde{u}_R}$, respectively, just as for the first- and second-family squarks.
2. Second, there are contributions equal to m_t^2 for each of $\tilde{t}_L^* \tilde{t}_L$ and $\tilde{t}_R^* \tilde{t}_R$. These come from Yukawa-squared contributions contained in the F -squared term, $W^i W_i^*$.

The resulting contributing to the scalar potential is of the form $y_t^2 H_u^{0*} H_u^0 \tilde{t}_L^* \tilde{t}_L$ and $y_t^2 H_u^{0*} H_u^0 \tilde{t}_R^* \tilde{t}_R$ (see Figures 10b and 10c), with the Higgs fields replaced by their VEVs.

(Of course, similar contributions are present for all of the squarks and sleptons, but they are too small to worry about except in the case of the top squarks.)

3. Third, there are contributions to the scalar potential from from the F -squared term, $W^i W_i^*$, of form $M_{in}^* y^{jkn} \phi^{*i} \phi_j \phi_k + \text{c.c.}$

In the MSSM, the only generic “ M^{ij} ” type term in the superpotential is $\mu H_u H_d = \mu(H_u^+ H_d^- - H_u^0 H_d^0)$. So, the relevant term for top

squark mixing in $M_{in}^* y^{jkn} \phi^{*i} \phi_j \phi_k$ is that with $i = H_d^0$, $n = H_u^0$, $j = \bar{t}$, $k = Q_3$ which yields the form $-\mu^* v y_t \cos \beta \tilde{t}_R^* \tilde{t}_L + \text{c.c.}$ when H_d^0 is replaced by its VEV.

4. Finally, there are contributions to the scalar potential from the soft (scalar)³ couplings $a_t \tilde{t} \tilde{Q}_3 H_u^0 + \text{c.c.}$ [see the first term of the second line of Eq. (379)], which become $a_t v \sin \beta \tilde{t}_L \tilde{t}_R^* + \text{c.c.}$ when H_u^0 is replaced by its VEV.

Putting these all together, we have a squared-mass matrix for the top squarks, which in the gauge-eigenstate basis $(\tilde{t}_L, \tilde{t}_R)$ is given by

$$\mathcal{L}_{\text{stop masses}} = - \begin{pmatrix} \tilde{t}_L^* & \tilde{t}_R^* \end{pmatrix} \mathbf{m}_{\tilde{t}}^2 \begin{pmatrix} \tilde{t}_L \\ \tilde{t}_R \end{pmatrix} \quad (502)$$

where

$$\mathbf{m}_{\tilde{t}}^2 = \begin{pmatrix} m_{Q_3}^2 + m_t^2 + \Delta_{\tilde{u}_L} & v(a_t^* \sin \beta - \mu y_t \cos \beta) \\ v(a_t \sin \beta - \mu^* y_t \cos \beta) & m_{u_3}^2 + m_t^2 + \Delta_{\tilde{u}_R} \end{pmatrix}. \quad (503)$$

This hermitian matrix can be diagonalized by a unitary matrix to give mass

eigenstates:

$$\begin{pmatrix} \tilde{t}_1 \\ \tilde{t}_2 \end{pmatrix} = \begin{pmatrix} c_{\tilde{t}} & -s_{\tilde{t}}^* \\ s_{\tilde{t}} & c_{\tilde{t}} \end{pmatrix} \begin{pmatrix} \tilde{t}_L \\ \tilde{t}_R \end{pmatrix}. \quad (504)$$

Here $m_{\tilde{t}_1}^2 < m_{\tilde{t}_2}^2$ are the eigenvalues of Eq. (503), and $|c_{\tilde{t}}|^2 + |s_{\tilde{t}}|^2 = 1$. If the off-diagonal elements of Eq. (503) are real, then $c_{\tilde{t}}$ and $s_{\tilde{t}}$ are the cosine and sine of a stop mixing angle $\theta_{\tilde{t}}$, which can be chosen in the range $0 \leq \theta_{\tilde{t}} < \pi$.

Because of the large RG effects, at the electroweak scale one finds that $m_{\tilde{u}_3}^2 < m_{\tilde{Q}_3}^2$, and both of these quantities are usually significantly smaller than the squark squared masses for the first two families.

The diagonal terms $m_{\tilde{t}}^2$ in Eq. (503) tend to mitigate this effect somewhat, but the off-diagonal entries will typically induce a significant mixing, which always reduces the lighter top-squark squared-mass eigenvalue.

Therefore, models often predict that \tilde{t}_1 is the lightest squark of all, and that it is predominantly \tilde{t}_R .

A very similar analysis can be performed for the bottom squarks and charged tau sleptons, which in their respective gauge-eigenstate bases $(\tilde{b}_L, \tilde{b}_R)$ and $(\tilde{\tau}_L, \tilde{\tau}_R)$ have squared-mass matrices:

$$\mathbf{m}_{\tilde{b}}^2 = \begin{pmatrix} m_{Q_3}^2 + \Delta_{\tilde{d}_L} & v(a_b^* \cos \beta - \mu y_b \sin \beta) \\ v(a_b \cos \beta - \mu^* y_b \sin \beta) & m_{\tilde{d}_3}^2 + \Delta_{\tilde{d}_R} \end{pmatrix}, \quad (505)$$

$$\mathbf{m}_{\tilde{\tau}}^2 = \begin{pmatrix} m_{L_3}^2 + \Delta_{\tilde{e}_L} & v(a_\tau^* \cos \beta - \mu y_\tau \sin \beta) \\ v(a_\tau \cos \beta - \mu^* y_\tau \sin \beta) & m_{\tilde{e}_3}^2 + \Delta_{\tilde{e}_R} \end{pmatrix}. \quad (506)$$

These can be diagonalized to give mass eigenstates \tilde{b}_1, \tilde{b}_2 and $\tilde{\tau}_1, \tilde{\tau}_2$ in exact analogy with Eq. (504).

The magnitude and importance of mixing in the sbottom and stau sectors depends on how big $\tan \beta$ is.

If $\tan \beta$ is not too large (in practice, this usually means less than about 10 or so, depending on the situation under study), the sbottoms and staus do not get a very large effect from the mixing terms and the RG effects.

In that case the mass eigenstates are very nearly the same as the gauge eigenstates $\tilde{b}_L, \tilde{b}_R, \tilde{\tau}_L$ and $\tilde{\tau}_R$. The latter three, and $\tilde{\nu}_\tau$, will be nearly degenerate with their first- and second-family counterparts with the same $SU(3)_C \times SU(2)_L \times U(1)_Y$ quantum numbers.

However, even in the case of small $\tan\beta$, \tilde{b}_L will feel the effects of the large top Yukawa coupling because it is part of the doublet containing \tilde{t}_L . In particular, top loop contributions in the RG equations act to decrease $m_{Q_3}^2$ as it is RG-evolved down from the input scale to the electroweak scale.

Therefore the mass of \tilde{b}_L can be significantly less than the masses of \tilde{d}_L and \tilde{s}_L .

For larger values of $\tan\beta$, the mixing in Eqs. (505) and (506) can be quite significant, because y_b, y_τ and a_b, a_τ are non-negligible.

Just as in the case of the top squarks, the lighter sbottom and stau mass eigenstates (denoted \tilde{b}_1 and $\tilde{\tau}_1$) can be significantly lighter than

their first- and second-family counterparts.

Furthermore, $\tilde{\nu}_\tau$ can be significantly lighter than the nearly degenerate $\tilde{\nu}_e, \tilde{\nu}_\mu$.

The requirement that the third-family squarks and sleptons should all have positive squared masses implies limits on the magnitudes of $a_t^* \sin \beta - \mu y_t \cos \beta$ and $a_b^* \cos \beta - \mu y_b \sin \beta$ and $a_\tau^* \cos \beta - \mu y_\tau \sin \beta$.

If they are too large, then the smaller eigenvalue of Eq. (503), (505) or (506) will be driven negative, implying that a squark or charged slepton gets a VEV, breaking $SU(3)_C$ or electromagnetism.

Since this is clearly unacceptable, one can put bounds on the (scalar)³ couplings, or equivalently on the parameter A_0 in minimal supergravity models.

Even if all of the squared-mass eigenvalues are positive, the presence of large (scalar)³ couplings can yield global minima of the scalar

potential, with non-zero squark and/or charged slepton VEVs, which are disconnected from the vacuum that conserves $SU(3)_C$ and electromagnetism.

However, it is not always immediately clear whether the mere existence of such disconnected global minima should really disqualify a set of model parameters, because the tunneling rate from our “good” vacuum to the “bad” vacua can easily be longer than the age of the universe.

- Summary: the MSSM sparticle spectrum

In the MSSM there are 32 distinct masses corresponding to undiscovered particles, not including the gravitino. In this section we have explained how the masses and mixing angles for these particles can be computed, given an underlying model for the soft terms at some input scale.

Assuming only that the mixing of first- and second-family squarks

and sleptons is negligible, the mass eigenstates of the MSSM are listed in Table 5.

Names	Spin	P_R	Gauge Eigenstates	Mass Eigenstates
Higgs bosons	0	+1	$H_u^0 \ H_d^0 \ H_u^+ \ H_d^-$	$h^0 \ H^0 \ A^0 \ H^\pm$
squarks	0	-1	$\tilde{u}_L \ \tilde{u}_R \ \tilde{d}_L \ \tilde{d}_R$	(same)
			$\tilde{s}_L \ \tilde{s}_R \ \tilde{c}_L \ \tilde{c}_R$	(same)
			$\tilde{t}_L \ \tilde{t}_R \ \tilde{b}_L \ \tilde{b}_R$	$\tilde{t}_1 \ \tilde{t}_2 \ \tilde{b}_1 \ \tilde{b}_2$
sleptons	0	-1	$\tilde{e}_L \ \tilde{e}_R \ \tilde{\nu}_e$	(same)
			$\tilde{\mu}_L \ \tilde{\mu}_R \ \tilde{\nu}_\mu$	(same)
			$\tilde{\tau}_L \ \tilde{\tau}_R \ \tilde{\nu}_\tau$	$\tilde{\tau}_1 \ \tilde{\tau}_2 \ \tilde{\nu}_\tau$
neutralinos	1/2	-1	$\tilde{B}^0 \ \tilde{W}^0 \ \tilde{H}_u^0 \ \tilde{H}_d^0$	$\tilde{N}_1 \ \tilde{N}_2 \ \tilde{N}_3 \ \tilde{N}_4$
charginos	1/2	-1	$\tilde{W}^\pm \ \tilde{H}_u^\pm \ \tilde{H}_d^\pm$	$\tilde{C}_1^\pm \ \tilde{C}_2^\pm$
gluino	1/2	-1	\tilde{g}	(same)
goldstino (gravitino)	1/2 (3/2)	-1	\tilde{G}	(same)

Table 5: The undiscovered particles in the Minimal Supersymmetric Standard Model (with sfermion mixing for the first two families assumed to be negligible).

A complete set of Feynman rules for the interactions of these particles with each other and with the Standard Model quarks, leptons, and gauge bosons can be found in Haber-Kane and Gunion-Haber.

Specific models for the soft terms typically predict the masses and the mixing angles for the MSSM in terms of far fewer parameters.

For example, in the minimal supergravity models, the only free parameters not already measured by experiment are m_0^2 , $m_{1/2}$, A_0 , μ , and b .

In gauge-mediated supersymmetry breaking models, the free parameters include at least the scale Λ , the typical messenger mass scale M_{mess} , the integer number N_5 of copies of the minimal messengers, the goldstino decay constant $\langle F \rangle$, and the Higgs mass parameters μ and b .

After RG evolving the soft terms down to the electroweak scale, one can demand that the scalar potential gives correct electroweak

symmetry breaking. This allows us to trade $|\mu|$ and b (or B_0) for one parameter $\tan \beta$ given the known value of m_Z , as in Eqs. (527)-(526).

So, to a reasonable approximation, the entire mass spectrum in minimal supergravity models is determined by only five unknown parameters: m_0^2 , $m_{1/2}$, A_0 , $\tan \beta$, and $\text{Arg}(\mu)$, while in the simplest gauge-mediated supersymmetry breaking models one can pick parameters Λ , M_{mess} , N_5 , $\langle F \rangle$, $\tan \beta$, and $\text{Arg}(\mu)$.

Both frameworks are highly predictive. Of course, it is easy to imagine that the essential physics of supersymmetry breaking is not captured by either of these two scenarios in their minimal forms. For example, the anomaly mediated contributions could play a role, perhaps in concert with the gauge-mediation or Planck-scale mediation mechanisms.

Figure 23 shows the RG running of scalar and gaugino masses in a typical model based on the minimal supergravity boundary conditions

imposed at $Q_0 = 2.5 \times 10^{16}$ GeV.

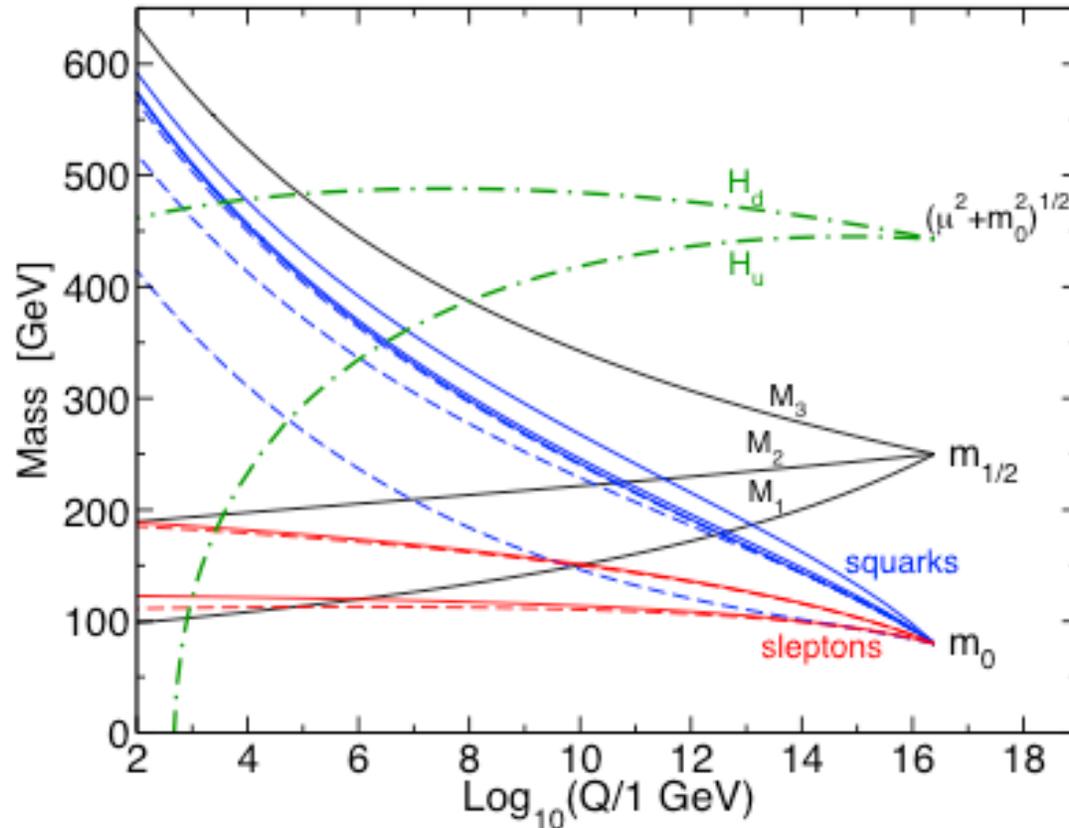


Figure 23: RG evolution of scalar and gaugino mass parameters in the MSSM with typical minimal supergravity-inspired boundary conditions imposed at $Q_0 = 2.5 \times 10^{16}$ GeV. The parameter $\mu^2 + m_{H_u}^2$ runs negative, provoking electroweak symmetry breaking.

[The parameter values used for this illustration were $m_0 = 80$ GeV, $m_{1/2} = 250$

GeV, $A_0 = -500$ GeV, $\tan \beta = 10$, and $\text{sign}(\mu) = +$.]

- The running gaugino masses are solid lines labeled by M_1 , M_2 , and M_3 .
- The dot-dashed lines labeled H_u and H_d are the running values of the quantities $(\mu^2 + m_{H_u}^2)^{1/2}$ and $(\mu^2 + m_{H_d}^2)^{1/2}$, which appear in the Higgs potential.
- The other lines are the running squark and slepton masses, with dashed lines for the square roots of the third family parameters $m_{d_3}^2$, $m_{Q_3}^2$, $m_{u_3}^2$, $m_{L_3}^2$, and $m_{e_3}^2$ (from top to bottom), and solid lines for the first and second family sfermions.

Note that $\mu^2 + m_{H_u}^2$ runs negative because of the effects of the large top Yukawa coupling as discussed above, providing for electroweak symmetry breaking.

At the electroweak scale, the values of the Lagrangian soft parameters can be used to extract the physical masses, cross-sections, and decay

widths of the particles, and other observables such as dark matter abundances and rare process rates. There are a variety of publicly available programs that do these tasks, including radiative corrections; see for example ISAJET, micrOMEGAs, FeynHiggs.

Figures 24—26 show deliberately qualitative sketches of sample MSSM mass spectrum obtained from three different types of model assumptions. These spectra are presented for entertainment purposes only! No warranty, expressed or implied, guarantees that they look anything like the real world.

1. Minimal Supergravity Inspired

The first is the output from a minimal supergravity-inspired model with relatively low m_0^2 compared to $m_{1/2}^2$ (in fact the same model parameters as used for Fig. 23).

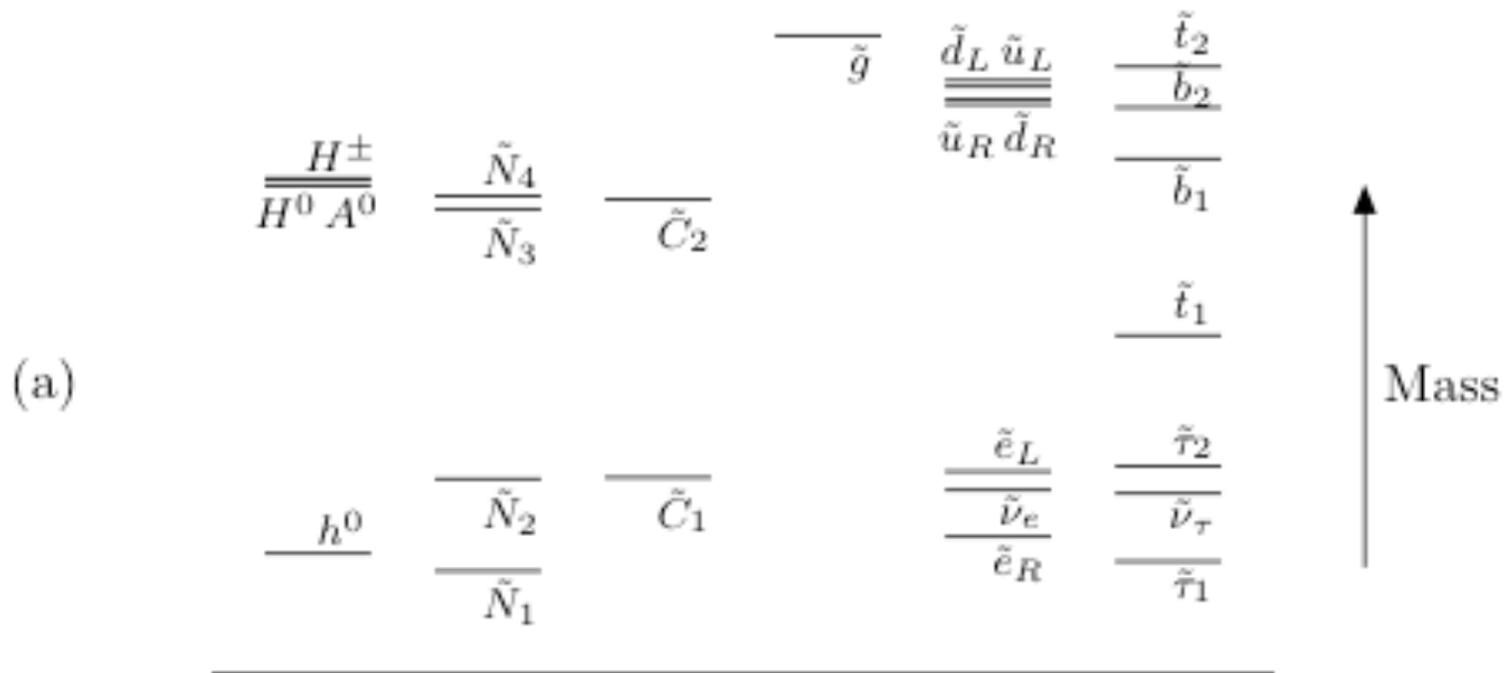


Figure 24: Mass spectra for the undiscovered particles in the MSSM, for minimal supergravity with $m_0^2 \ll m_{1/2}^2$.

This model features a near-decoupling limit for the Higgs sector, and a bino-like \tilde{N}_1 LSP, nearly degenerate wino-like \tilde{N}_2, \tilde{C}_1 , and higgsino-like $\tilde{N}_3, \tilde{N}_4, \tilde{C}_2$.

The gluino is the heaviest superpartner.

The squarks are all much heavier than the sleptons, and the lightest

sfermion is a stau.

Variations in the model parameters have important and predictable effects.

For example, taking larger m_0^2 in minimal supergravity models will tend to squeeze together the spectrum of squarks and sleptons and move them all higher compared to the neutralinos, charginos and gluino.

Taking larger values of $\tan\beta$ with other model parameters held fixed will usually tend to lower \tilde{b}_1 and $\tilde{\tau}_1$ masses compared to those of the other sparticles.

2. Gauge Mediated Supersymmetry Breaking Model

The second sample sketch in Fig. 25 is obtained from a typical minimal GMSB model, [with $N_5 = 1$, $\Lambda = 150$ TeV, $\tan\beta = 15$, and $\text{sign}(\mu) = +$ at a scale $Q_0 = M_{\text{mess}} = 300$ TeV for the illustration].

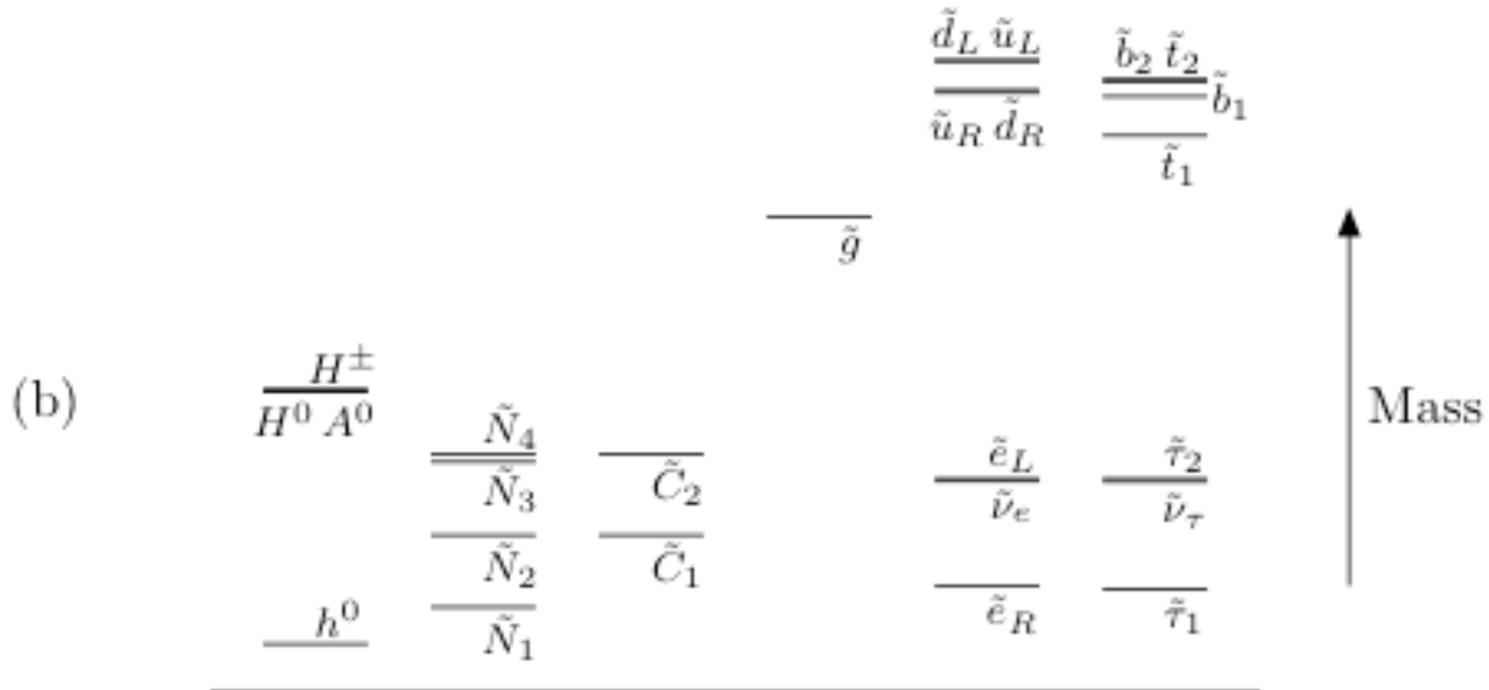


Figure 25: Mass spectra for the undiscovered particles in the MSSM for minimal GMSB with $N_5 = 1$.

Here we see that the hierarchy between strongly interacting sparticles and weakly interacting ones is quite large. Changing the messenger scale or Λ does not reduce the relative splitting between squark and slepton masses, because there is no analog of the universal m_0^2 contribution here.

Increasing the number of messenger fields tends to decrease the squark and slepton masses relative to the gaugino masses, but still keeps the hierarchy between squark and slepton masses intact.

In the model shown, the NLSP is a bino-like neutralino, but for larger number of messenger fields it could be either a stau, or else co-NLSPs $\tilde{\tau}_1$, \tilde{e}_L , $\tilde{\mu}_L$, depending on the choice of $\tan\beta$.

3. Anomaly-Mediated Supersymmetry Breaking Model

The third sample sketch in Fig. 26 is obtained from an AMSB model with an additional universal scalar mass $m_0 = 450$ GeV? added at $Q_0 = 2 \times 10^{16}$ GeV to rescue the sleptons, and with $m_{3/2} = 60$ TeV, $\tan\beta = 10$, and $\text{sign}(\mu) = +$ for the illustration.

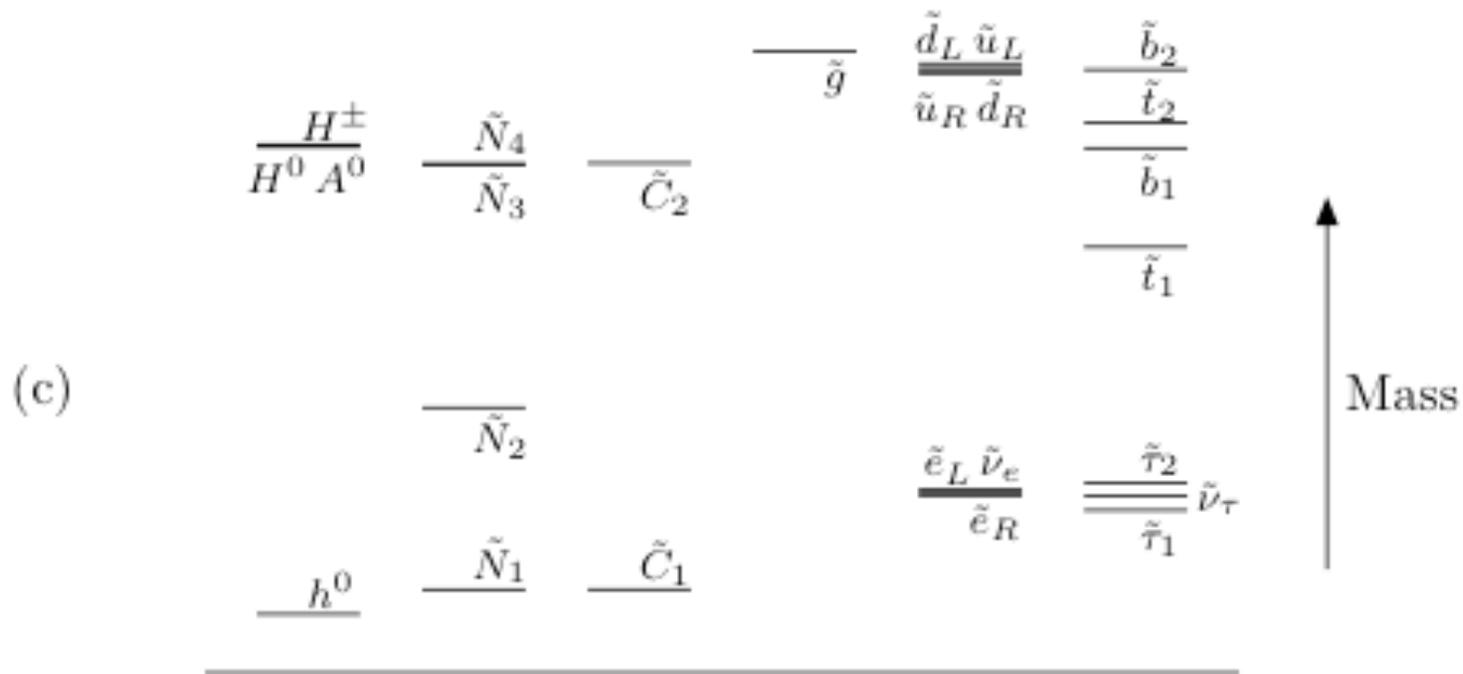


Figure 26: Mass spectra for the undiscovered particles in the MSSM for AMSB with an extra m_0^2 for scalars.

Here the most striking feature is that the LSP is a wino-like neutralino, with $m_{\tilde{C}_1} - m_{\tilde{N}_1}$ only about 160 MeV.

It would be a mistake to rely too heavily on specific scenarios for the MSSM mass and mixing spectrum, and the above illustrations are only a tiny fraction of the available possibilities.

General Lessons

It is useful to keep in mind some general lessons that often recur in various different models.

Indeed, there has emerged a sort of folklore concerning likely features of the MSSM spectrum, partly based on theoretical bias and partly on the constraints inherent in most known viable softly-broken supersymmetric theories.

We remark on these features mainly because they represent the prevailing prejudices among supersymmetry theorists, which is certainly a useful thing to know even if one wisely decides to remain skeptical. For example, it is perhaps not unlikely that:

- The LSP is the lightest neutralino \tilde{N}_1 , unless the gravitino is lighter or R -parity is not conserved.

If $M_1 < M_2, |\mu|$, then \tilde{N}_1 is likely to be bino-like, with a mass roughly 0.5 times the masses of \tilde{N}_2 and \tilde{C}_1 in many well-motivated

models.

If, instead, $|\mu| < M_1, M_2$, then the LSP \tilde{N}_1 has a large higgsino content and \tilde{N}_2 and \tilde{C}_1 are not much heavier.

And, if $M_2 \ll M_1, |\mu|$, then the LSP will be a wino-like neutralino, with a chargino only very slightly heavier.

- The gluino will be much heavier than the lighter neutralinos and charginos. This is certainly true in the case of the “standard” gaugino mass relation Eq. (388); more generally, the running gluino mass parameter grows relatively quickly as it is RG-evolved into the infrared because the QCD coupling is larger than the electroweak gauge couplings.

So even if there are big corrections to the gaugino mass boundary conditions, *e.g.* Eq. (389), the gluino mass parameter M_3 is likely to come out larger than M_1 and M_2 .

- The squarks of the first and second families are nearly degenerate and much heavier than the sleptons. This is because each squark

mass gets the same large positive-definite radiative corrections from loops involving the gluino.

The left-handed squarks $\tilde{u}_L, \tilde{d}_L, \tilde{s}_L$ and \tilde{c}_L are likely to be heavier than their right-handed counterparts $\tilde{u}_R, \tilde{d}_R, \tilde{s}_R$ and \tilde{c}_R , because of the effect parameterized by K_2 in Eqs. (494)-(500).

- The squarks of the first two families cannot be lighter than about 0.8 times the mass of the gluino in minimal supergravity models, and about 0.6 times the mass of the gluino in the simplest gauge-mediated models if the number of messenger squark pairs is $N_5 \leq 4$. In the minimal supergravity case this is because the gluino mass feeds into the squark masses through RG evolution; in the gauge-mediated case it is because the gluino and squark masses are tied together.
- The lighter stop \tilde{t}_1 and the lighter sbottom \tilde{b}_1 are probably the lightest squarks.

This is because stop and sbottom mixing effects and the effects of

y_t and y_b in the RG equations both tend to decrease the lighter stop and sbottom masses.

- The lightest charged slepton is probably a stau $\tilde{\tau}_1$.
The mass difference $m_{\tilde{e}_R} - m_{\tilde{\tau}_1}$ is likely to be significant if $\tan\beta$ is large, because of the effects of a large tau Yukawa coupling.
For smaller $\tan\beta$, $\tilde{\tau}_1$ is predominantly $\tilde{\tau}_R$ and it is not so much lighter than $\tilde{e}_R, \tilde{\mu}_R$.
- The left-handed charged sleptons \tilde{e}_L and $\tilde{\mu}_L$ are likely to be heavier than their right-handed counterparts \tilde{e}_R and $\tilde{\mu}_R$.
This is because of the effect of K_2 in Eq. (498). (Note also that $\Delta_{\tilde{e}_L} - \Delta_{\tilde{e}_R}$ is positive but very small because of the numerical accident $\sin^2\theta_W \approx 1/4$.)
- The lightest neutral Higgs boson h^0 should be lighter than about 150 GeV, and may be much lighter than the other Higgs scalar mass eigenstates A^0, H^\pm, H^0 .

The most important point is that by measuring the masses and mixing

angles of the MSSM particles we will be able to gain a great deal of information that can rule out or bolster evidence for competing proposals for the origin and mediation of supersymmetry breaking.

Dark matter and its detection

As we have reviewed, evidence from experimental cosmology has now solidified to the point that, with some plausible assumptions, the cold dark matter density is known to be

$$\Omega_{\text{DM}}^0 h_0^2 \approx 0.11. \quad (507)$$

with statistical errors of order 5%, and systematic errors that are less clear, where Ω_{DM} is the average energy density in non-baryonic dark matter divided by the total critical density that would lead to a spatially flat homogeneous universe. For $h_0^2 \approx 0.5$ (with an error of order 10%), this translates into a cold dark matter density

$$\rho_{\text{DM}} \approx 1.2 \times 10^{-6} \text{ GeV}/\text{cm}^3, \quad (508)$$

averaged over very large distance scales.

One of the nice features of supersymmetry with exact R -parity conservation is that a stable electrically neutral LSP might be this cold dark matter. There are three obvious candidates: the lightest sneutrino, the gravitino, and the lightest neutralino.

1. The possibility of a sneutrino LSP making up the dark matter with a cosmologically interesting density has been largely ruled out by direct searches.
2. If the gravitino is the LSP, as in many gauge-mediated supersymmetry breaking models, then gravitinos from reheating after inflation or from other sparticle decays might be the dark matter, but they would be impossible to detect directly even if they have the right cosmological density today.

They interact too weakly.

3. The most attractive prospects for direct detection of supersymmetric

dark matter, therefore, are based on the idea that the lightest neutralino \tilde{N}_1 is the LSP.

Dark Matter

In the early universe, sparticles existed in thermal equilibrium with the ordinary Standard Model particles. As the universe cooled and expanded, the heavier sparticles could no longer be produced, and they eventually annihilated or decayed into neutralino LSPs. Some of the LSPs pair-annihilated into final states not containing sparticles.

If there are other sparticles that are only slightly heavier, then they existed in thermal equilibrium in comparable numbers to the LSP, and their co-annihilations are also important in determining the resulting dark matter density.

Eventually, as the density decreased, the annihilation rate became small compared to the cosmological expansion, and the \tilde{N}_1 experienced “freeze out”, with a density today determined by this small rate and the subsequent dilution due to the expansion of the universe.

In order to get the observed dark matter density today, the thermal-averaged effective annihilation cross-section times the relative speed of the LSPs should be about

$$\langle \sigma |v| \rangle \sim 1 \text{ pb} \sim \alpha^2 / (150 \text{ GeV})^2. \quad (509)$$

Let us check this out in detail. Recall Eq. (242), repeated below:

$$\begin{aligned} \rho_\chi^0 &= n_\chi^0 m_\chi = 1.101 \times 10^4 \frac{(n+1)x_f^{n+1}}{(g_* s / g_*^{1/2}) M_{\text{P}} \sigma_0} \\ \Omega_\chi^0 h_0^2 &= \left(\frac{\rho_\chi^0}{\rho_c^0} \right) h_0^2 = 1.042 \times 10^9 \frac{(n+1)x_f^{n+1} \text{ GeV}^{-1}}{(g_* s / g_*^{1/2}) M_{\text{P}} \sigma_0}. \end{aligned} \quad (510)$$

In the above, σ_0 was defined by

$$\langle \sigma_\chi |v| \rangle \equiv \sigma_0 (T/m_\chi)^n = \sigma_0 x^{-n}, \quad \text{for } x \gtrsim 3. \quad (511)$$

If we assume that the cross section is S-wave dominated ($n = 0$), as is commonly the case, and use $g_{*S} = g_*$ (as is inevitably the case if the freezeout temperature is above a GeV or so), then the above reduces to

$$\Omega_{\chi}^0 h_0^2 = 1.042 \times 10^9 \frac{x_f \text{ GeV}^{-1}}{g_*^{1/2} M_{\text{P}} \sigma_0}. \quad (512)$$

If there is a mixture of S-wave and P-wave annihilation so that

$$\langle \sigma |v| \rangle = a + b|v|^2 + \dots \quad (513)$$

then the above generalizes to

$$\Omega_{\chi}^0 h_0^2 = 1.042 \times 10^9 \frac{x_f \text{ GeV}^{-1}}{g_*^{1/2} M_{\text{P}} (a + 3b/x_f)}. \quad (514)$$

with

$$x_f \equiv \frac{m_\chi}{T_f} \approx \ln \left[c(c+2) \sqrt{\frac{45}{8} \frac{g_\chi}{2\pi^3} \frac{m_\chi M_{\text{P}} (a + 6b/x_f)}{g_*^{1/2} x_f^{1/2}}} \right], \quad (515)$$

where $c \sim 0.5$ providing a good approximation to the exact solution (depending upon precise relative weight of a and b terms).

As regards g_* , recall the result given earlier:

$$g_* = \text{neutrinos} + \text{photon} + \text{charged-leptons} + \text{gluons} + (W^\pm, Z) + \text{quarks} + \text{Higgs} \quad (516)$$

yielding

$$g_* = \frac{7}{8}(3 \times 2) + 2 + \frac{7}{8}(3 \times 2 \times 2) + 8 \times 2 + 3 \times 3 + \frac{7}{8}(3 \times 3 \times 2 \times 2 \times 2) + 1 = 106.75. \quad (517)$$

Now, which of these terms are present clearly depends on the temperature, T_f at freeze-out. For example, if $m_\chi < 100$ GeV and $x_f = m_\chi/T_f \sim 20 - 30$ (a typical value as you have learned), then $T_f < 5$ GeV. In this case, we should remove the Higgs boson, the

W^\pm , Z and the 3rd family quark terms, with the result that

$$g_* = \frac{7}{8}(3 \times 2) + 2 + \frac{7}{8}(3 \times 2 \times 2) + 8 \times 2 + \frac{7}{8}(2 \times 3 \times 2 \times 2 \times 2) = 75.75. \quad (518)$$

Of course, 5 GeV is very close to the b -quark mass and so we are on the verge of needing to include the b -quark piece of

$$g_*(b) = \frac{7}{8}(1 \times 3 \times 2 \times 1 \times 2) = \frac{21}{2} = 10.5, \quad (519)$$

which would increase g_* to $g_* = 86.25$. This value would be fairly appropriate for $m_\chi \gtrsim 150$ GeV.

Inserting $M_P = 1.22 \times 10^{19}$ GeV and setting $\Omega_\chi^0 h_0^2 = 0.11$ in Eq. (512) gives

$$\sigma_0 = 1.74 \times 10^{-9} \left(\frac{x_f}{20}\right) \left(\frac{80}{g_*}\right)^{1/2} \text{GeV}^{-2} \quad (520)$$

as the needed annihilation cross section. As stated earlier, this compares nicely in order of magnitude with the crude estimate of an electroweak

cross section (using a ~ 5 GeV energy scale value of α):

$$\begin{aligned}\sigma_{EW} &\sim \frac{\alpha^2}{m_\chi^2} = \left(\frac{1}{137}\right)^2 \left(\frac{150 \text{ GeV}}{m_\chi}\right)^2 \left(\frac{1}{150 \text{ GeV}}\right)^2 \\ &= 2.37 \times 10^{-9} \text{ GeV}^{-2} \left(\frac{150 \text{ GeV}}{m_\chi}\right)^2.\end{aligned}\quad (521)$$

To repeat, this coincidence is called the “**WIMP Miracle**”. We must now turn to whether or not it is easily realized in the context of a supersymmetric model.

A neutralino LSP naturally has, very roughly, the correct (electroweak) interaction strength and mass to give the required $\langle\sigma|v|\rangle$. More detailed and precise estimates can be obtained with publicly available computer programs such as DarkSUSY and micrOMEGAs. These allow the predictions of specific candidate models of supersymmetry breaking to be compared to Eq. (507).

Some of the diagrams that are typically important for neutralino LSP pair annihilation are shown in Fig. 27.

Depending on the mass of \tilde{N}_1 , various other processes including $\tilde{N}_1\tilde{N}_1 \rightarrow ZZ, Zh^0, h^0h^0$ or even $W^\pm H^\mp, ZA^0, h^0A^0, h^0H^0, H^0A^0, H^0H^0, A^0A^0$, or H^+H^- may also have been important.

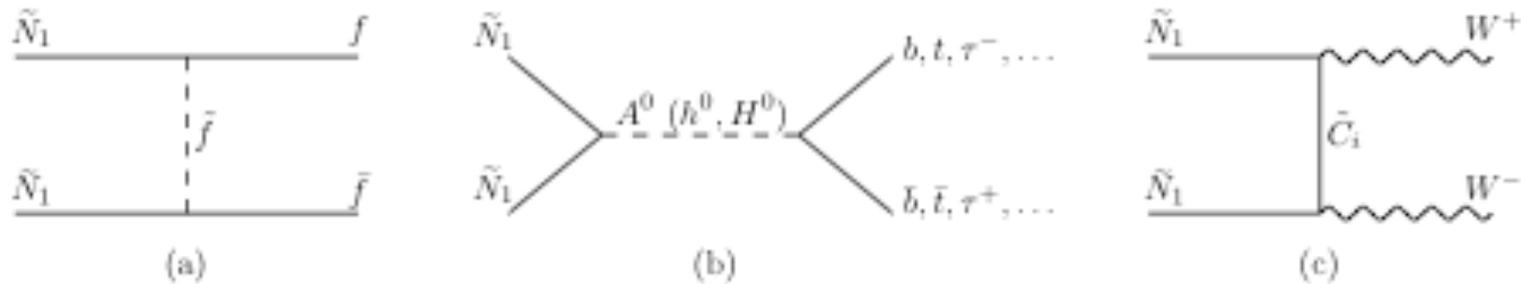
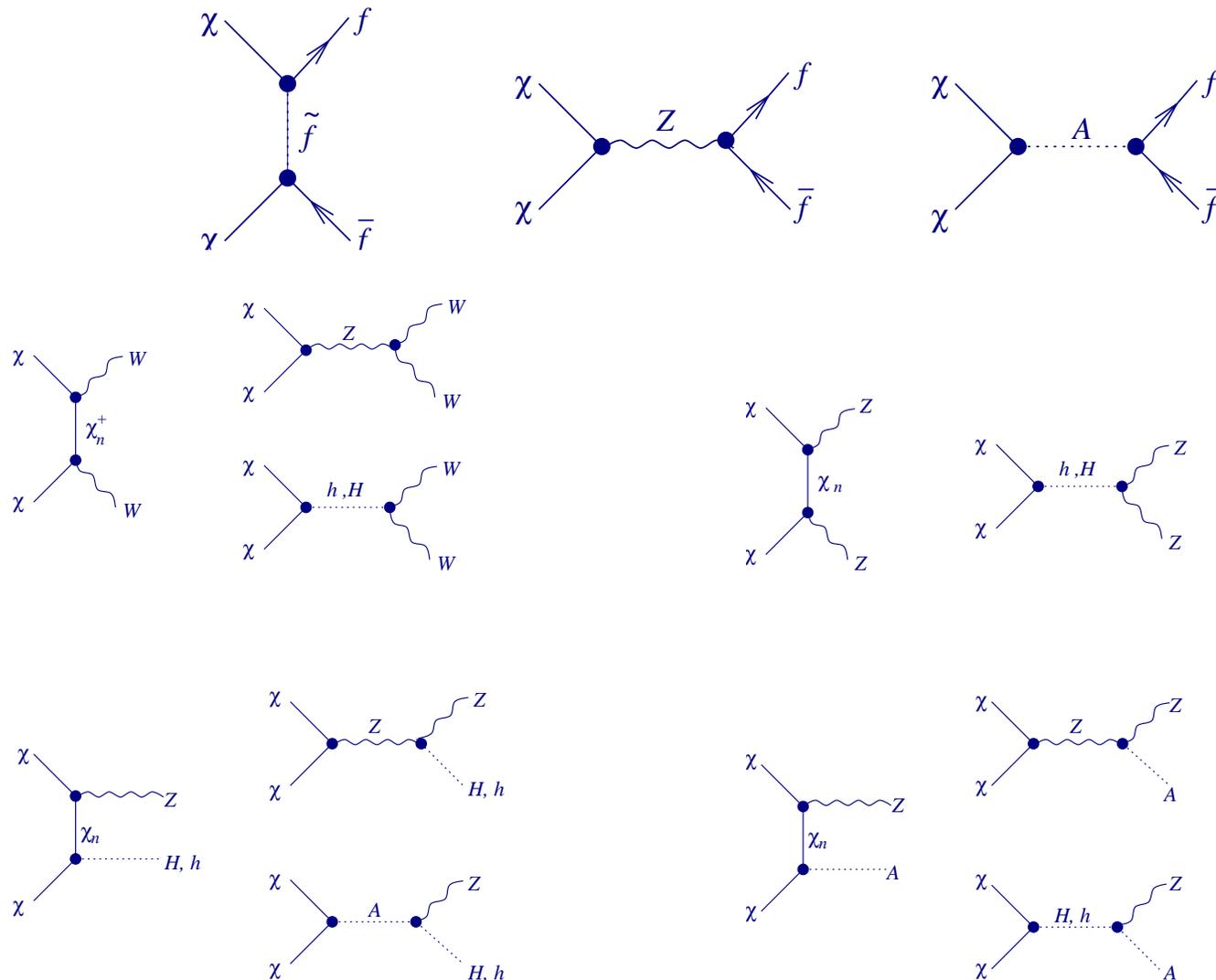
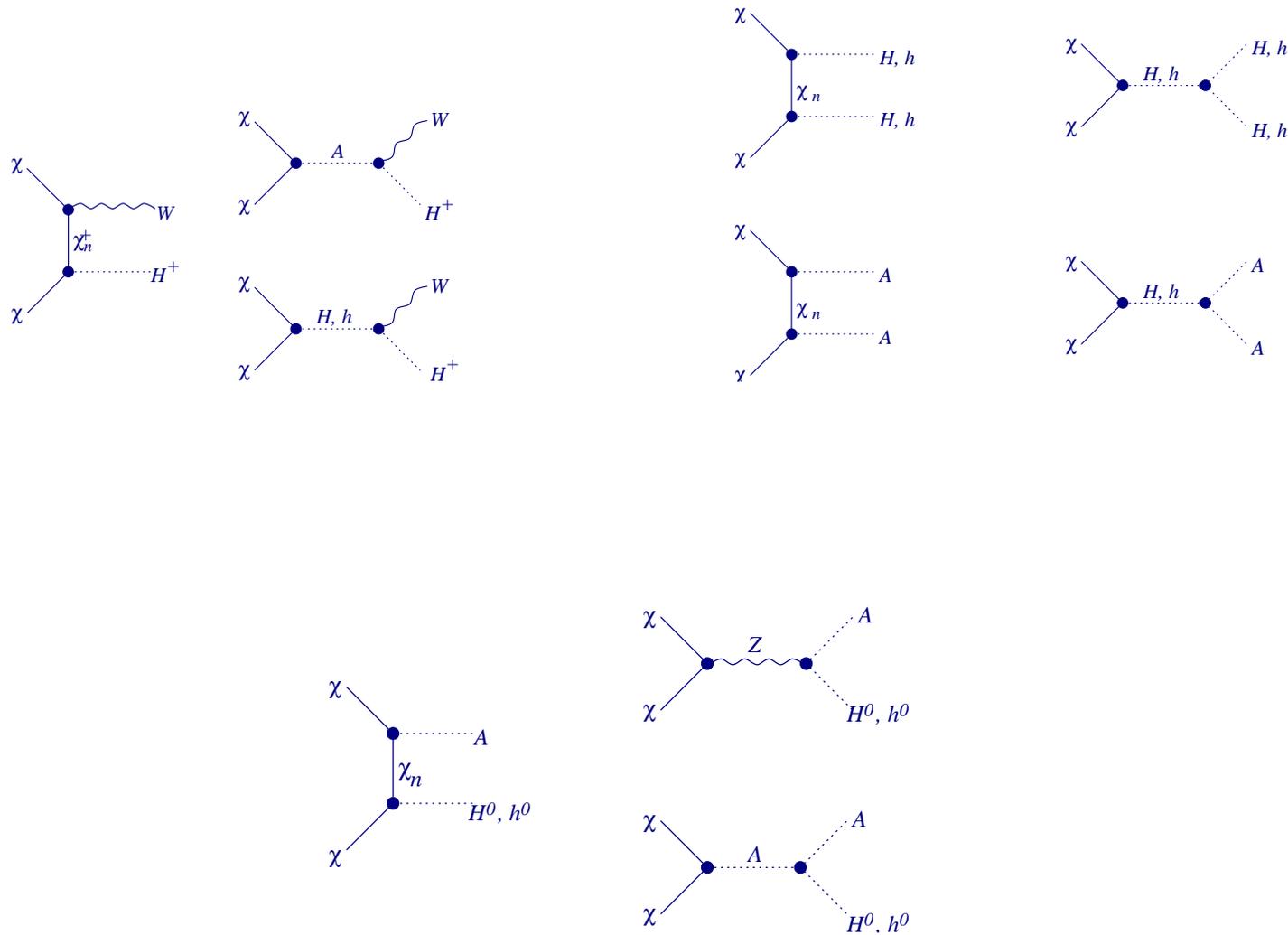


Figure 27: Contributions to the annihilation cross-section for neutralino dark matter LSPs from (a) t -channel slepton and squark exchange, (b) near-resonant annihilation through a Higgs boson (s -wave for A^0 , and p -wave for h^0, H^0), and (c) t -channel chargino exchange.

The complete set of possibly important diagrams for direct $\tilde{N}_1 = \chi$ annihilation is given below.





Some of the diagrams that can lead to co-annihilation of the LSPs with slightly heavier sparticles are shown in Figs. 28 and 29.

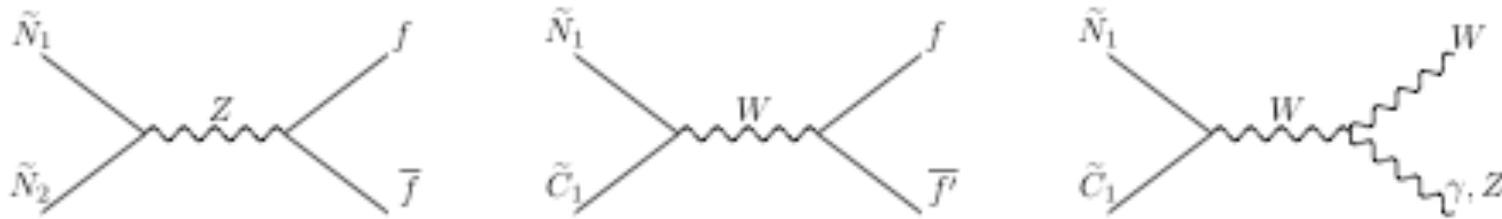


Figure 28: Some contributions to the co-annihilation of dark matter \tilde{N}_1 LSPs with slightly heavier \tilde{N}_2 and \tilde{C}_1 . All three diagrams are particularly important if the LSP is higgsino-like, and the last two diagrams are important if the LSP is wino-like.

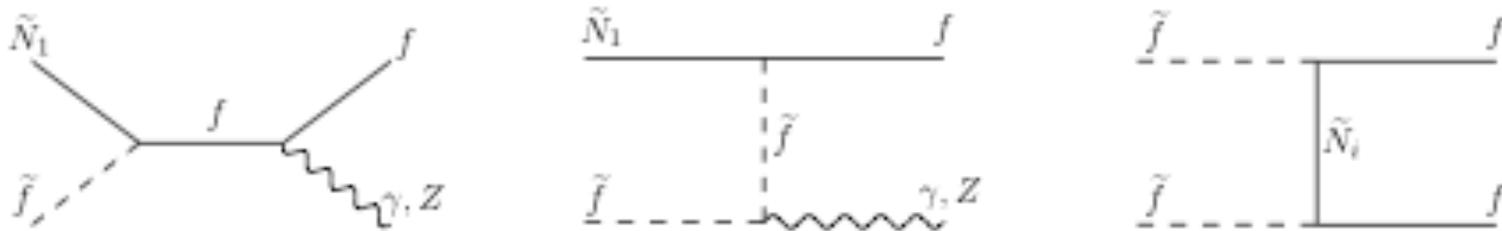


Figure 29: Some contributions to the co-annihilation of dark matter \tilde{N}_1 LSPs with slightly heavier sfermions, which in popular models are most plausibly staus (or perhaps top squarks).

Remarks

- If \tilde{N}_1 is **mostly higgsino or mostly wino**, then the annihilation diagram Fig. 27c ($\tilde{N}_1\tilde{N}_1 \rightarrow W^+W^-$ via \tilde{C}_1 exchange, involving the SUSY

couplings that are the analogues of HW^+W^- in the higgsino case or $W^0W^+W^-$ in the wino case) and the co-annihilation mechanisms provided by Fig. 28 are typically **much too efficient** (too large $\langle\sigma|v|\rangle$) to provide the full required cold dark matter density, unless the LSP is very heavy, of order 1 TeV or more.

This is often considered to be somewhat at odds with the idea that supersymmetry is the solution to the hierarchy problem.

However, for lighter higgsino-like or wino-like LSPs, non-thermal mechanisms can be invoked to provide the right dark matter abundance.

- A recurring feature of many models of supersymmetry breaking is that **the lightest neutralino is mostly bino**.

It turns out that in much of the parameter space not already ruled out by LEP with a bino-like \tilde{N}_1 , **the predicted relic density is too high**, either because the LSP couplings are too small, or the sparticles

are too heavy, or both, leading to an annihilation cross-section that is too low.

To avoid this, there must be significant additional contributions to $\langle\sigma|v|\rangle$. The possibilities can be classified qualitatively in terms of the diagrams that contribute most strongly to the annihilation.

1. First, if at least one sfermion is not too heavy, the diagram of Fig. 27a ($\tilde{N}_1\tilde{N}_1 \rightarrow f\bar{f}$ via \tilde{f} exchange) is effective in reducing the dark matter density.
 - In models with a bino-like \tilde{N}_1 , the most important such contribution usually comes from \tilde{e}_R , $\tilde{\mu}_R$, and $\tilde{\tau}_1$ slepton exchange. The region of parameter space where this works out right is often referred to by the jargon “**bulk region**”, because it corresponded to the main allowed region with dark matter density less than the critical density, before $\Omega_{\text{DM}}^0 h_0^2$ was accurately known and before the highest energy LEP searches had happened.
 - However, the diagram of Fig. 27a is subject to a p -wave

suppression, and so sleptons that are light enough to reduce the relic density sufficiently are, in many models, also light enough to be excluded by LEP, or correspond to light Higgs bosons that are excluded by LEP, or have difficulties with other indirect constraints.

- In the minimal supergravity inspired framework described earlier, the remaining viable bulk region usually has m_0 and $m_{1/2}$ less than about 100 GeV and 250 GeV respectively, depending on other parameters.

In fact, this region is now excluded by the latest CMS and ATLAS data just reported at the Aspen Winter Conference.

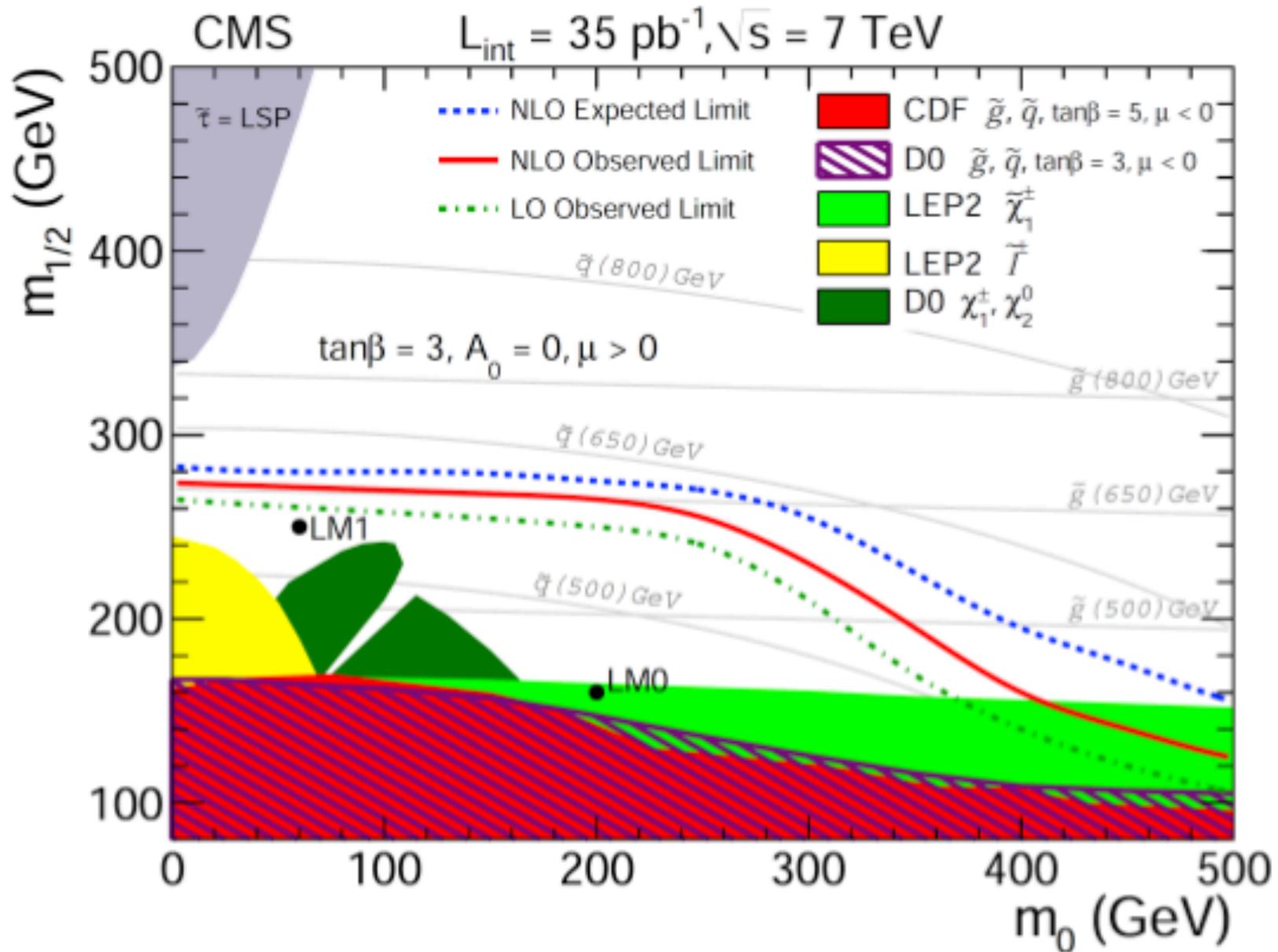


Figure 30: Latest CMS limits on the CMSSM. ATLAS limits are similar.

– If the final state of neutralino pair annihilation is instead $t\bar{t}$, then there is no p -wave suppression.

This typically requires a top squark that is less than about 150 GeV heavier than the LSP, which in turn has $m_{\tilde{N}_1}$ between about m_t and $m_t + 100$ GeV.

This situation does not occur in the minimal supergravity inspired framework, but can be natural if the ratio of gluino and wino mass parameters, M_3/M_2 , is smaller than the unification prediction of Eq. (480) by a factor of a few.

2. A second way of annihilating excess bino-like LSPs to the correct density is obtained if $2m_{\tilde{N}_1} \approx m_{A^0}$, or m_{h^0} , or m_{H^0} , as shown in fig. 27b, so that the cross-section is near a resonance pole. For example, if $2m_{\tilde{N}_1} \approx m_{A^0}$ then the cross section is proportional to

$$\frac{1}{(s - m_{A^0}^2)^2 + m_{A^0}^2 \Gamma_{A^0}^2} \quad (522)$$

which is largest when $s \sim (2m_{\tilde{N}_1})^2$ if $2m_{\tilde{N}_1} \approx m_{A^0}$.

An A^0 resonance annihilation will be s -wave, and so more efficient than a p -wave h^0 or H^0 resonance.

Therefore, the most commonly found realization involves annihilation through A^0 . Because the $A^0 b\bar{b}$ coupling is proportional to $m_b \tan \beta$, this usually entails large values of $\tan \beta$. (Annihilation through h^0 is also possible.)

The region of parameter space where this happens is often called the “ A -funnel” or “Higgs funnel” or “Higgs resonance region”.

3. A third effective annihilation mechanism is obtained if \tilde{N}_1 mixes so as to obtain a significant higgsino or wino admixture.

Then both Fig. 27c ($\tilde{N}_1 \tilde{N}_1 \rightarrow W^+ W^-$ via \tilde{C}_1 exchange) and the co-annihilation diagrams of Fig. 28 can be important.

In this “focus point” region of parameter space, where $|\mu|$ is not too large, the LSP can have a significant higgsino content and can yield the correct relic abundance even for very heavy squarks

and sleptons. (This is motivated by focusing properties of the renormalization group equations, which allow $|\mu| \ll m_0^2$ in minimal supergravity inspired models.)

It is also possible to arrange for just enough wino content in the LSP to do the job by choosing M_1/M_2 appropriately.

Of course, too much higgsino or wino content will yield to much annihilation and too small (rather than too large) a relic density.

4. A fourth possibility, the “**sfermion co-annihilation region**” of parameter space, is obtained if there is a sfermion that happens to be less than a few GeV heavier than the LSP.

In many model frameworks, this is most naturally the lightest stau, $\tilde{\tau}_1$, but it could also be the lightest top squark, \tilde{t}_1 .

A significant density of this sfermion will then coexist with the LSP around the freeze-out time, and so annihilations involving the sfermion with itself or with the LSP, including those of the type shown in Fig. 29 ((a,b) $\tilde{N}_1 \tilde{f} \rightarrow f + (\gamma, Z)$ via f exchange in

s -channel of \tilde{f} exchange in the t -channel and (c) $\tilde{f}\tilde{f} \rightarrow ff$ via \tilde{N}_1 t -channel exchange) will further dilute the number of sparticles and so the eventual dark matter density.

- An aside on co-annihilation

In some cases, particles other than the WIMP itself can play an important role in the freeze-out process. Before such a particle can significantly impact the relic density of a WIMP, however, it must first manage to be present at the temperature of freeze-out.

The relative abundances of two species at freeze-out can be very roughly estimated by

$$\frac{n_Y}{n_X} \sim \frac{e^{-m_Y/T_f}}{e^{-m_X/T_f}}. \quad (523)$$

Considering, for example, a particle with a mass twice that of the WIMP and a typical freeze-out temperature of $m_X/T_f \approx 20$, there

will be only $\sim e^{-40}/e^{-20} \sim 10^{-9}$ Y particles for every X at freeze-out, thus making Y completely irrelevant.

If m_Y were only 10% larger than m_X , however, we estimate $n_Y/n_X \sim e^{-22}/e^{-20} \sim 10^{-1}$. In this quasi-degenerate case, the additional particle species can potentially have a significant impact on the dark matter relic abundance.

To quantitatively account for other species in the calculation of the relic abundance of a WIMP, we make the following substitution (for both a and b) into Eqs. 515 and 514:

$$\sigma_{\text{Ann}} \rightarrow \sigma_{\text{Eff}}(x) = \sum_{i,j} \sigma_{i,j} \frac{g_i g_j}{g_{\text{Eff}}^2(x)} (1 + \Delta_i)^{3/2} (1 + \Delta_j)^{3/2} e^{-x(\Delta_i + \Delta_j)}, \quad (524)$$

where the double sum is over all particle species ($i, j = 1$ denoting the WIMP itself) and $\sigma_{i,j}$ is the cross section for the coannihilation of species i and j (or self-annihilation in the case of $i = j$) into

Standard Model particles.

As the effective annihilation cross section has a strong dependence on x , we must integrate Eqs. 515 and 514 over x (or T).

The quantities $\Delta_i = (m_i - m_1)/m_1$ denote the fractional mass splittings between the species i and the WIMP. The effective number of degrees of freedom, $g_{\text{Eff}}(x)$, is given by:

$$g_{\text{Eff}}(x) = \sum_i g_i (1 + \Delta_i)^{3/2} e^{-x\Delta_i}. \quad (525)$$

To better understand how the introduction of particles other than the WIMP can effect the process of freeze-out, let's consider a few simple cases.

1. First, consider one additional particle with a mass only slightly above that of the WIMPs ($\Delta_2 \ll 1$), and with a comparatively large coannihilation cross section, such that $\sigma_{1,2} \gg \sigma_{1,1}$.

In this case, $g_{\text{Eff}} \approx g_1 + g_2$, and $\sigma_{\text{Eff}} \approx \sigma_{1,2} g_1 g_2 / (g_1 + g_2)^2$.

Since σ_{Eff} is much larger than the WIMP's self-annihilation cross section, the relic density of WIMPs will be sharply suppressed. This is the case that is usually meant by the term “**coannihilation**”.

2. Alternatively, consider the opposite case in which the WIMP and the additional quasi-degenerate particle do not coannihilate efficiently ($\sigma_{1,2} \ll \sigma_{1,1}, \sigma_{2,2}$).

Here, $\sigma_{\text{Eff}} \approx \sigma_{1,1} g_1^2 / (g_1 + g_2)^2 + \sigma_{2,2} g_2^2 / (g_1 + g_2)^2$, which in some cases can actually be *smaller* than that for the process of self-annihilation alone, leading to an *enhanced* relic abundance.

Physically speaking, what is going on here is that the two species are each freezing out independently of each other, after which the heavier species decays, producing additional WIMPs as a byproduct.

3. As an extreme version of this second case, consider a scenario in which the lightest state is not a WIMP, but is instead a purely gravitationally interacting particle.

- A slightly heavier particle with weak interactions will self-annihilate much more efficiently than it will coannihilate with the lightest particle ($\sigma_{1,2}$ is negligible), leading the two states to freeze-out independently.
- The gravitationally interacting particle, however, never reaches thermal equilibrium, so could potentially have not been produced in any significant quantities up until this point.
- Well after freezing out, the heavier particles will eventually decay, producing the stable gravitationally interacting lightest state.
- Although the resulting particles are not WIMPs (they do not have weak interactions), they are naturally produced with approximately the measured dark matter abundance because of the WIMP-like properties of the heavier state.

In other words, this case – known as the “**superWIMP**” **scenario** – makes use of the coincidence between the electroweak scale and the measured dark matter abundance **without the dark matter actually**

consisting of WIMPs.

Because gravitationally interacting particles and other much less than weakly interacting particles are almost impossible to detect astrophysically, superWIMPs are among the dark matter hunter's worst nightmares.

- Quick Review and the CMSSM $m_0 - m_{1/2}$ plane

In the constrained MSSM we have universal m_0 and $m_{1/2}$ for scalar and gaugino soft masses, respectively, at M_U .

This yields a ratio at the electroweak scale of $M_1 = \frac{5}{3} \tan^2 \theta_W M_2 \approx 0.5 M_2$.

In this case, the lightest neutralino has only a small wino fraction and is largely bino-like (higgsino-like) for $M_1 \ll |\mu|$ ($M_1 \gg |\mu|$).

We must also keep in mind that $|\mu|$ is not an independent parameter — $|\mu|$ and b (the soft relative to μ) are determined by the minimization

conditions

$$m_{H_u}^2 + |\mu|^2 - b \cot \beta - (m_Z^2/2) \cos(2\beta) = 0, \quad (526)$$

$$m_{H_d}^2 + |\mu|^2 - b \tan \beta + (m_Z^2/2) \cos(2\beta) = 0, \quad (527)$$

in terms of m_Z^2 and $\tan \beta$, given values of $m_{H_u}^2$ and $m_{H_d}^2$ obtained by RG equations from m_0^2 . In practice, this means that if m_0 and $\tan \beta$ are held fixed then $|\mu|$ will remain relatively constant as $m_{1/2}$ is increased. This in turn implies that at some value of $m_{1/2}$ the values of M_1 and $|\mu|$ will be similar, allowing a LSP with significant higgsino content.

As we said earlier, since a bino-like LSP is very weakly interacting, over much or most of the supersymmetric parameter space, the relic abundance of neutralinos is predicted to be in excess of the observed dark matter density.

To avoid this, we are forced to consider the regions of parameter

space which lead to especially efficient neutralino annihilation in the early universe.

In particular, the following scenarios are among those which can lead to a phenomenologically viable density of neutralino dark matter:

- If the lightest neutralino has a significant higgsino or wino fraction, it can have fairly large couplings and, as a result, annihilate very efficiently.
- If the mass of the lightest neutralino is near a resonance, such as the CP-odd Higgs pole, it can annihilate efficiently, even with relatively small couplings.
- If the lightest neutralino is only slightly lighter than another superpartner, such as the lightest stau, coannihilations between these two states can very efficiently deplete the dark matter abundance.

To illustrate these regions in CMSSM parameter space, it is convenient

to begin with some older figures shown in Fig. 31. We will then progress to more recent figures where additional constraints and the most recent Ω_{DM} measurements (with smaller errors) are incorporated. In these early figures, note the following:

1. In each frame, the narrow blue regions denote the parameter space in which neutralino dark matter is predicted to be generated with the desired abundance ($0.0913 < \Omega_{\chi_0} h^2 < 0.1285$).
2. In the corridor along side of the LEP chargino bound ($m_{\chi^\pm} > 104$ GeV), μ and M_1 are comparable in magnitude, leading to a mixed bino-higgsino LSP with large couplings.

Within the context of the CMSSM, this is often called the “focus point” region.

3. In the bottom portion of each frame, the lightest stau ($\tilde{\tau}_1$) is the LSP, and thus does not provide a viable dark matter candidate. Just outside of this region, however, the $\tilde{\tau}_1$ is slightly heavier than the lightest neutralino, leading to a neutralino LSP which efficiently

coannihilates with the nearly degenerate stau.

4. In the lower right frame, a viable region also appears along the CP-odd Higgs resonance ($m_{\chi^0} \approx m_A/2$). This is often called the “*A*-funnel region”.

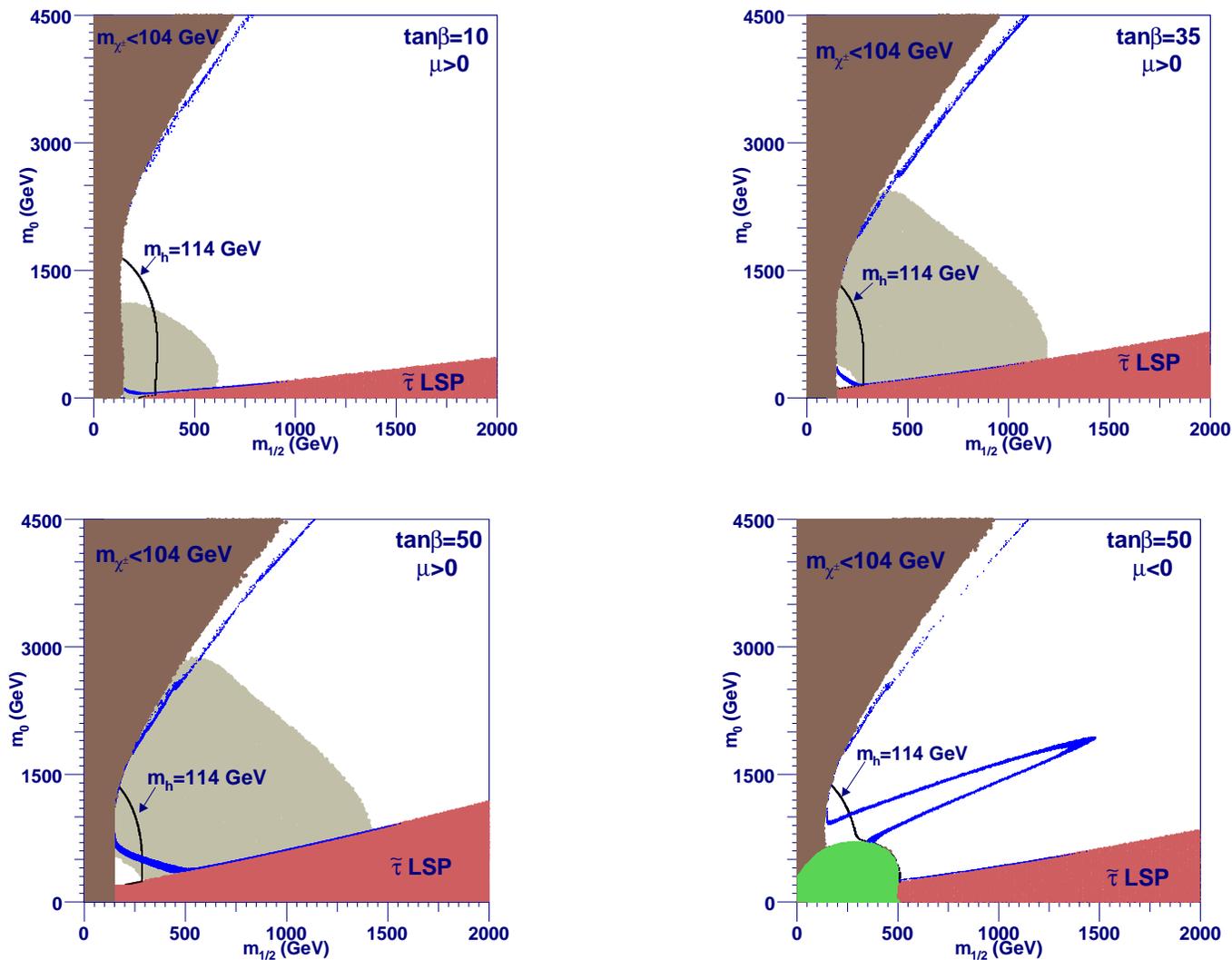


Figure 31: Representative regions of the CMSSM parameter space. The blue regions predict a neutralino density consistent with the measured dark matter abundance. The shaded region to the upper left has $m_{\tilde{\chi}^0_1}$ below the LEP limit and that to the lower right has $\tilde{\tau}_1$ as the LSP. The LEP bound on the light Higgs mass is shown as a solid line ($m_h = 114$ GeV). The region favored (at 3σ) by measurements of $(g - 2)_\mu$ are shown as a light shaded region. In a,b,c, we have used $A_0 = 0$ and $\mu > 0$. In d, $\mu < 0$ and $(g - 2)_\mu$ is bad.

The above plot, from Hooper's TASI notes, is representative of the 2004 situation. It fairly clearly displays the focus, funnel and coannihilation regions.

Of course, measurements have gotten more precise since then and theorists have become more demanding about the model being consistent with all known measurements.

The Ellis + Olive paper of January 2010 (arXiv:1001.3651) summarizes something close to the current situation, but is not quite up to date.

As above, for given values of $\tan\beta$, A_0 , and $\text{sgn}(\mu)$, the regions of the CMSSM parameter space that yield an acceptable relic density and satisfy the other phenomenological constraints are conveniently displayed in the $(m_{1/2}, m_0)$ plane. Fig. 32 displays, for $\tan\beta = 10$ (a) and $\tan\beta = 50$ (b), the impacts of the most relevant constraints.

Constraints

1. The LEP lower limits on the chargino mass: $m_{\chi^\pm} > 104$ GeV,

on the selectron mass: $m_{\tilde{e}} > 99$ GeV and on the Higgs mass: $m_h > 114$ GeV.

The former two constrain $m_{1/2}$ and m_0 directly via the sparticle masses, and the latter indirectly via the sensitivity of radiative corrections to the m_{h0} to the sparticle masses, principally $m_{\tilde{t},\tilde{b}}$.

– Here the code “FeynHiggs” is used for the calculation of m_h . It would be prudent to assign an uncertainty of 3 GeV to this calculation.

– Nevertheless, the Higgs limit imposes important constraints, principally on $m_{1/2}$ and particularly at low $\tan\beta$.

2. Another constraint is the requirement that the branching ratio for $b \rightarrow s\gamma$ be consistent with the experimental measurement.

– These measurements agree with the Standard Model, and therefore provide bounds on MSSM particles, such as the chargino and charged Higgs bosons, in particular.

– Typically, the $b \rightarrow s\gamma$ constraint is more important for $\mu < 0$,

- but it is also relevant for $\mu > 0$, particularly when $\tan\beta$ is large.
- The constraint imposed by measurements of $b \rightarrow s\gamma$ also exclude small values of $m_{1/2}$.
3. Finally, there are regions of the $(m_{1/2}, m_0)$ plane that are favoured by the Brookhaven National Laboratory measurement of $g_\mu - 2$. Here we assume the Standard Model calculation of $g_\mu - 2$ using e^+e^- data, and indicate by dashed and solid lines the contours of 1- and 2- σ level deviations induced by supersymmetry.
 4. The most precise constraint on supersymmetry may be that provided by the density of cold dark matter, as determined from astrophysical and cosmological measurements by WMAP and other experiments:

$$\Omega_{CDM} = 0.1099 \pm 0.0062. \quad (528)$$

- Applied straightforwardly to the relic LSP density $\Omega_{LSP}h^2$, this would give a very tight relation between supersymmetric model parameters, fixing some combination of them at the %

level, which would essentially reduce the dimensionality of the supersymmetric parameter space by one unit.

- Assuming that the LSP is the lightest neutralino $\chi = \tilde{N}_1$ and the freeze-out formalism, respecting the constraint Eq. (528) would force the CMSSM into one of the narrow WMAP ‘strips’ in planar projections of the parameters, as illustrated by the narrow light (turquoise) regions in Fig. 32.
- If supersymmetry is not the only contribution to the cold dark matter, Eq. (528) should be interpreted as an upper limit on $\Omega_{LSP} h^2$.

However, most of the supersymmetric parameter space in the CMSSM gives a supersymmetric relic density that exceeds the WMAP range Eq. (528), e.g., above the WMAP ‘strip’ in Fig. 32, and the regions with lower density generally correspond to *lower* values of the sparticle masses, i.e., below the WMAP ‘strip’ in Fig. 32.

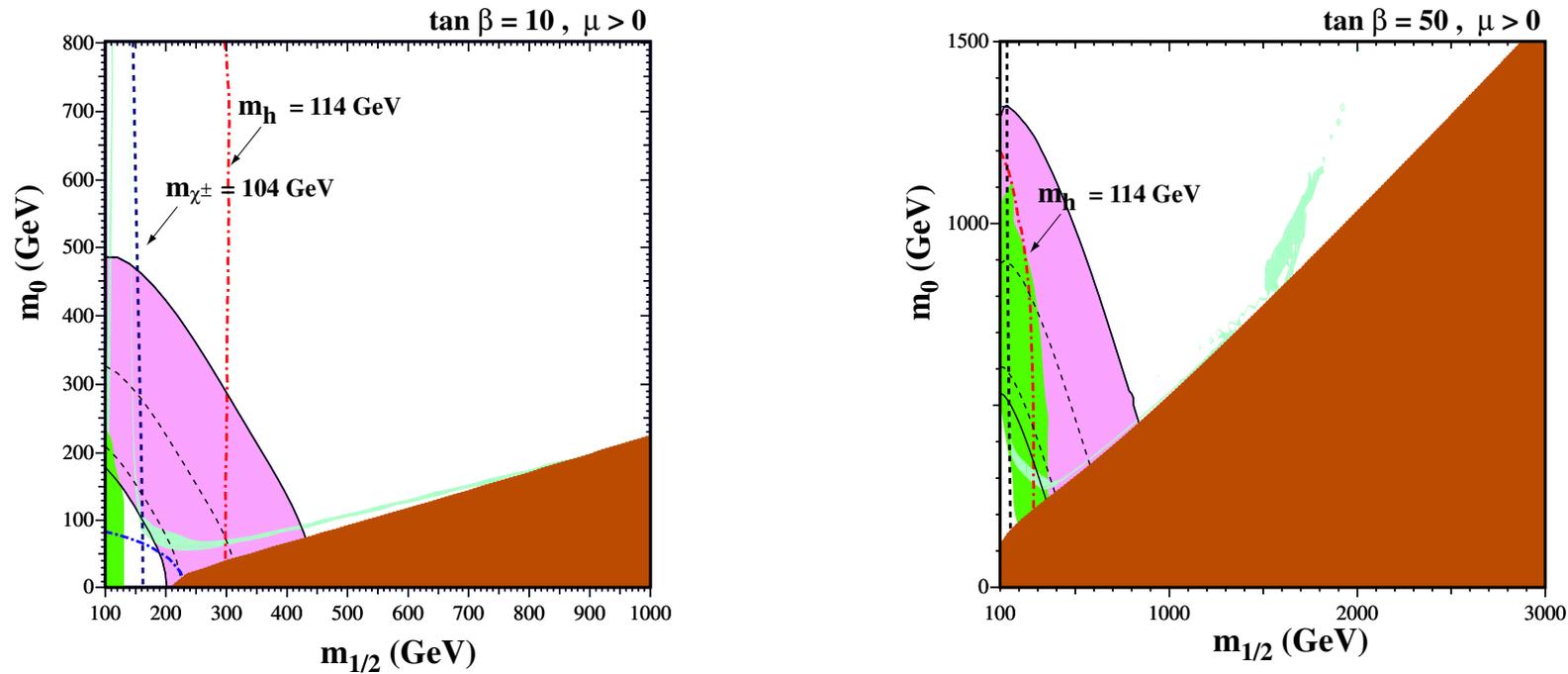


Figure 32: The $(m_{1/2}, m_0)$ planes for (a) $\tan \beta = 10$ and (b) $\tan \beta = 50$, assuming $\mu > 0$, $A_0 = 0$, $m_t = 175$ GeV and $m_b(m_b)_{\overline{MS}}^{SM} = 4.25$ GeV. The near-vertical (red) dot-dashed lines are the contours for $m_h = 114$ GeV, and the near-vertical (black) dashed line is the contour $m_{\chi^\pm} = 104$ GeV. Also shown by the dot-dashed curve in the lower left is the region excluded by the LEP bound $m_{\tilde{e}} > 99$ GeV. The medium (dark green) shaded region is excluded by $b \rightarrow s\gamma$, and the light (turquoise) shaded area is the cosmologically preferred region. In the dark (brick red) shaded region, the LSP is the charged $\tilde{\tau}_1$. The region allowed by the E821 measurement of a_μ at the $2\text{-}\sigma$ level, is shaded (pink) and bounded by solid black lines, with dashed lines indicating the $1\text{-}\sigma$ ranges. From Ellis+Olive, arXiv:1001.3651.

More Discussion of the WMAP strip

- The locations of these WMAP ‘strips’ do vary significantly with the choices of other supersymmetric parameters, as can be seen by comparing the cases of $\tan\beta = 10, 50$ in Fig. 32(a, b).

As one varies $\tan\beta$, the WMAP ‘strips’ cover much of the $(m_{1/2}, m_0)$ plane.

- Once again, several different regions of the WMAP ‘strips’ in the CMSSM $(m_{1/2}, m_0)$ plane can be distinguished, in which different dynamical processes are dominant.

1. At low values of $m_{1/2}$ and m_0 , simple $\chi - \chi$ annihilations via crossed-channel sfermion exchange are dominant, but this ‘bulk’ region is now largely excluded by the LEP lower limit on the Higgs mass, m_h .

2. At larger $m_{1/2}$, but relatively small m_0 , close to the boundary of the region where the $\tilde{\tau}_1$ is lighter than the lightest neutralino: $m_{\tilde{\tau}_1} < m_\chi$, coannihilation between the χ and sleptons are

important in suppressing the relic χ density into the WMAP range of Eq. (528), as seen in Fig. 32.

3. At larger $m_{1/2}, m_0$ and $\tan\beta$, the relic χ density may be reduced by rapid annihilation through direct-channel H, A Higgs bosons, as seen in Fig. 32(b) (the isolate turquoise blob).
4. Finally, the relic density can again be brought down into the WMAP range (528) at large m_0 (not really visible in Fig. 32 since the strip is so narrow), in the ‘focus-point’ region close the boundary where electroweak symmetry breaking ceases to be possible and the lightest neutralino χ acquires a significant higgsino component.

Reconciling the WMAP Strip with other constraints

- As seen in Fig. 32, the relic density constraint is compatible with relatively large values of $m_{1/2}$ and m_0 , and it is interesting to look for any indication where the supersymmetric mass scale might lie within this range, using the available phenomenological and

cosmological constraints.

- In this regard, Fig. 32 shows that $(g - 2)_\mu$ and the Higgs mass constraint of $m_{h^0} > 114$ GeV are very important.
- * The $\tan \beta = 10$ plot shows that, **within the WMAP strip**, a point that is close to the 1σ band for $(g - 2)_\mu$ and just past the $m_{h^0} > 114$ GeV LEP limit is located at $m_{1/2} \sim 300$ GeV and $m_0 \sim 80$ GeV.
- * At $\tan \beta = 50$ things have shifted only somewhat. If we want to keep m_{h^0} as close to 114 GeV as possible (as precision data suggests) while remaining in the 1σ band for $(g - 2)_\mu$ and inside the WMAP strip, then $m_{1/2} = 400$ GeV and $m_0 \sim 250$ GeV is preferred.

To be more quantitative, it is desirable to use a global likelihood analysis to pin down the available parameter space in the CMSSM and related models.

One can avoid the dependence on priors (*e.g.* precise value of m_t) by

performing a pure likelihood analysis or a purely χ^2 -based fit.

Let me summarize the results from one such analysis, which used a Markov-Chain Monte Carlo (MCMC) technique to explore efficiently the likelihood function in the parameter space of the CMSSM.

A full list of the observables and the values assumed for them in this global analysis are given in *Buchmueller:2007zk*, as updated in *Buchmueller:2008qe*.

- The 68% and 95% confidence-level (C.L.) regions in the $(m_{1/2}, m_0)$ plane of the CMSSM are shown in Fig. 33.
- Also shown for comparison are the physics reaches of ATLAS and CMS with 1/fb of integrated luminosity. (MET stands for missing transverse energy, SS stands for same-sign dilepton pairs, and the sensitivity for finding the lightest Higgs boson in cascade decays of supersymmetric particles is calculated for 2/fb of data.)

You should of course compare to the earlier Fig. 30 showing the

latest CMS limits that have just become available based on just 35 pb^{-1} of data.

The preferred fit point is close to the excluded region!

- The likelihood analysis assumed $\mu > 0$, as motivated by the sign of the apparent discrepancy in $g_\mu - 2$, but sampled all values of $\tan \beta$ and A_0 : the experimental sensitivities were estimated assuming $\tan \beta = 10$ and $A_0 = 0$, but are probably not very sensitive to these assumptions.
- The global maxima of the likelihood function (indicated by the black dot) is at $m_{1/2} = 310 \text{ GeV}$, $m_0 = 60 \text{ GeV}$, $A_0 = 240 \text{ GeV}$, $\tan \beta = 11$ and $\chi^2/N_{dof} = 20.4/19$ (37% probability).
- It is encouraging that the best-fit point lies well within the LHC discovery range, as do the 68% and most of the 95% C.L. regions.
- All points with similarly good χ^2 have similar values of $m_{1/2}$, m_0 and $\tan \beta$, the most important parameters for the sparticle spectrum.

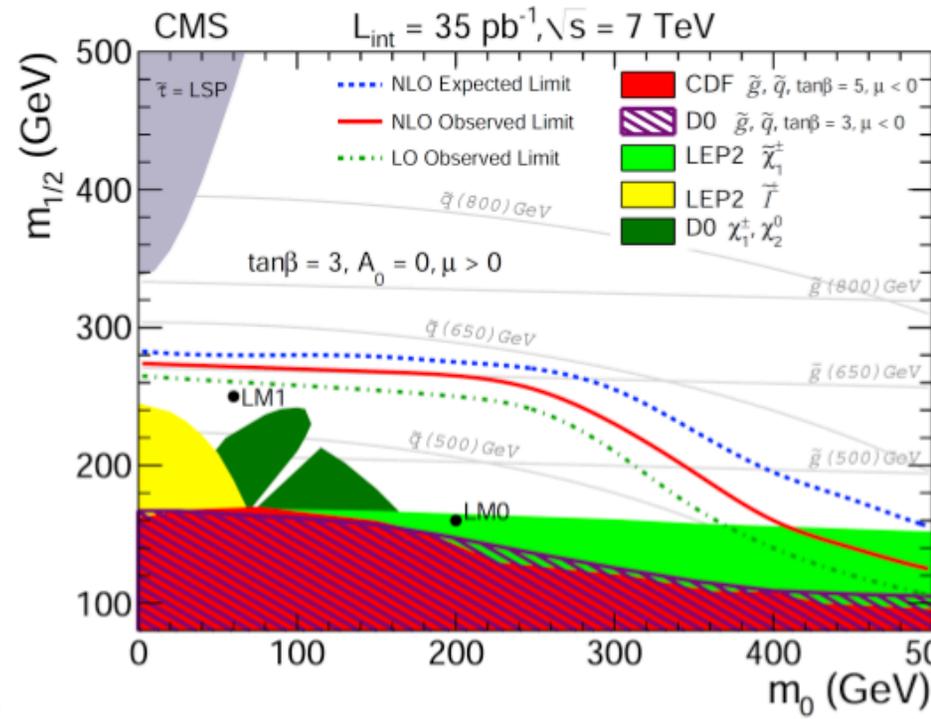
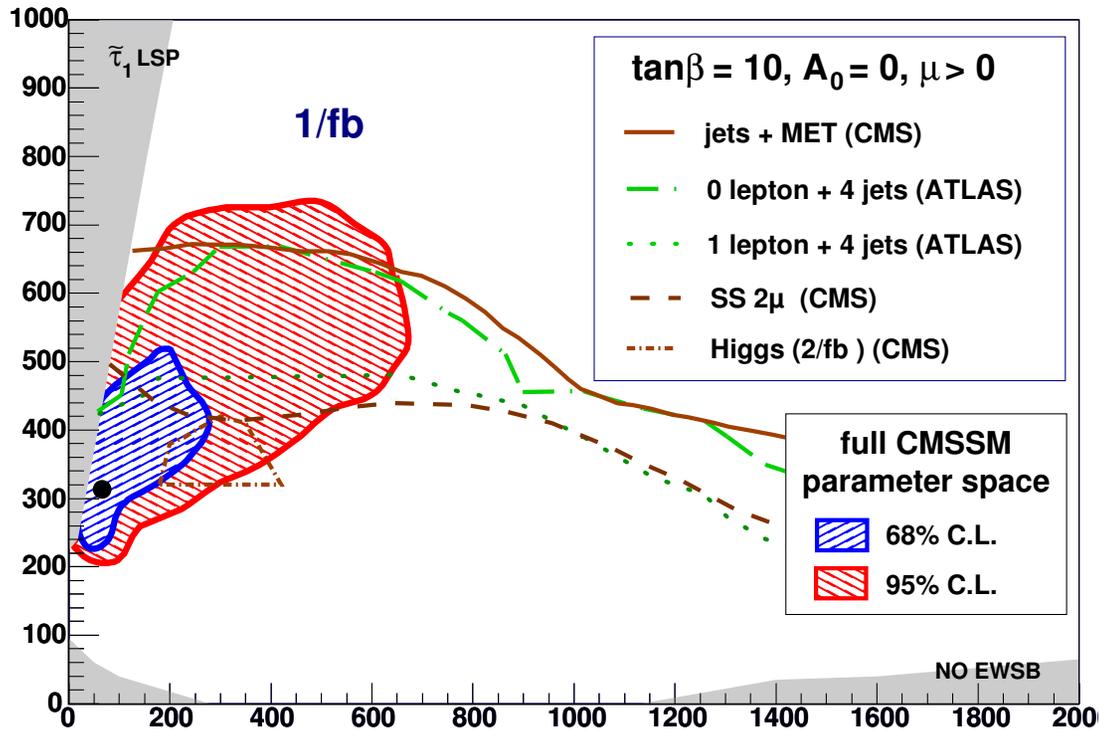


Figure 33: Left: The $(m_0, m_{1/2})$ plane (m_0 is on x -axis and $m_{1/2}$ is on y -axis, contrary to earlier plots) in the CMSSM showing the regions favoured in a likelihood analysis at the 68% (blue) and 95% (red) confidence levels. The best-fit point is shown as the black point. Also shown are the discovery contours in different channels for the LHC with 1/fb (2/fb for the Higgs search in cascade decays of sparticles). Right: A repeat of the limits from Aspen, winter 2011.

- Beyond the CMSSM — the 19 parameter space

In arXiv:1012.0248, Baer implements a linear scan over the following 19 GUT scale parameters **in particular, without assuming unification.**

- Gaugino masses: $M_1, M_2, M_3 : 0 - 3.5$ TeV
- First/second generation scalar masses: $m_{Q_1}, m_{U_1}, m_{D_1}, m_{L_1}, m_{E_1} : 0 - 3.5$ TeV,
- Third generation scalar masses: $m_{Q_3}, m_{U_3}, m_{D_3}, m_{L_3}, m_{E_3} : 0 - 3.5$ TeV,
- Higgs soft masses: $m_{H_u}, m_{H_d} : 0 - 3.5$ TeV,
- trilinear soft terms: $A_t, A_b, A_\tau : -3.5$ TeV \rightarrow 3.5 TeV,
- ratio of weak scale Higgs vevs $\tan \beta : 2 - 60$.

He adopts a common mass for first and second generation scalars so as to avoid SUSY FCNC processes.

To gain an acceptable sparticle mass solution, one requires.

1. the lightest SUSY particle (LSP) is the neutralino $\tilde{\chi}_1^0$,
2. the lightest chargino, if non-wino-like, obeys the LEP2 limit $m_{\tilde{\chi}_1} > 103.5$ GeV,
3. the lightest chargino, if wino-like, obeys the LEP2 limit $m_{\tilde{\chi}_1} > 91.9$ GeV,
4. the light Higgs mass obeys the LEP2 limit $m_h > 111$ GeV (which allows for a roughly 3 GeV uncertainty in the theory calculation as applied to the actual limit where $m_h > 114.4$ GeV).

For each acceptable solution, the neutralino relic density $\Omega_{\tilde{\chi}_1^0} h^2$ is calculated using the IsaReD program.¹⁷

A few basic points

- Since we assume the neutralino to be in thermal equilibrium, our relic density results do not explicitly depend on the value of the

¹⁷IsaReD calculates all relevant neutralino annihilation and co-annihilation reactions, as obtained using CalcHEP, and then calculates the relativistic thermally-averaged (co)-annihilation cross sections times relative velocity. Once the freeze-out temperature is determined, then the relic density at the present time is found by integrating the Boltzmann equation as formulated for a FRW universe.

re-heat temperature of the Universe T_R after inflation.

- However, we must assume $T_R > T_f \sim m_{\tilde{\chi}_1^0}/20$ so that T_R is above the neutralino freeze-out temperature.
- Further, if $T_R \gtrsim 10^{10}$ GeV, then thermal production of gravitinos in the early Universe, followed by decays to the LSP, will overproduce neutralino dark matter.

Hence, for the neutralino CDM relic density calculations, we must assume here that $m_{\tilde{\chi}_1^0}/20 \lesssim T_R \lesssim 10^{10}$ GeV.

Results from a linear scan over the above SUGRA-19 parameter space is shown in Fig. 34, in the $\Omega_{\tilde{\chi}_1^0} h^2$ vs. $m_{\tilde{\chi}_1^0}$ plane.

- The various solutions are color coded according to the gaugino/higgsino content of the neutralino.
 - * If the bino-component $|N_{11}| > 0.9$, then the neutralino is labeled as bino-like (blue diamonds);
 - * if the wino-component $|N_{12}| > 0.9$, then it is labeled wino-like

(purple \times);

- * if the higgsino components $\sqrt{|N_{13}|^2 + |N_{14}|^2} > 0.9$, then it is labeled as higgsino-like (red squares).
- * If the neutralino falls into none of these categories, then it is labeled as “mixed” DM: (orange circles).

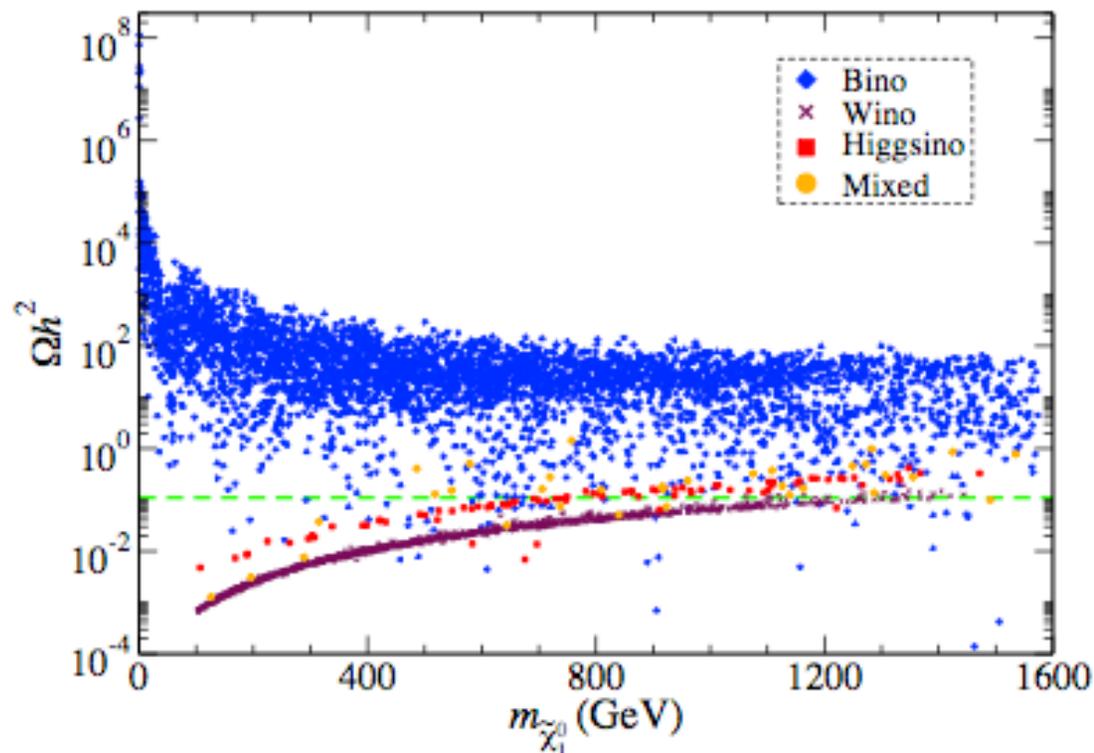


Figure 34: Thermal abundance of neutralino cold dark matter from a linear scan over the SUGRA-19 parameter space. We plot versus the neutralino mass. Models with mainly bino, wino, higgsino or a mixture are indicated by the various color and symbol choices. There are 5252 points in the figure.

- We see from Fig. 34 the following.
 - * the bino-like neutralinos tend to populate the region with $\Omega_{\tilde{\chi}_1^0} h^2 \gg 0.1$, *i.e.* usually about 2-3 orders of magnitude too high.
 - * For low values of $m_{\tilde{\chi}_1^0}$, the abundance tends to be more like 3-5 orders of magnitude too high.
 - * A few bino-like points do tend to make it into the $\Omega_{\tilde{\chi}_1^0} h^2 \sim 0.1$ region; these solutions tend to come from various co-annihilation or resonance annihilation processes.
 - * To obtain the required relic abundance via co-annihilation, the LSP-NLSP mass gap must be tuned to just the right value.
 - * To obtain the required relic abundance via resonance annihilation, the LSP mass must be adjusted to be close to half the mass of the resonance.
 - * These co-annihilation and resonance annihilation points are quite hard, but not impossible, to generate using a random scan over GUT scale parameters.

- * The higgsino-like and wino-like CDM bands also show up as distinct lines, typically with $\Omega_{\tilde{\chi}_1^0} h^2$ too low by 1-2 orders of magnitude unless $m_{\tilde{\chi}_1^0} \gtrsim 800 - 1200$ GeV.
- * The wino-like band is relatively well-populated, as this just requires M_2 to be the lightest of the gaugino masses at the weak scale.
- * The higgsino-like band is relatively less populated, showing that higgsino-like CDM is rather fine-tuned if one starts with GUT scale parameters.
- * The points with the lowest population are those with mixed bino-higgsino-wino CDM.
 These “well-tempered neutralino” points most naturally tend to populate the $\Omega_{\tilde{\chi}_1^0} h^2 \sim 0.1$ line, but they do require a fine-tuning to avoid a bino, wino or higgsino dominance.
- * Especially at low $m_{\tilde{\chi}_1^0}$, relatively few solutions are found with $\Omega_{\tilde{\chi}_1^0} h^2 \sim 0.1$.

- To appreciate more clearly the dark matter probability distribution for a linear scan of SUGRA-19 parameter space, we project the model points of Fig. 34 onto the $\Omega_{\tilde{\chi}_1^0} h^2$ axis in Fig. 35a).

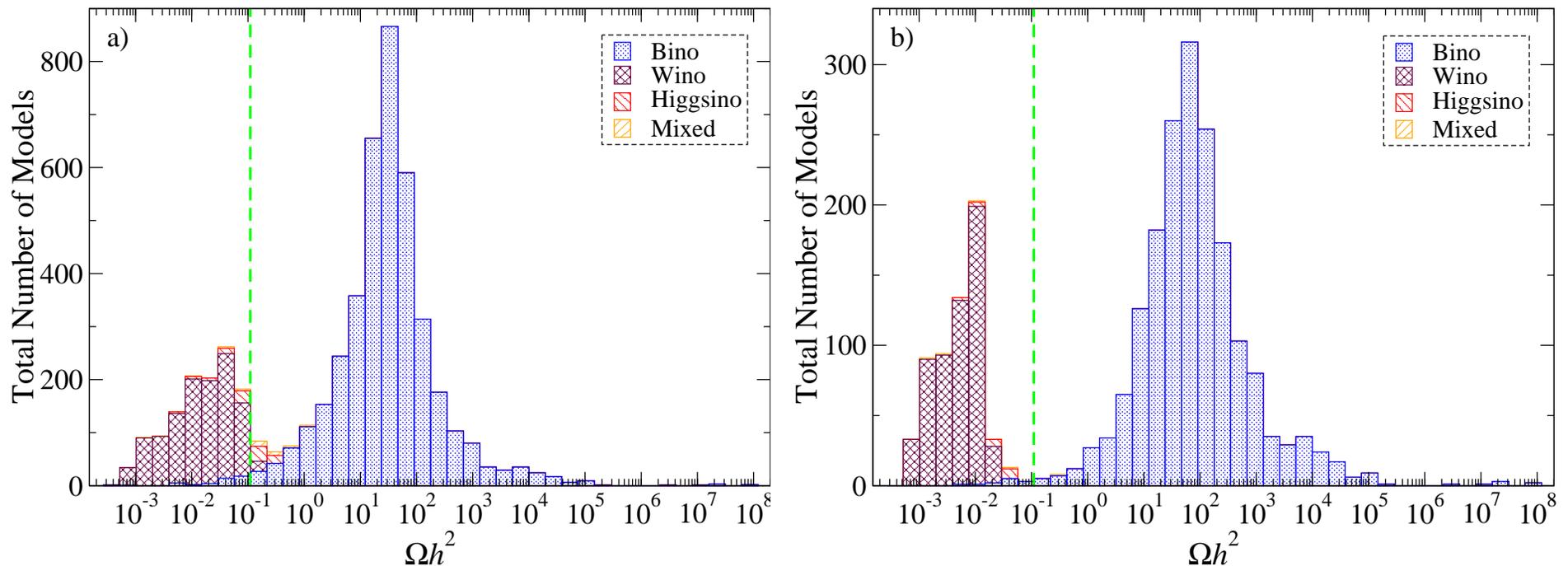


Figure 35: Projection of the number of models generated by a linear scan over SUGRA-19 parameters, versus neutralino relic density $\Omega_{\tilde{\chi}_1^0} h^2$. Models with mainly bino, wino, higgsino or a mixture are indicated by the various color and symbol choices. In frame b)., we require only models with $m_{\tilde{\chi}_1^0} < 500$ GeV to avoid too large a fine-tuning of the SUSY parameters.

- Here we see the most probable value of $\Omega_{bino}h^2$ is $\sim 10 - 100$ for bino-like dark matter (blue histogram), while the most probable value for wino-like dark matter is $\Omega_{wino}h^2 \sim 0.005 - 0.05$.
 - The dip between these two cases is partially filled in by cases of bino, higgsino or wino, or a mixture, with the minimum probability lying around $\Omega_{\tilde{\chi}_1^0}h^2 \sim 0.2 - 0.4$, *i.e.* just above the measured value.
 - A large number of the wino and higgsino dark matter solutions with $\Omega_{\tilde{\chi}_1^0}h^2 \sim 0.1$ come from models with very heavy neutralinos: $m_{\tilde{\chi}_1^0} \gtrsim 800$ GeV.
 - If the $\tilde{\chi}_1^0$ is the LSP, as assumed here, then all other sparticles are heavier— and usually much heavier— than this value, and will likely lead to solutions with high electroweak fine-tuning.
- In Fig. 35b)., we plot the number of model solutions from the SUGRA-19 scan versus $m_{\tilde{\chi}_1^0}$, where in addition we require (somewhat

arbitrarily) that $m_{\tilde{\chi}_1^0} < 500$ GeV, so the solutions are not too fine-tuned with regard to electroweak symmetry breaking.

- In this case, the higgsino and wino LSP models which naturally give $\Omega_{\tilde{\chi}_1^0} h^2 \sim 0.1$ are all excluded.
- The maxima of bino-like solutions has moved up slightly to $\Omega_{\text{bino}} h^2 \sim 50 - 100$, while wino-like solutions peak at $\Omega_{\text{wino}} h^2 \sim 0.01$.
- The minimum of the probability distribution lies very close to the measured value $\Omega_{\text{CDM}} h^2 \sim 0.1$.
- With sparticle masses generally at the TeV or below scale, the measured relic density lies at the *least likely value* as predicted by the SUGRA-19 model.
- In this case, it would be *extremely fortuitous* if the lightest neutralino of SUGRA theories was in fact the dark matter particle.

- **Cautions**

However, it is important to keep in mind that a set of MSSM Lagrangian parameters that fails to predict the correct relic dark matter abundance by the standard thermal mechanisms is **not** ruled out as a model for collider physics.

This is because simple extensions can completely change the relic abundance prediction without changing the predictions for colliders much or at all.

1. For example, if the model predicts a neutralino dark matter abundance that is too small, one need only assume another sector (even a completely disconnected one) with a stable neutral particle, or that the dark matter is supplied by some non-thermal mechanism such as out-of-equilibrium decays of heavy particles.
2. If the predicted neutralino dark matter abundance appears to be too large, one can assume that R -parity is slightly broken, so that the offending LSP decays before nucleosynthesis; this would require some other unspecified dark matter candidate.

3. Or, the dark matter LSP might be some particle that the lightest neutralino decays into.

In such cases, the dark matter density after the lightest neutralino decays would be reduced compared to its naively predicted value by a factor of $m_{\text{LSP}}/m_{\tilde{N}_1}$, provided that other sources for the LSP relic density are absent.

(a) One possibility is a gravitino LSP.

In this case, the gravitino is much lighter than the \tilde{N}_1 and so the density of \tilde{N}_1 's would need to be very high (small annihilation cross section) in order to achieve the observed dark matter density.

(b) Another example is obtained by extending the model to solve the strong CP problem with an invisible axion, which can allow the LSP to be a very weakly-interacting axino (the fermionic supersymmetric partner of the axion).

Axinos need not have a small mass and so this case would fall

- into the “superWIMP” category of scenario discussed earlier.
4. A correct density for neutralino LSPs can also be obtained by assuming that they are produced non-thermally in reheating of the universe after neutralino freeze-out but before nucleosynthesis.
 5. Finally, in the absence of a compelling explanation for the apparent cosmological constant, it seems possible that the standard model of cosmology will still need to be modified in ways not yet imagined.

Dark Matter Detection Overview

- If neutralino LSPs really make up the cold dark matter, then their local mass density in our neighborhood ought to be of order $0.3 \text{ GeV}/\text{cm}^3$ [much larger than the density averaged over the largest scales, Eq. (508)] in order to explain the dynamics of our own galaxy.

LSP neutralinos could be detectable directly through their weak interactions with ordinary matter, or indirectly by their ongoing annihilations.

However, the halo is subject to significant uncertainties in overall size, velocity, and clumpiness, so even if the Lagrangian parameters were known exactly, the signal rates would be quite indefinite, possibly even by orders of magnitude.

- The direct detection of \tilde{N}_1 depends on their elastic scattering off of

heavy nuclei in a detector.

At the parton level, \tilde{N}_1 can interact with a quark by virtual exchange of squarks in the s -channel, or Higgs scalars or a Z boson in the t -channel.

It can also scatter off of gluons through one-loop diagrams.

The scattering mediated by neutral Higgs scalars is suppressed by tiny Yukawa couplings, but is coherent for the quarks and so can actually be the dominant contribution for nuclei with larger atomic weights, if the squarks are heavy.

The energy transferred to the nucleus in these elastic collisions is typically of order tens of keV per event.

There are important backgrounds from natural radioactivity and cosmic rays, which can be reduced by shielding and pulse-shape analysis. A wide variety of current or future experiments are sensitive

to some, but not all, of the parameter space of the MSSM that predicts a dark matter abundance in the range of Eq. (507).

- Another, more indirect, way to detect neutralino LSPs is through ongoing annihilations. This can occur in regions of space where the density is greatly enhanced.

If the LSPs lose energy by repeated elastic scattering with ordinary matter, they can eventually become concentrated inside massive astronomical bodies like the Sun or the Earth. In that case, the annihilation of neutralino pairs into final states leading to neutrinos is the most important process, since no other particles can escape from the center of the object where the annihilation is going on.

In particular, muon neutrinos and antineutrinos from $\tilde{N}_1\tilde{N}_1 \rightarrow W^+W^-$ or ZZ , (or possibly $\tilde{N}_1\tilde{N}_1 \rightarrow \tau^+\tau^-$ or $\nu\bar{\nu}$, although these are p -wave suppressed) will travel large distances, and can be detected in neutrino telescopes.

The neutrinos undergo a charged-current weak interaction in the earth, water, or ice under or within the detector, leading to energetic upward-going muons pointing back to the center of the Sun or Earth.

Another possibility is that neutralino LSP annihilation in the galactic center (or the halo) could result in high-energy photons from cascade decays of the heavy Standard Model particles that are produced. These photons could be detected in air Cerenkov telescopes or in space-based detectors. There are also interesting possible signatures from neutralino LSP annihilation in the galactic halo producing detectable quantities of high-energy positrons or antiprotons.

More information on these possibilities, and the various experiments that can exploit them can be found in the literature and we will hopefully discuss a few of them.

Direct Detection of Neutralinos: Details

Turning our attention now to dark matter detection, we begin with those experiments which attempt to detect dark matter particles through their elastic scattering with nuclei, including CDMS, XENON, ZEPLIN, EDELWEISS, CRESST CoGeNT, DAMA/LIBRA, COUPP, WARP, and KIMS. This class of techniques is collectively known as direct detection, in contrast to indirect detection efforts which attempt to observe the annihilation products of dark matter particles.

- Kinematics

Envision an incoming χ with momentum p scattering to an outgoing χ with momentum p' off a nucleon target initially at rest via exchange of some mediator carrying momentum $q = p - p'$.

The at rest nucleon has momentum $p_N = (m_N, \vec{0})$ and the final

nucleon's momentum is p'_N . We have

$$p'_N = q + p_N \quad (529)$$

from which

$$m_N^2 = (q + p_N)^2 = q_0^2 - |\vec{q}|^2 + 2q_0m_N + m_N^2, \quad (530)$$

leading to the conclusion that (for low energies as appropriate here)

$$q_0 \ll |\vec{q}|, \quad \Rightarrow \quad |\vec{q}|^2 = 2q_0m_N, \quad \Rightarrow \quad \frac{|\vec{q}|}{m_N} \sim \sqrt{\frac{2q_0}{m_N^2}}. \quad (531)$$

The quantity q_0 is called the recoil energy which we see is given by

$$E_{\text{recoil}} = \frac{|\vec{q}|^2}{2m_N} \quad (532)$$

For $m_X \gg m_N$ and a velocity of ~ 300 km/s, we expect typical recoil energies of $E_{\text{recoil}} \sim m_N v^2 \sim 1\text{-}100$ keV.

In the center of mass frame, and using the (very good) non-relativistic approximation, one writes

$$\begin{aligned} p &= (m_\chi + \frac{1}{2}m_\chi v_\chi^{cm2}, 0, 0, m_\chi v_\chi^{cm}), \\ p' &= (m_\chi + \frac{1}{2}m_\chi v_\chi^{cm2}, m_\chi v_\chi^{cm} \sin \theta, 0, m_\chi v_\chi^{cm} \cos \theta) \end{aligned} \quad (533)$$

so that, using $\sin^2 \theta + (1 - \cos \theta)^2 = 2(1 - \cos \theta)$,

$$q^2 = -2m_\chi^2 v_\chi^{cm2} (1 - \cos \theta). \quad (534)$$

However, we wish to express v_χ^{cm} in terms of the laboratory (N rest

frame) velocity of the χ , v_χ^{lab} . We have, since $v_N^{lab} = 0$,

$$V^{cm}(m_\chi + m_N) = m_\chi v_\chi^{lab}, \quad \Rightarrow v_\chi^{cm} = v_\chi^{lab} - V^{cm} = \frac{m_N}{m_\chi + m_N} v_\chi^{lab}. \quad (535)$$

Using the shorthand notation $v \equiv v_\chi^{lab}$ we then have

$$|\vec{q}|^2 \simeq -q^2 = 2 \frac{m_\chi^2 m_N^2 v^2 (1 - \cos \theta)}{(m_\chi + m_N)^2} \quad (536)$$

yielding

$$E_{recoil} = \frac{|\vec{q}|^2}{2m_N} = \frac{2\mu^2 v^2 (1 - \cos \theta)}{2m_N}, \quad (537)$$

where, to repeat, v is the WIMP velocity in the target rest frame, and μ is the reduced mass, $\mu \equiv m_\chi m_N / (m_\chi + m_N)^2$.

The rate in a detector depends upon the energy/mass density ρ_χ of the WIMPs near the Earth and the velocity distribution $f(v)$ of WIMPs in the Galactic halo near the Earth. As a function of the

energy deposited, Q , direct-detection experiments measure the number of events per day per kilogram of detector material. Qualitatively, this event rate is simply

$$R \simeq n\sigma\langle v\rangle/m_N, \quad (538)$$

where $n = \rho_\chi/m_\chi$ is the WIMP number density, σ is the elastic cross section, and $\langle v\rangle$ is the average speed of the WIMPs relative to the target, and we divide the detector mass M_{det} by the target nucleon mass, m_N , to get the number of target nuclei (per kilogram).

More accurately, one needs to take into account the fact that the WIMPs move in the halo with velocities determined by $f(v)$, that the differential cross section depends upon $f(v)$ through a form factor, $d\sigma/d|\vec{q}|^2 \propto F^2(E_{recoil})$, and that detectors have a threshold energy, E_T , below which they are insensitive to WIMP-nuclear recoils. (In fact, there is an efficiency that is a function of the deposit energy.)

In addition, the earth moves through the Galactic halo and this

motion should be taken into account via $f(v)$.¹⁸ In the end we get (neglecting efficiency variation and a possible nuclear form factor and using $d|\vec{q}|^2 = 2m_N dE_{recoil}$)

$$R \approx \int_{E_{\min}}^{E_{\max}} \int_{v_{\min}}^{v_{\max}} \frac{2\rho_\chi}{m_\chi} \frac{d\sigma}{d|\vec{q}|^2} v f(v) dv dE_{recoil}, \quad (539)$$

where ρ_χ is the dark matter density, σ is the WIMP-nuclei elastic scattering cross section, and $f(v)$ is the velocity distribution of WIMPs. The limits of integration are set by:

1. the galactic escape velocity, $v_{\max} \approx 650$ km/s,

2. by setting $\cos \theta = -1$ in Eq. (537): $v_{\min} = \left(\frac{E_{recoil} m_N}{2\mu^2} \right)^{1/2}$.

3. The minimum energy is set by the energy threshold of the detector,

¹⁸The potential field of the Sun and the Galactic disk will also affect the local ρ_χ and $f(v)$, but these effects are relatively small and not considered in what follows.

which is typically in the range of several keV to several tens of keV.

4. E_{\max} is actually set by v_{\max} through $v = \left(\frac{E_{\text{recoil}} m_N}{2\mu^2} \right)^{1/2}$, yielding $E_{\max} = 2v_{\max}^2 \mu^2 / m_N$.

In practice, the observable rate will also need to include in R the efficiency for seeing a certain value of E_{recoil} , which efficiency declines significantly as E_{recoil} decreases. This is why it is useful to give the general formula as above.

However, it is useful to write down a result valid in the limit where:

- No efficiency factor is included.
- No lower threshold cutoff is imposed — *i.e.* integrate down to $E_{\text{recoil}} = 0$.
- No nuclear form factor is present (as already assumed above).

- Variation of the invariant matrix element contained in $d\sigma/d|\vec{q}|^2$ is neglected — a good approximation so long as m_χ and masses associated with internal propagators are $\gtrsim \text{few GeV}$.

In this limit, we can integrate over $d|\vec{q}|^2$ as follows.

- From PDG handbook or Peskin and Schroeder, we have

$$\frac{d\sigma}{dt} = \frac{1}{64\pi s} \frac{1}{|\vec{p}_1^{cm}|^2} |\mathcal{M}|^2, \quad (540)$$

where $|\vec{p}_1^{cm}| = |\vec{p}_1^{lab}| m_N / \sqrt{s}$ and $|\vec{p}_1^{lab}| = m_\chi v$ implying that $|\vec{p}_1^{cm}| = m_\chi m_N v / \sqrt{s}$.

- Now $t \simeq |\vec{q}|^2$ and if \mathcal{M} is independent of $|\vec{q}|^2$ we can integrate above from $|\vec{q}|^2 = 0$ to $|\vec{q}|^2 = 4\mu^2 v^2$ to obtain

$$\sigma = \frac{1}{64\pi m_N^2 m_\chi^2 v^2} (4\mu^2 v^2) |\mathcal{M}|^2 = \frac{\mu^2}{16\pi m_N^2 m_\chi^2} |\mathcal{M}|^2. \quad (541)$$

- Next, we must recall that the above cross section expression assumes that \mathcal{M} is computed in relativistic state normalization. We will be operating below in non-relativistic state normalization. The relation is $\mathcal{M} = (2m_\chi)(2m_N)\mathcal{M}_{NR}$, yielding

$$\sigma = \frac{\mu^2 v^2}{\pi} |\mathcal{M}_{NR}|^2. \quad (542)$$

WIMPs can potentially scatter with nuclei through both spin-independent and spin-dependent interactions.

- The experimental sensitivity to spin-independent couplings benefits from coherent scattering, which leads to cross sections (and rates) proportional to the square of the atomic mass of the target nuclei.
- The cross sections for spin-dependent scattering, in contrast, are proportional to $J(J+1)$, where J is the spin of the target nucleus, and thus do not benefit from large target nuclei.

- As a result, the current experimental sensitivity to spin-dependent scattering is far below that of spin-independent interactions.

For this reason, we consider mainly the case of spin-independent scattering of WIMPs with nuclei.

- The diagrams of importance for the spin-independent scattering of the neutralino of the MSSM (I will be using χ or $\tilde{\chi}_1^0$ as my notation) are Higgs exchange in the t -channel and \tilde{q} exchange in the s -channel.

Both are “scalar” type interactions. For these the effective interaction Lagrangian can be written (at low energy)

$$\mathcal{L} = \sum_q f_q \bar{\chi} \chi \bar{q} q + b \alpha_s \bar{\chi} \chi G_{\mu\nu}^a G^{a\mu\nu} + \dots \quad (543)$$

where $G_{\mu\nu}^a$ is the gluonic field strength tensor.

- It is then necessary to evaluate things like

$$\langle N | m_q \bar{q}q | N \rangle. \quad (544)$$

- The matrix elements of the light-quark currents are obtained in chiral perturbation theory from the measurements of the pion-nucleon sigma term. For each of $q = u, d, s$, we write

$$\langle N | m_q \bar{q}q | N \rangle = m_N f_T^{(N)} q. \quad (545)$$

We will not worry about the small differences between u and d in the discussion below, but, of course, full analyses take these differences into account.

In this approximation, one defines (in the state normalization conventions employed below)

$$\sigma_{\pi N} = \hat{m} \langle N | \bar{u}u + \bar{d}d | N \rangle, \quad \hat{m} = \frac{1}{2}(m_u + m_d). \quad (546)$$

leading to

$$f_{T u} + f_{T d} = \frac{\hat{m}}{m_N} \sigma_{\pi N}. \quad (547)$$

As we discuss, the determination of the pion-nucleon sigma term from the data is fraught with significant uncertainties, which lead to uncertainties in the parameters $f_{T q}$.

- Let us also define

$$y = \frac{2\langle p | \bar{s}s | p \rangle}{\langle p | \bar{u}u + \bar{d}d | p \rangle}. \quad (548)$$

Then

$$\sigma_{\pi N} = \hat{m} \frac{\langle p | \bar{u}u + \bar{d}d - 2\bar{s}s | p \rangle}{1 - y}, \quad (549)$$

which form is useful because the numerator is proportional to the octet breaking piece in the Hamiltonian.

To see how to evaluate the numerator and then proceed to get values for the $f_{T q}$ requires a bit of work. The steps are below.

1. First, recall the old Feynman-Hellmann theorem stating

$$\frac{\partial E(\lambda)}{\partial \lambda} = \langle \psi(\lambda) | \frac{\partial H(\lambda)}{\partial \lambda} | \psi(\lambda) \rangle \quad (550)$$

where λ is some parameter and $H(\lambda)|\psi(\lambda)\rangle = E(\lambda)|\psi(\lambda)\rangle$ in (NR) normalization $\langle \psi(\lambda) | \psi(\lambda) \rangle = 1$.

2. Here, we want to isolate the dependence of the QCD Hamiltonian on the quark masses by treating

$$H^1 = m_u \bar{u}u + m_d \bar{d}d + m_s \bar{s}s \sim \hat{m}(\bar{u}u + \bar{d}d) + m_s \bar{s}s \quad (551)$$

as a perturbation on some quark-mass independent contribution to nucleon masses.

Then, to first order

$$\frac{\partial m_h(m_q)}{\partial m_q} = \langle h | \frac{\partial H^1}{\partial m_q} | h \rangle. \quad (552)$$

3. Applying to the nucleon, which has mass even in the chiral limit of $\hat{m} \rightarrow 0$, we have

$$\begin{aligned}\frac{\partial m_N}{\partial \hat{m}} &= \langle N | \bar{u}u + \bar{d}d | N \rangle, \\ \frac{\partial \hat{m}}{\partial m_s} &= \langle N | \bar{s}s | N \rangle.\end{aligned}\tag{553}$$

These results imply

$$m_N = A + \hat{m} \langle N | \bar{u}u + \bar{d}d | N \rangle + m_s \langle N | \bar{s}s | N \rangle.\tag{554}$$

Here, A is some constant contribution to the nucleon mass not related to the quark masses.

4. We may proceed similarly with other members of the standard octet.

To first order in $SU(3)$ -breaking and using the sample octet members $p = uud$, $\Sigma^+ = uus$, $\Xi^0 = ssu$, we have, using shorthand

of $\langle p|\bar{q}q|p\rangle \equiv \langle \bar{q}q\rangle$ (we always reference the proton)

$$\begin{aligned}
 m_p &= A + m_u\langle\bar{u}u\rangle + m_d\langle\bar{d}d\rangle + m_s\langle\bar{s}s\rangle \\
 m_n &= A + m_u\langle\bar{d}d\rangle + m_d\langle\bar{u}u\rangle + m_s\langle\bar{s}s\rangle \\
 m_{\Sigma^+} &= A + m_u\langle\bar{u}u\rangle + m_d\langle\bar{s}s\rangle + m_s\langle\bar{d}d\rangle \\
 m_{\Sigma^-} &= A + m_u\langle\bar{s}s\rangle + m_d\langle\bar{u}u\rangle + m_s\langle\bar{d}d\rangle \\
 m_{\Xi^0} &= A + m_u\langle\bar{d}d\rangle + m_d\langle\bar{s}s\rangle + m_s\langle\bar{u}u\rangle \\
 m_{\Xi^-} &= A + m_u\langle\bar{s}s\rangle + m_d\langle\bar{d}d\rangle + m_s\langle\bar{u}u\rangle.
 \end{aligned} \tag{555}$$

As a result, we have

$$\begin{aligned}
 \Delta m &\equiv \frac{1}{2}(m_{\Sigma^+}m_{\Sigma^-}) + \frac{1}{2}(m_{\Xi^0} + m_{\Xi^-}) - (m_p + m_n) \\
 &= (m_s - \hat{m})\langle p|\bar{u}u + \bar{d}d - 2\bar{s}s|p\rangle.
 \end{aligned} \tag{556}$$

To good approximation, we can compute the left-hand side as

$$\Delta m = m_{\Sigma^+} + m_{\Xi^0} - 2m_p \simeq (1190 + 1315 - 2 \cdot 939) \text{ MeV} \simeq 627 \text{ MeV}. \quad (557)$$

5. We also need the ratio of the quark masses themselves. These can be extracted from the meson masses when the mesons are considered as goldstone bosons. One finds (no time to derive)

$$m_{\pi^+}^2 = 2\hat{m}B, \quad m_{K^+}^2 = (m_u + m_s)B, \quad m_{K^0}^2 = (m_d + m_s)B, \quad (558)$$

where $B = |\langle 0 | \bar{u}u | 0 \rangle| / f_\pi^2$, where f_π is called the pion decay constant. Thus, we can write

$$\frac{m_s}{\hat{m}} \simeq \frac{m_{K^0}^2 + m_{K^+}^2}{m_{\pi^+}^2} - 1 \simeq 25. \quad (559)$$

Lattice calculations yield a fairly similar result.

For later reference, we note that the values of $|\langle 0|\bar{u}u|0\rangle|$ and f_π can be gotten from experiment and theory in order to obtain the (current) quark masses themselves with the results:

$$m_u = 4.2 \text{ MeV}, \quad m_d = 7.5 \text{ MeV}, \quad m_s = 150 \text{ MeV}. \quad (560)$$

6. Altogether, we obtain

$$\begin{aligned} \sigma_{\pi N} &= \frac{\hat{m}}{1-y} \langle p|\bar{u}u + \bar{d}d - 2\bar{s}s|p\rangle \\ &= \frac{\hat{m}}{(m_s - \hat{m})(1-y)} \Delta m \\ &\simeq \frac{1}{24} \frac{637 \text{ MeV}}{1-y} \simeq \frac{26 \text{ MeV}}{1-y}. \end{aligned} \quad (561)$$

Higher order terms of order $m_q^{3/2}$ and m_q^2 are argued to increase the 26 MeV to perhaps as much as 35 MeV, the value employed in arXiv:0801.3656.

Below, we will use as a proxy

$$y = 1 - \frac{35 \text{ MeV}}{\sigma_{\pi N}}. \quad (562)$$

The bottom line is that if $\sigma_{\pi N}$ can be determined from data or lattice calculations, not only can we determine the u, d content of the N , we can also determine y , which gives the strangeness content of the nucleon.

7. We could go beyond the isospin average.

First, note that if there were no sea quarks, then $\langle \bar{u}u \rangle = 2\langle \bar{d}d \rangle$ and $\langle \bar{s}s \rangle = 0$ would be the expectation. However, using observed octet masses and the computed quark masses,

$$\begin{aligned} \langle \bar{u}u \rangle - \langle \bar{s}s \rangle &= \frac{m_{\Xi^0} + m_{\Xi^-} - m_p - m_n}{2m_s - m_u - m_d} = 2.63 \\ \langle \bar{d}d \rangle - \langle \bar{s}s \rangle &= \frac{m_{\Sigma^+} + m_{\Sigma^-} - m_p - m_n}{2m_s - m_u - m_d} = 1.77, \end{aligned} \quad (563)$$

from which we conclude $\langle \bar{u}u \rangle / \langle \bar{d}d \rangle \simeq 1.5$ if $\langle \bar{s}s \rangle = 0$, implying the need for sea quarks of all types.

Writing

$$\langle \bar{u}u \rangle = 2V + S, \quad \langle \bar{d}d \rangle = V + S, \quad \langle \bar{s}s \rangle = S', \quad (564)$$

and substituting above, one finds $V = 0.86$ and $S - S' = 0.91$.

This indicates a large sea quark content of the proton and neutron with more $\bar{u}u$ and $\bar{d}d$ sea pairs than $\bar{s}s$ pairs.

Further, if $\langle \bar{s}s \rangle = 0$, then $y = 0$ and Eq. (561) would imply $\sigma_{\pi N} = 26$ MeV as compared to the larger experimental values discussed below.

8. The determination of $\sigma_{\pi N}$ usually begins with what is called the Cheng-Dashen theorem, according to which

$$\Sigma = \sigma_{\pi N} \left[1 + \mathcal{O}(m_q^{1/2}) + \dots \right] \quad (565)$$

where Σ is determined by an analytic extrapolation of the scattering amplitude for $\pi N \rightarrow \pi N$ to an unphysical point in energy and momentum-transfer. This is a long story with the result that

$$\Sigma = (56 - 74) \text{ MeV} \quad (566)$$

with the corrections to the $[1 + \dots]$ indicated above subtracting perhaps as much as 16 MeV or as little as 4 MeV leaving

$$\sigma_{\pi N} \sim (40 - 70) \text{ MeV}. \quad (567)$$

This uncertainty in the pion-nucleon sigma term is the largest source of uncertainty for getting all the f_{Tq} .

Higher orders in chiral perturbation theory lead to small corrections.

9. Commonly used values are $\sigma_{\pi N} = 60 \text{ MeV}$ and 45 MeV , leading to $y = 1 - \frac{35 \text{ MeV}}{\sigma_{\pi N}} = \frac{5}{12} \simeq 0.42$ and $\frac{2}{9} \simeq 0.22$, respectively.

For smaller $\sigma_{\pi N}$, both the s content and the u, d content of the proton is smaller!

10. Taking the value of $\sigma_{\pi N} = 55 \text{ MeV}$, we can combine our various equations to conclude the following:

$$\begin{aligned}
 m_u \langle \bar{u}u \rangle &= 21.5 \text{ MeV}, & m_d \langle \bar{d}d \rangle &= 32.1 \text{ MeV}, & m_s \langle \bar{s}s \rangle &= 376.2 \text{ MeV}, \\
 \frac{\langle \bar{u}u \rangle}{\langle \bar{u}u \rangle + \langle \bar{d}d \rangle + \langle \bar{s}s \rangle} &= 0.43, & \frac{\langle \bar{u}u \rangle}{\langle \bar{u}u \rangle + \langle \bar{d}d \rangle + \langle \bar{s}s \rangle} &= 0.36, & \frac{\langle \bar{s}s \rangle}{\langle \bar{u}u \rangle + \langle \bar{d}d \rangle + \langle \bar{s}s \rangle} &= 0.21,
 \end{aligned}
 \tag{568}$$

leading to

$$S : S' : V = 3.9 : 2.9 : 1.
 \tag{569}$$

Summing the top line of Eq. (568), we see that the quark masses only account for 430 MeV of the 939 MeV mass of the proton. The rest comes from the A term, which is related to the gluon component of the proton mass by

$$A = \left\langle -9 \frac{\alpha_s}{8\pi} GG \right\rangle,
 \tag{570}$$

which includes the heavy quark anomaly contributions. See below.

11. A very systematic survey of all the possibilities for the πN sigma term and the implications for dark matter appeared in arXiv:0801.3656.

- We can now use the above to determine values for the f_{Tq} 's.

We obtain an average value for $f_{T u,d}$ (neglecting as we did above splitting between u and d) of

$$\begin{aligned} \frac{1}{2}(f_{T u} + f_{T d}) &\sim \frac{1}{2} \left(\frac{\hat{m}}{m_p} \langle p | \bar{u}u + \bar{d}d | p \rangle \right) = \frac{1}{2} \left(\frac{\sigma_{\pi N}}{939 \text{ MeV}} \right) \\ &\sim \frac{1}{2} \left(\frac{(60, 45) \text{ MeV}}{939 \text{ MeV}} \right) \sim (0.032, 0.024). \quad (571) \end{aligned}$$

For the strange quark, we have

$$f_{T s}^{(p)} \equiv \frac{m_s}{m_p} \langle p | \bar{s}s | p \rangle$$

$$\begin{aligned}
&= \frac{1 m_s}{2 m_p} \langle p | \bar{u}u + \bar{d}d | p \rangle \frac{2 \langle p | \bar{s}s | p \rangle}{\langle p | \bar{u}u + \bar{d}d | p \rangle} \\
&= \frac{1 m_s}{2 m_p} \frac{\sigma_{\pi N}}{\hat{m}} y \\
&= 12.5 \frac{\sigma_{\pi N}}{m_p} y \\
&= 12.5 \times (0.032, 0.024) \times \left(\frac{5}{12}, \frac{2}{9} \right) \\
&= (0.167, 0.067) . \tag{572}
\end{aligned}$$

- An old table from Jungman et al (Physics Reports, 267, p195) is:

Table 6

Estimates for the nucleon parameters f_{Tq} . The u- and d-quark values are obtained from Ref. [284], and we list values reported for f_{Tq} by several authors

Nucleon	f_{T_u}	f_{T_d}	f_{T_s} [285]	F_{T_s} [283, 284]	F_{T_s} [288]
n	0.023	0.034	0.14	0.46	0.08
p	0.019	0.041	0.14	0.46	0.08

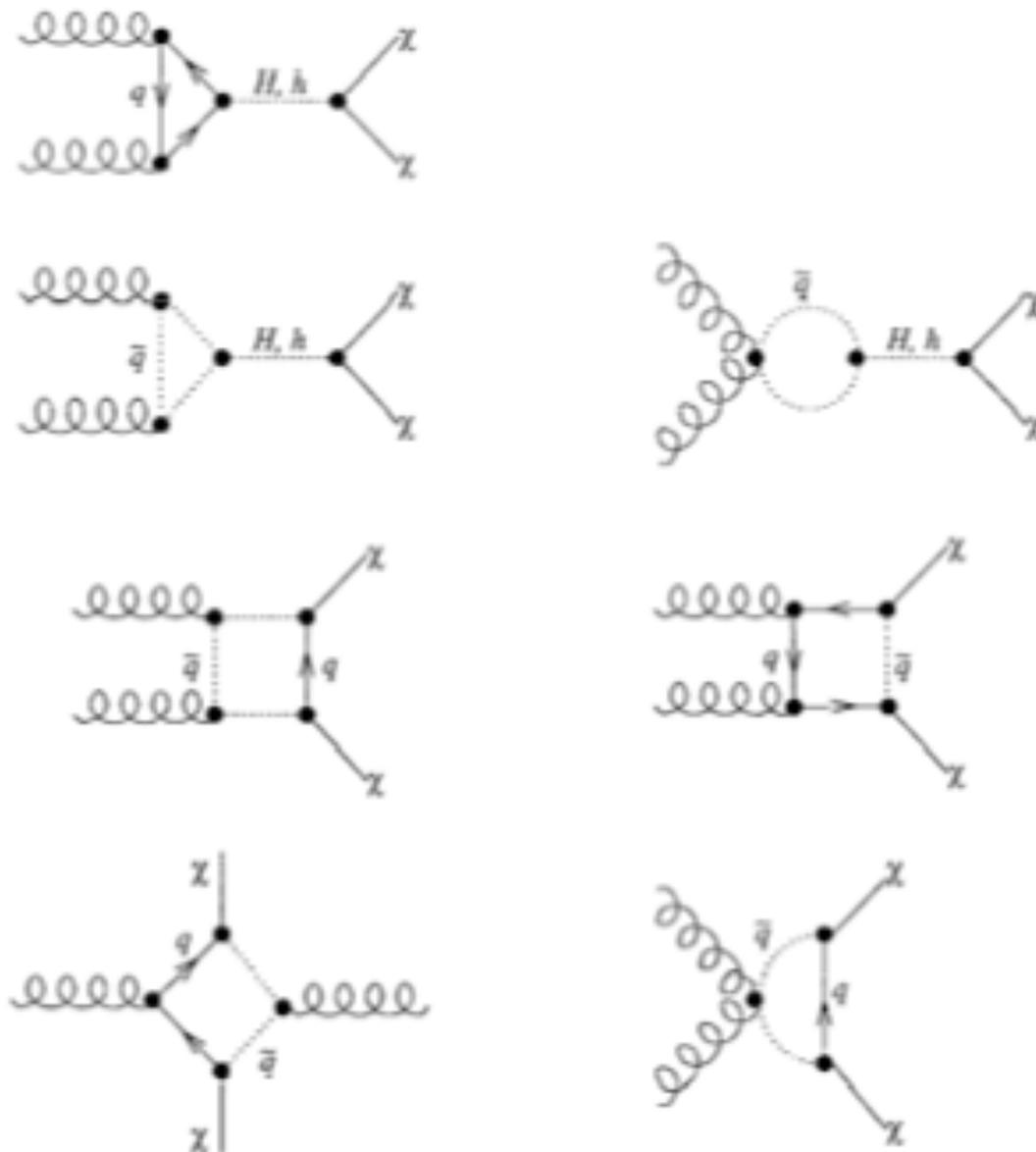
In arXiv:0801.3656, they conclude that reasonable ranges for $\sigma_{\pi N}$ and $(1 - y)\sigma_{\pi N} \equiv \sigma_0$ are: $\sigma_{\pi N} = (64 \pm 8)$ MeV and $\sigma_0 = (36 \pm 7)$ MeV,

yielding results for the f_{Tq} that are near those for our $\sigma_{\pi N} = 60 \text{ MeV}$ and $\sigma_0 = 35 \text{ MeV}$ choices.

However, they also conclude that if one allows for $48 \text{ MeV} < \sigma_{\pi N} < 80 \text{ MeV}$ (the 2σ error range for the above) keeping σ_0 fixed, then the spin-independent dark matter scattering cross section can vary by almost an order of magnitude!

This is a warning that unless the errors on $\sigma_{\pi N}$ and σ_0 can be greatly improved, precision checking of the dark matter scattering cross section against some collider measurements of dark matter properties (esp. couplings) may not be possible.

- Heavy quarks contribute to the mass of the nucleon through the anomaly. This is equivalent to the set of diagrams shown below containing a $Q = c, b, t$.



For heavy squarks only the first diagram is significant. It can be

evaluated using the heavy quark expansion, according to which the following substitution can be made for each of $Q = c, t, b$ in a nucleon matrix element:

$$m_Q \bar{Q}Q \rightarrow -\frac{2\alpha_s}{24\pi} GG. \quad (573)$$

The trace of the QCD energy-momentum tensor can then be written

$$\begin{aligned} \theta_{\mu}^{\mu} &= m_u \bar{u}u + m_d \bar{d}d + m_s \bar{s}s + \sum_{Q=b,c,t} m_Q \bar{Q}Q - \frac{7\alpha_s}{8\pi} GG \\ &= m_u \bar{u}u + m_d \bar{d}d + m_s \bar{s}s - \frac{9\alpha_s}{8\pi} GG. \end{aligned} \quad (574)$$

The matrix element of this quantity is the nucleon mass,

$$m_N = \sum_{q=u,d,s} \langle N | m_q \bar{q}q | N \rangle + \langle N | -\frac{9\alpha_s}{8\pi} GG | N \rangle \quad (575)$$

This can be rewritten as

$$\begin{aligned}
 m_N &= \sum_{u,d,s} m_N f_{Tq} + \frac{27}{2} \left(-\frac{2\alpha_s}{24\pi} \langle N | GG | N \rangle \right) \\
 &= \sum_{u,d,s} m_N f_{Tq} + \frac{27}{2} \langle N | m_Q \bar{Q}Q | N \rangle
 \end{aligned} \tag{576}$$

for any given Q , so that for each of the heavy quarks we find

$$\langle N | m_Q \bar{Q}Q | N \rangle = \frac{2}{27} m_N \left[1 - \sum_{q=u,d,s} f_{Tq}^{(N)} \right] \equiv \frac{2}{27} m_N f_{TG}, \tag{577}$$

which defines f_{TG} . By this definition,

$$f_{TG} = 1 - \sum_{q=u,d,s} f_{Tq}. \tag{578}$$

Note that relatively large values of f_{TG} are typical.

- The effective Lagrangian given earlier suggests that the elastic-scattering cross section should generally increase as the strangeness content of the nucleon is increased.

Meanwhile, the coupling to heavy quarks is maximized if the strangeness content is zero.

But, as we increase the strangeness content, the coupling to the strange-quark scalar density more than makes up for the decrease in the heavy-quark coupling.

- In principle, one should bring in the squark loop stuff. We have no time for this. We will work in the approximation that the squarks are very heavy, in which case the previously shown diagrams involving \tilde{q} 's are suppressed by $1/m_{\tilde{q}}^2$.

In the limit of heavy squark masses, we obtain (with $\langle \chi | \bar{\chi} \chi | \chi \rangle = 1$

as in NR normalization)

$$\begin{aligned}
 f_p &\equiv \langle \chi p | \mathcal{L} | \chi p \rangle \\
 &= \langle \chi p | \sum_q f_q \bar{\chi} \chi \bar{q} q | \chi p \rangle \\
 &= \sum_q \frac{m_p}{m_q} f_q \frac{\langle p | m_q \bar{q} q | p \rangle}{m_p} \\
 &= \sum_{q=u,d,s} \frac{m_p}{m_q} f_q f_{Tq} + \sum_{q=c,t,b} \frac{m_p}{m_q} f_q \frac{2}{27} f_{TG}. \quad (579)
 \end{aligned}$$

Summarizing for p, n we have:

$$\frac{f_{p,n}}{m_{p,n}} \simeq \sum_{q=u,d,s} \frac{f_{Tq} f_q}{m_q} + \frac{2}{27} f_{TG} \sum_{q=c,b,t} \frac{f_q}{m_q}. \quad (580)$$

More recent work suggests and $f_{T_u}^{(p)} \approx 0.020 \pm 0.004$, $f_{T_d}^{(p)} \approx 0.026 \pm 0.005$, $f_{T_s}^{(p)} \approx 0.118 \pm 0.062$, $f_{T_u}^{(n)} \approx 0.014 \pm 0.003$, $f_{T_d}^{(n)} \approx 0.036 \pm$

0.008, $f_{T_s}^{(n)} \approx 0.118 \pm 0.062$. And, as we have seen, $f_{TG}^{(p)}$ is given by $1 - f_{T_u}^{(p)} - f_{T_d}^{(p)} - f_{T_s}^{(p)} \approx 0.84$, and analogously, $f_{TG}^{(n)} \approx 0.83$.

- For a nuclear target, the full \mathcal{M}_{NR} for spin-independent WIMP-nucleus elastic scattering is then

$$\mathcal{M}_{NR} = Z f_p + (A - Z) f_n \quad (581)$$

and the cross section (integrated over $|\vec{q}|^2$, neglecting any nucleon form factor and assuming detection efficiency independent of E_{recoil} all the way down to $E_{\text{recoil}} = 0$) is then given by

$$\sigma \approx \frac{4\mu^2}{\pi} [Z f_p + (A - Z) f_n]^2, \quad (582)$$

where Z and A are the atomic number and atomic mass of the nucleus. f_p and f_n are the WIMP's couplings to protons and

neutrons, given earlier

$$f_{p,n} = \sum_{q=u,d,s} f_{T_q}^{(p,n)} f_q \frac{m_{p,n}}{m_q} + \frac{2}{27} f_{TG}^{(p,n)} \sum_{q=c,b,t} f_q \frac{m_{p,n}}{m_q}, \quad (583)$$

where f_q are the WIMP-quark couplings. As we have seen, the first term in Eq. (583) corresponds to interactions with the quarks in the target nuclei. The second term corresponds to interactions with the gluons in the target through colored loop diagrams containing a heavy quark.

Cross Section Details

Neutralinos can elastically scatter with quarks through either t -channel CP-even Higgs exchange, or s -channel squark exchange:

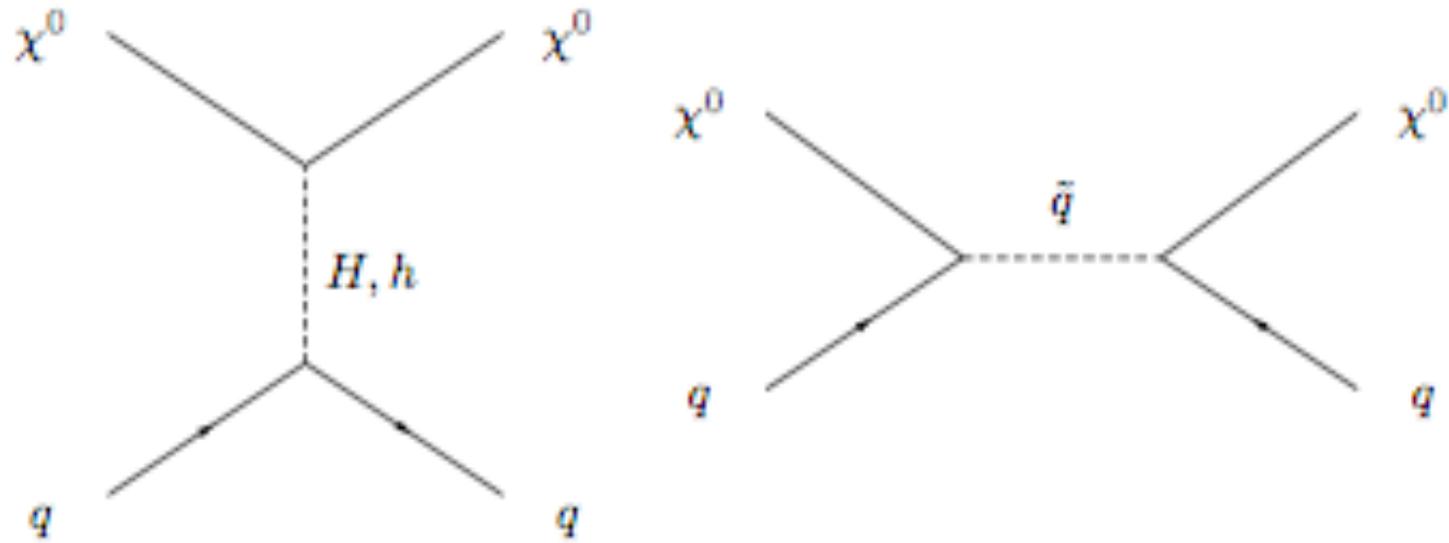


Figure 36: Diagrams for $\tilde{\chi}_1^0$ scattering off a quark.

In addition to these diagrams, we can write analogous processes in which the WIMP couples to gluons in the target through a quark/squark loop. By calculating the WIMP-quark couplings, f_q , we can also implicitly include the interactions of neutralinos with gluons in the target nuclei as well (see Eq. (583)).

The neutralino-quark coupling, in which all of the supersymmetry model-dependent information is contained, is given by

$$\begin{aligned}
 f_q = & -\frac{1}{2(m_{1i}^2 - m_\chi^2)} \text{Re} [(X_i) (Y_i)^*] - \frac{1}{2(m_{2i}^2 - m_\chi^2)} \text{Re} [(W_i) (V_i)^*] \\
 & - \frac{g_2 m_q}{4m_W B} \left[\text{Re} (\delta_1 [g_2 N_{12} - g_1 N_{11}]) DC \left(-\frac{1}{m_H^2} + \frac{1}{m_h^2} \right) \right. \\
 & \left. + \text{Re} (\delta_2 [g_2 N_{12} - g_1 N_{11}]) \left(\frac{D^2}{m_h^2} + \frac{C^2}{m_H^2} \right) \right], \tag{584}
 \end{aligned}$$

where

$$\begin{aligned}
 X_i & \equiv \eta_{11}^* \frac{g_2 m_q N_{1,5-i}^*}{2m_W B} - \eta_{12}^* e_i g_1 N_{11}^*, \\
 Y_i & \equiv \eta_{11}^* \left(\frac{y_i}{2} g_1 N_{11} + g_2 T_{3i} N_{12} \right) + \eta_{12}^* \frac{g_2 m_q N_{1,5-i}^*}{2m_W B},
 \end{aligned}$$

$$\begin{aligned}
W_i &\equiv \eta_{21}^* \frac{g_2 m_q N_{1,5-i}^*}{2m_W B} - \eta_{22}^* e_i g_1 N_{11}^*, \\
V_i &\equiv \eta_{22}^* \frac{g_2 m_q N_{1,5-i}^*}{2m_W B} + \eta_{21}^* \left(\frac{y_i}{2} g_1 N_{11} + g_2 T_{3i} N_{12} \right), \quad (585)
\end{aligned}$$

where throughout $i = 1$ for up-type quarks and $i = 2$ for down type quarks. m_{1i}, m_{2i} denote the squark mass eigenvalues and η is the matrix which diagonalizes the squark mass matrices, $\text{diag}(m_1^2, m_2^2) = \eta M^2 \eta^{-1}$. y_i , T_{3i} and e_i denote hypercharge, isospin and electric charge of the quarks. For scattering off of up-type quarks

$$\delta_1 = N_{13}, \quad \delta_2 = N_{14}, \quad B = \sin \beta, \quad C = \sin \alpha, \quad D = \cos \alpha, \quad (586)$$

whereas for down-type quarks

$$\delta_1 = N_{14}, \quad \delta_2 = -N_{13}, \quad B = \cos \beta, \quad C = \cos \alpha, \quad D = -\sin \alpha \quad (587)$$

The quantity α is the angle that diagonalizes the CP-even Higgs

mass matrix.

The first two terms in Eq. (584) correspond to interactions through the exchange of a squark, while the final term is generated through Higgs exchange.

To help develop some intuition for what size neutralino-nucleon cross sections we might expect, let's consider a few simple limits.

1. First, there is the case in which squarks are heavy and $\tan\beta$ is large. In this case the scattering is dominated by heavy Higgs (H) exchange through its couplings to strange and bottom quarks. For moderate to large m_H , Eq. (436) implies $m_A \sim m_H \sim m_{H^\pm}$. If $\tan\beta$ is large, then $\cos\beta$ is small and $\sin\beta \sim 1$. Further from Eq. (438) we find that $\cos\alpha \approx 1$.

In this case, the leading contribution to the neutralino-nucleon cross section derives from the b and s (down-type) contributions proportional to $\frac{1}{B} \frac{C^2}{m_H^2} \sim \frac{\cos^2\alpha}{\cos\beta m_H^2}$. Assuming a dominantly bino

neutralino, *i.e.* large N_{11} , we find

$$f_{s,b} = \frac{g_1 g_2 m_{s,b} \text{Re}(N_{13} N_{11}) \cos^2 \alpha}{4m_W \cos \beta m_H^2}. \quad (588)$$

Substituting into Eq. (583), we find

$$f_{p,n} = \frac{g_1 g_2 m_{p,n} \text{Re}(N_{13} N_{11}) \cos^2 \alpha}{4m_W \cos \beta m_H^2} \left(f_{T_s} + \frac{2}{27} f_{TG} \right), \quad (589)$$

which is then substituted into Eq. (582) to obtain (using $\mu^2 \simeq m_{p,n}^2$ for $m_\chi \gg m_{p,n}$ in the single nucleon target case and writing just N for p, n)

$$\sigma_{\chi N} \sim \frac{g_1^2 g_2^2 |\text{Re}(N_{11} N_{13})|^2 m_N^4}{4\pi m_W^2 \cos^2 \beta m_H^4} \left(f_{T_s} + \frac{2}{27} f_{TG} \right)^2, \quad (m_{\tilde{q}} \text{ large, } \cos \alpha \approx 1). \quad (590)$$

Here, the cross section scales with m_H^{-4} and with $\tan^2 \beta$, and very large rates are possible.

For $m_\chi \sim 100$ GeV, for example, $m_H \sim 200$ GeV and moderate to large $\tan\beta$ leads to cross sections with nucleons on the order of 10^{-5} to 10^{-7} pb for $|\mu| \sim 200$ GeV, or 10^{-7} to 10^{-9} pb for $|\mu| \sim 1$ TeV. The dependence on $|\mu|$ arises because N_{14} decreases with increasing $|\mu|$ – *i.e.* as the χ becomes more and more purely bino.

Tracking down the couplings

We did not end up having time to work out all the couplings intrinsically contained in Eq. (584), but let's look at the main ones contributing in this specific case.

- (a) First, there is the $H^0 b\bar{b}$ coupling associated with the “bottom” vertex of $\chi b \rightarrow \chi b$ scattering via H^0 exchange.

Recall that the bottom quark mass comes from $\mathcal{L} \ni -y_b H_d^0 b\bar{b}$ (the minus sign is standard for the Lagrangian) when H_d^0 is

replaced by $\langle H_d^0 \rangle = v_d$ which leads to

$$y_b = \frac{m_b}{v_d} = \frac{m_b}{v \cos \beta} = \frac{m_b g_2}{\sqrt{2} m_W \cos \beta}, \quad (591)$$

after using $m_W = g_2 v / \sqrt{2}$ — cf Eq. (422). Now, referring back to Eq. (433) we find that the Langrangian state H_d^0 is written in terms of the mass eigenstates as

$$H_d^0 = v_d + \frac{1}{\sqrt{2}}(-h^0 \sin \alpha + H^0 \cos \alpha + iA^0 \sin \beta). \quad (592)$$

For later reference we note that

$$H_u^0 = v_u + \frac{1}{\sqrt{2}}(h^0 \cos \alpha + H^0 \sin \alpha + iA^0 \cos \beta). \quad (593)$$

Thus, for the CP-even Higgs bosons we find

$$\mathcal{L} \ni -y_b H_d^0 b \bar{b} \rightarrow -\frac{m_b g_2}{\sqrt{2} m_W \cos \beta} \frac{1}{\sqrt{2}} (-h^0 \sin \alpha + H^0 \cos \alpha) b \bar{b} \quad (594)$$

implying an $H^0 b \bar{b}$ term in the Lagrangian of

$$\mathcal{L} \ni -\frac{m_b g_2 \cos \alpha}{2 m_W \cos \beta} H^0 b \bar{b} \quad (595)$$

Looking at f_b in Eq. (588), we find this factor, but we have yet to convert the \mathcal{L} expression to a Feynman rule for the vertex (see later).

(b) Now, what about the $H^0 \chi \bar{\chi}$ vertex at the “top” of the H^0 exchange diagram.

It is useful to first think “intuitively” about how a vertex for a χ that is mainly bino arises.

First, because of the structure of supersymmetry and dimensional

analysis, there is no Higgs-bino-bino vertex. This is because such a vertex would have to arise from a Higgs-Higgs-bino-bino coupling with one Higgs acquiring a VEV (in analogy to the manner in which a Higgs- B - B coupling arises). However, a Higgs-Higgs-bino-bino term in the Lagrangian would have dimension 5 (rather than dimension 4) and thus would be non-renormalizable. This same argument applies to show that the Higgs-Higgs-bino-neutral-wino coupling cannot exist. This means that if the χ were absolutely pure bino, or only a mixture of bino and neutral wino, then it would not couple to H^0 .

So, we must make use of the (small in the limit being considered) components of the χ that are not bino or wino? Recall that we have trilinear couplings in the SM of Higgs-Higgs- B form that just come from expanding the covariant derivative $(D_\mu H_d)^\dagger (D^\mu H_d)$ or H_u version thereof, where we keep a derivative on one H and multiply the other H by the B field contained in D_μ . The

supersymmetric analogue of such a coupling is a Higgs-Higgsino-bino coupling (note that the product of the fields has dimension 4).

More explicitly, we may search through the different types of interaction vertices that were found as part of the earlier discussion regarding interactions in the MSSM. We find the following.

- i. First, no explicit gaugino-gaugino-Higgs interaction was found, as anticipated above.
- ii. Second, we found a vertex of form gaugino-higgsino-scalar, where the scalar is a Higgs. These were the terms of form

$$\mathcal{L} \ni -\sqrt{2}g(\phi_i^* T_{ij}^a \psi_j) \lambda^a + h.c. \quad (596)$$

appearing in Eq. (356).

This interaction gives a contribution to the $H^0 \chi \bar{\chi}$ vertex when one of the χ or $\bar{\chi}$ has a gaugino component and the other has

a Higgsino component. We will work this out in detail in a moment.

- iii. Thirdly, we note that a susy-fermion — susy-fermion — scalar interaction could also have potentially arisen from the

$$-\frac{1}{2}y_{ijk}\phi_i\psi_j\psi_k + h.c. \quad (597)$$

structure appearing in Eq. (354) originating from Eq. (336). However, all our MSSM Yukawa's (assuming R -parity) contain at least one quark field and will not contribute to Higgs- $\chi\bar{\chi}$ interactions.

So, now let us work out the Eq. (596) piece in more detail. We have, including both the $SU(2)$ and $U(1)$ group components:

$$\mathcal{L} \ni -\sqrt{2}g_2\frac{\tau_{ij}^1}{2} \left[H_d^{i*}\psi_{H_d}^j + H_u^{i*}\psi_{H_u}^j \right] \lambda^1$$

$$\begin{aligned}
& -\sqrt{2}g_2\frac{\tau_{ij}^2}{2}\left[H_d^{i*}\psi_{H_d}^j + H_u^{i*}\psi_{H_u}^j\right]\lambda^2 \\
& -\sqrt{2}g_2\frac{\tau_{ij}^3}{2}\left[H_d^{i*}\psi_{H_d}^j + H_u^{i*}\psi_{H_u}^j\right]\lambda^3 \\
& -\sqrt{2}g_1\frac{\delta_{ij}}{2}\left[-H_d^{i*}\psi_{H_d}^j + H_u^{i*}\psi_{H_u}^j\right]\lambda' \\
& + \text{h.c.}
\end{aligned} \tag{598}$$

Now, above you see a structure of the form

$$\begin{aligned}
\frac{1}{\sqrt{2}}(\tau^1\lambda^1 + \tau^2\lambda^2) &= \frac{\tau^1 + i\tau^2}{2}\frac{\lambda^1 - i\lambda^2}{\sqrt{2}} + \frac{\tau^1 - i\tau^2}{2}\frac{\lambda^1 + i\lambda^2}{\sqrt{2}} \\
&= \tau^+\lambda^+ + \tau^-\lambda^-,
\end{aligned} \tag{599}$$

where τ^+ raises index 2 to 1 and τ^- lowers index 1 to 2. Substituting into Eq. (598) we have

$$\mathcal{L} \ni -g_2\left(\lambda^+H_d^{1*}\psi_{H_d}^2 + H_d^{2*}\lambda^-\psi_{H_d}^1 + \lambda^+H_u^{1*}\psi_{H_u}^2 + H_d^{2*}\lambda^-\psi_{H_u}^1\right)$$

$$\begin{aligned}
& -\frac{1}{\sqrt{2}} \left[(g_2 \tau_{ij}^3 \lambda^3 - g_1 \delta_{ij} \lambda') \psi_{H_d}^j H_d^{i*} + (g_2 \tau_{ij}^3 \lambda^3 + g_1 \delta_{ij} \lambda') \psi_{H_u}^j H_u^{j*} \right] \\
+ & \quad h.c. \tag{600}
\end{aligned}$$

The λ^+ and λ^- terms obviously relate to charge changing interactions which are not of interest to us for the $H^0 \chi \bar{\chi}$ interactions. So we focus on the terms of the 2nd line above, dropping i, j choices involving charged Higgs fields. The terms with neutral Higgs fields we then rewrite showing explicit charges of the fields

$$\begin{aligned}
& -\frac{1}{\sqrt{2}} \left[(g_2 \tau_{11}^3 \lambda^3 - g_1 \lambda') \psi_{H_d^1} H_d^{1*} + (g_2 \tau_{22}^3 \lambda^3 + g_1 \lambda') \psi_{H_u^2} H_u^{2*} \right] + h.c. \\
= & -\frac{1}{\sqrt{2}} \left[(g_2 (+1) \lambda^3 - g_1 \lambda') \psi_{H_d^0} H_d^{0*} + (g_2 (-1) \lambda^3 + g_1 \lambda') \psi_{H_u^0} H_u^{0*} \right] \\
& + h.c. \tag{601}
\end{aligned}$$

Now, we go to four component fields using

$$\begin{aligned} \tilde{B} &= \begin{pmatrix} \lambda' \\ \lambda' \\ \lambda' \end{pmatrix} & \tilde{W}_3 &= \begin{pmatrix} \lambda^3 \\ \lambda^3 \\ \lambda^3 \end{pmatrix} \\ \tilde{H}_d^0 &= \begin{pmatrix} \psi_{H_d^0} \\ \overline{\psi}_{H_d^0} \end{pmatrix} & \tilde{H}_u^0 &= \begin{pmatrix} \psi_{H_u^0} \\ \overline{\psi}_{H_u^0} \end{pmatrix} \end{aligned} \quad (602)$$

Now writing $\chi_i^0 = N_{ij}\psi_j^0$, where N is the diagonalizing matrix and the χ_i^0 are the (still 2-component) mass eigenstates, we invert to write $\psi_j^0 = N_{ji}^\dagger \chi_i^0 = \chi_i^0 N_{ij}^*$ and use

$$\begin{aligned} \lambda' &= P_L \tilde{B} = P_L (N_{i1}^* \tilde{\chi}_i^0) \\ \lambda^3 &= P_L \tilde{W}_3 = P_L (N_{i2}^* \tilde{\chi}_i^0) \\ \psi_{H_d^0} &= P_L \tilde{H}_d^0 = P_L (N_{i3}^* \tilde{\chi}_i^0) \\ \psi_{H_u^0} &= P_L \tilde{H}_u^0 = P_L (N_{i4}^* \tilde{\chi}_i^0) \end{aligned} \quad (603)$$

where $\tilde{\chi}_i^0 = \begin{pmatrix} \chi_i^0 \\ \bar{\chi}_i^0 \end{pmatrix}$ is the Dirac spinor mass eigenstate. We next insert all this into our earlier form of Eq. (604) to obtain

$$\begin{aligned}
 & - \frac{1}{\sqrt{2}} \left[(g_2 P_L(N_{i2}^* \tilde{\chi}_i^0) - g_1 P_L(N_{i1}^* \tilde{\chi}_i^0)) P_L(N_{j3}^* \tilde{\chi}_j^0) H_d^{0*} \right. \\
 & \quad \left. + (-g_2 P_L(N_{i2}^* \tilde{\chi}_i^0) + g_1 P_L(N_{i1}^* \tilde{\chi}_i^0)) P_L(N_{j4}^* \tilde{\chi}_j^0) H_u^{0*} \right] \\
 & \quad + h.c. \tag{604}
 \end{aligned}$$

where i and j are summed over. Now, we write (as before)

$$\begin{aligned}
 H_d^0 &= v_d + \frac{1}{\sqrt{2}} (H^0 \cos \alpha - h^0 \sin \alpha + i A^0 \sin \beta) \\
 H_u^0 &= v_u + \frac{1}{\sqrt{2}} (H^0 \sin \alpha + h^0 \cos \alpha + i A^0 \cos \beta) \tag{605}
 \end{aligned}$$

and pick off the terms multiplying $H^0 \tilde{\chi}_1^0 \tilde{\chi}_1^0$ to obtain a piece of

\mathcal{L} of form

$$\mathcal{L} \ni -\frac{1}{2}H^0 P_L \tilde{\chi}_1^0 P_L \tilde{\chi}_1^0 \left[(g_2 N_{12}^* N_{13}^* - g_1 N_{11}^* N_{13}^*) \cos \alpha + (-g_2 N_{12}^* N_{14}^* + g_1 N_{11}^* N_{14}^*) \sin \alpha \right] + h.c. \quad (606)$$

Now, $P_L \tilde{\chi}_1^0 P_L \tilde{\chi}_1^0 = \chi_1^0 \chi_1^0 = \overline{\tilde{\chi}_1^0} P_L \tilde{\chi}_1^0$, where χ_1^0 is just the two-component spinor. Further, the *h.c.* part just gives you the same thing without the stars and with $P_L \rightarrow P_R$. Next, we note that our earlier expression is making use only of the scalar (non- γ_5) part of this vertex, in which case we can use $P_L \rightarrow \frac{1}{2}$ and $P_R = \frac{1}{2}$ and write

$$\mathcal{L} \ni -\frac{1}{2}H^0 \overline{\tilde{\chi}_1^0} \tilde{\chi}_1^0 \left[\text{Re} (g_2 N_{12} N_{13} - g_1 N_{11} N_{13}) \cos \alpha + \text{Re} (-g_2 N_{12} N_{14} + g_1 N_{11} N_{14}) \sin \alpha \right]. \quad (607)$$

So, we now isolate the $\cos \alpha$ piece, neglecting the (small in the limit being considered) $\sin \alpha$ piece, and assume that the $\tilde{\chi}_1^0$ is mainly bino so that, in particular $|N_{11}| \gg |N_{12}|$, in which case we obtain the approximate result

$$\mathcal{L} \ni -\frac{1}{2} H^0 \overline{\tilde{\chi}_1^0} \tilde{\chi}_1^0 g_1 \text{Re} (N_{11} N_{13}) \cos \alpha + \dots \quad (608)$$

which gives the Feynman rule vertex factor (for Feynman rules it is $i\mathcal{L}$ that enters perturbatively) of

$$-ig_1 \text{Re} (N_{11} N_{13}) \cos \alpha ; \quad (609)$$

the factor of $1/2$ disappears because there are two ways of contracting external creation / annihilation operators with the two $\tilde{\chi}_1^0$ fields.

This Feynman rule vertex now combines with the $H^0 b \bar{b}$ vertex

factor of

$$\frac{-im_b g_2 \cos \alpha}{2m_W \cos \beta} \quad (610)$$

and the $\frac{1}{-m_{H^0}^2}$ propagator factor to give the expression for f_b isolated earlier

$$f_b = \frac{g_1 g_2 m_b \operatorname{Re}(N_{13} N_{11}) \cos^2 \alpha}{2m_W \cos \beta m_{H^0}^2} \quad (611)$$

appropriate in the limit under consideration. (Factor of 2 off?)

At this point, we can go back to Eq. (461) and look up the approximate results for N_{13} and N_{11} in the CP-conserving case and in the limit of $M_1 < M_2 < \mu$ and all $\gg m_Z$:

$$\begin{aligned} N_{11} &\sim 1 \\ N_{13} &\sim \frac{-m_Z s_W (M_1 c_\beta + \mu s_\beta)}{M_1^2 - \mu^2} \end{aligned}$$

$$N_{14} \sim \frac{m_Z s_W (M_1 s_\beta + \mu c_\beta)}{M_1^2 - \mu^2} \quad (612)$$

from which we see that N_{13} is suppressed basically by a factor of $m_Z s_W s_\beta / \mu$ when $|\mu|$ is large. This will tend to make the direct detection cross section (proportional to the square of f_b) kind of small.

2. Second, we can consider the case in which the cross section is dominated by light Higgs boson (h) exchange through its couplings to up-type quarks.

This is often found in the case of heavy squarks and heavy to moderate H . In this limiting case

$$\sigma_{\chi N} \sim \frac{g_1^2 g_2^2 |Re(N_{11} N_{14})|^2 m_N^4}{4\pi m_W^2 m_h^4} \left(f_{T_u} + \frac{4}{27} f_{TG} \right)^2, \quad (m_{\tilde{q}}, m_H \text{ large}, \cos \alpha \approx 1). \quad (613)$$

The expression for N_{14} given above shows that we again have a suppression by a factor of the square of something proportional

to m_Z/μ . Also, notice that we do not have the $\cos^2 \beta$ in the denominator that was present in the case of H^0 exchange. **So, the h^0 exchange is not enhanced $\propto \tan^2 \beta$ when $\tan \beta$ is large!**

If the heavy Higgs (H) is heavier than about ~ 500 GeV, exchange of the light Higgs generally dominates, leading to cross sections of around 10^{-8} to 10^{-10} pb for $|\mu|$ in the range of 200 GeV to 1 TeV.

3. Third, consider the case in which the elastic scattering cross section is dominated by the exchange of squarks through their couplings to strange and bottom quarks.

This is found for large to moderate $\tan \beta$ and squarks with masses well below 1 TeV. In this limiting case, and with approximately diagonal squark mass matrices,

$$\sigma_{\chi N} \sim \frac{g_1^2 g_2^2 |Re(N_{11} N_{13})|^2 m_N^4}{4\pi m_W^2 \cos^2 \beta m_{\tilde{q}}^4} \left(f_{T_s} + \frac{2}{27} f_{TG} \right)^2, \quad (\tilde{q} \text{ dominated, } \tan \beta \gg 1). \quad (614)$$

For squarks lighter than ~ 1 TeV, squark exchange can potentially provide the dominant contribution to neutralino-nuclei elastic

scattering.

LHC limits are currently pushing the CMSSM scenarios to $(m_{1/2}, m_0)$ mass scales that imply $m_{\tilde{q}} > 700 - 750$ GeV!

Using Eq. (539), we can crudely estimate the minimum target mass required to potentially detect neutralino dark matter.

A detector made up of Germanium targets (such as CDMS or Edelweiss, for example) would expect a WIMP with a nucleon-level cross section of 10^{-6} pb (10^{-42} cm²) to yield approximately 1 elastic scattering event per kilogram-day of exposure.

Such a target mass could thus be potentially sensitive to strongly mixed gaugino-higgsino neutralinos with light m_H and large $\tan\beta$.

The strongest current limits on spin-independent scattering have been obtained using $\sim 10^2$ kilogram-days of exposure, which can probe $\sigma_{\chi N} \sim 10^{-7}$ pb, *i.e.* $\sim 10^{-43}$ cm².

Reaching sensitivities near the 10^{-10} pb level, the natural prediction

when the $\tilde{\chi}_1^0$ is primarily bino (as opposed to being a strongly mixed state) will require ton-scale detectors capable of operating for weeks, months or longer with very low backgrounds.

Such detectors are under construction or being planned.

Current Limits

There are many operating experiments and they are continually updating the limits on the spin-independent cross section (always referenced to the “per-nucleon” cross section, σ_{SI}).

These are plotted in the $m_\chi - \sigma_{SI}$ plane.

The extracted σ_{SI} values depend on understanding the efficiencies in the experiment and especially the sensitivity of a given detector as a function of E_{recoil} .

For example, if no signal is seen at low m_χ , interpretation of the lack of signal is critically dependent on the sensitivity of the experiment

to low E_{recoil} values — the greater the sensitivity of the experiment at low E_{recoil} the stronger the limit of σ_{SI} .

Most experiments have not measured their sensitivity below $E_{\text{recoil}} \sim 10$ keV. Often, it is assumed that the sensitivity is roughly constant from 10 keV down to something like 2 – 3 keV. **As a result, the limits at low m_χ values could be too strong.**

There are other issues of interpretation. The most important are:

1. Channeling

For a few years it was thought that a given σ_{SI} could create a bigger signal as a result of the ability of the recoiling ion to “channel” its way out of the lattice (in some preferred direction) and thus deposit more energy in the sensors on the outside of the crystal.

However, it is now convincingly argued (Gondolo and Gelmini) that this effect is negligible. As a result, if a tentative signal is seen,

this will require a larger σ_{SI} than if it is assumed that channeling occurs.

2. Local Density

Most experiments assume a local χ density of order $\rho_\chi \sim 0.3 \text{ GeV}/\text{cm}^3$. However, quite a few papers now argue for ρ_χ as large as $\sim 0.4 - 0.5 \text{ GeV}/\text{cm}^3$.

There could also be a local “cusp” in ρ_χ leading to a still higher value of ρ_χ in the vicinity of our solar system.

Any increase in ρ_χ will imply a larger signal for given σ_{SI} .

Conversely, the larger the ρ_χ assumed, the stronger the limit on σ_{SI} if nothing is seen.

So, now let us step through the most recent experimental results, beginning in early 2010.

Remember: 10^{-4} pb is equivalent to 10^{-40} cm^2 .

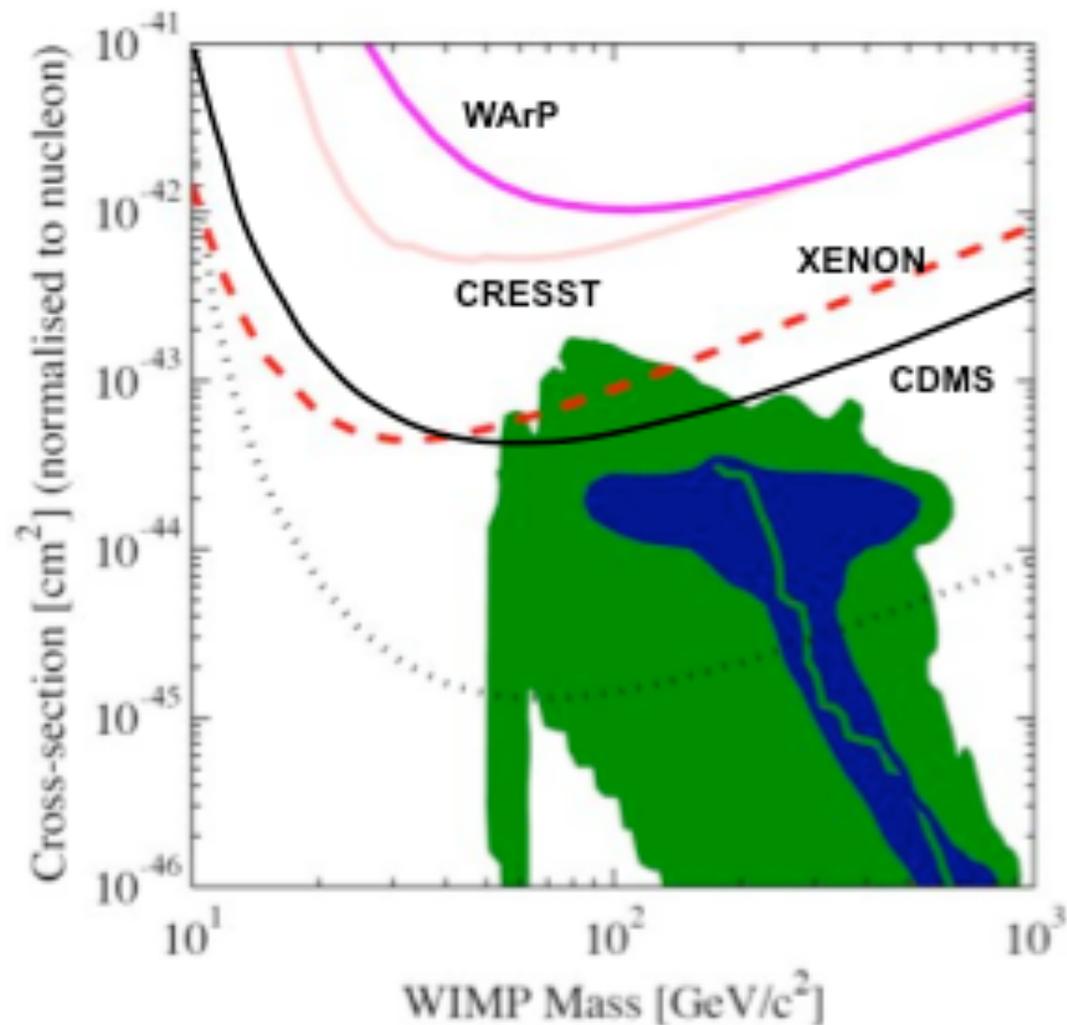


Figure 37: Current constraints on dark matter direct detection cross section from Chardin conference talk. The CDMS (full black line); XENON-10 (dashed line) — ZEPLIN-III is similar; CRESST (light solid line); and WArP (medium solid line). The light dotted line indicates expected sensitivity of operating experiments after a few more years. MSSM model predictions **without imposing the $B_s \rightarrow \mu^+ \mu^-$ constraints from the Tevatron** are shown by the solid regions.

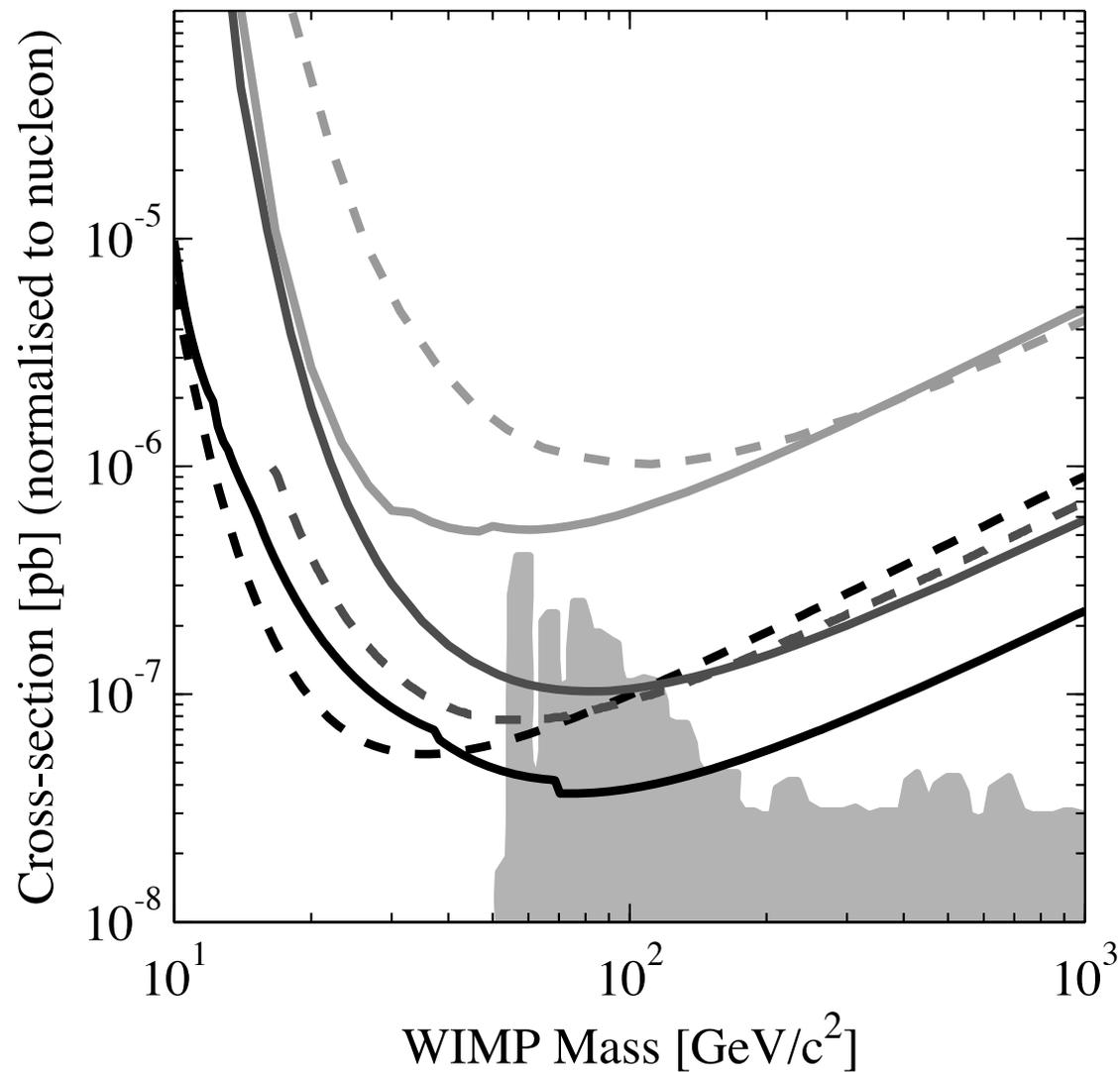


Figure 38: Upper limits from review of Schnee on the spin-independent WIMP-nucleon coupling σ_{SI} under the standard assumptions about the Galactic halo. Most sensitive limits are from cryogenic experiments (solid) CDMS (black), EDELWEISS-II (medium gray), and CRESST (light gray), and two-phase noble experiments (dashed) XENON10 (black), ZEPLIN-III (medium gray), and WArP (light gray). Current experiments already exclude part of the parameter space of MSSM models (shaded). Figure made using the Dark Matter Limit Plotter.

Above was situation in early 2010. Then came CoGeNT. Below is their recoil energy spectrum.

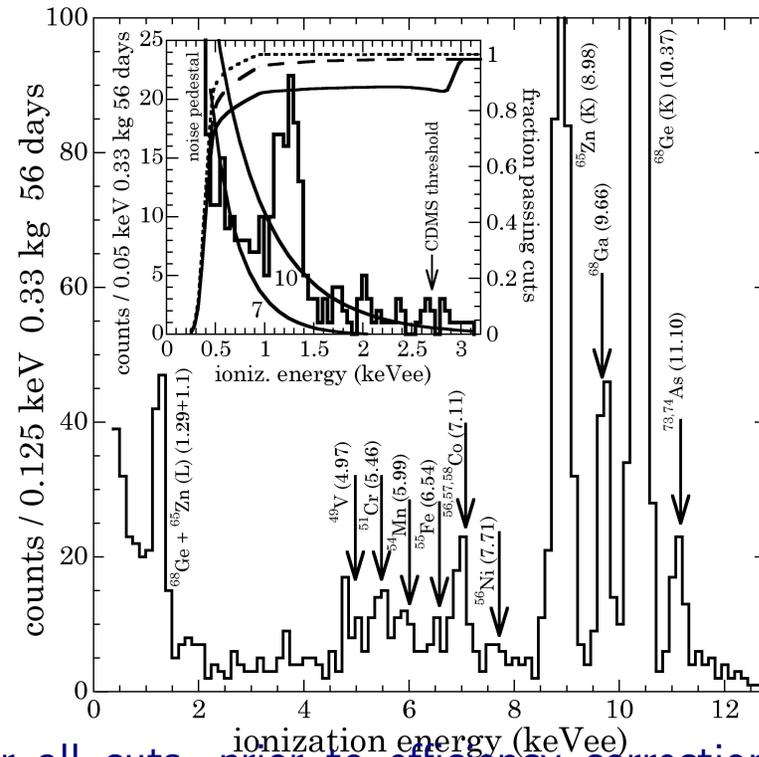


Figure 39: Low-energy spectrum after all cuts, prior to efficiency corrections. Arrows indicate expected energies for all viable cosmogenic peaks. *Inset:* Expanded threshold region, showing the ^{65}Zn and ^{68}Ge L-shell EC peaks. Overlapped on the spectrum are the sigmoids for triggering efficiency (dotted), trigger + microphonic PSD cuts (dashed) and trigger + PSD + rise time cuts (solid), obtained via high-statistics electronic pulser calibrations. Also shown are reference signals (exponentials) from $7 \text{ GeV}/c^2$ and $10 \text{ GeV}/c^2$ WIMPs with $\sigma_{SI} = 10^{-4} \text{ pb}$.

CoGeNT interprets the rise at low ionization energy as a dark matter signal. The next figures shows the CoGeNT, DAMA and CDMS 2-event positive signal regions.

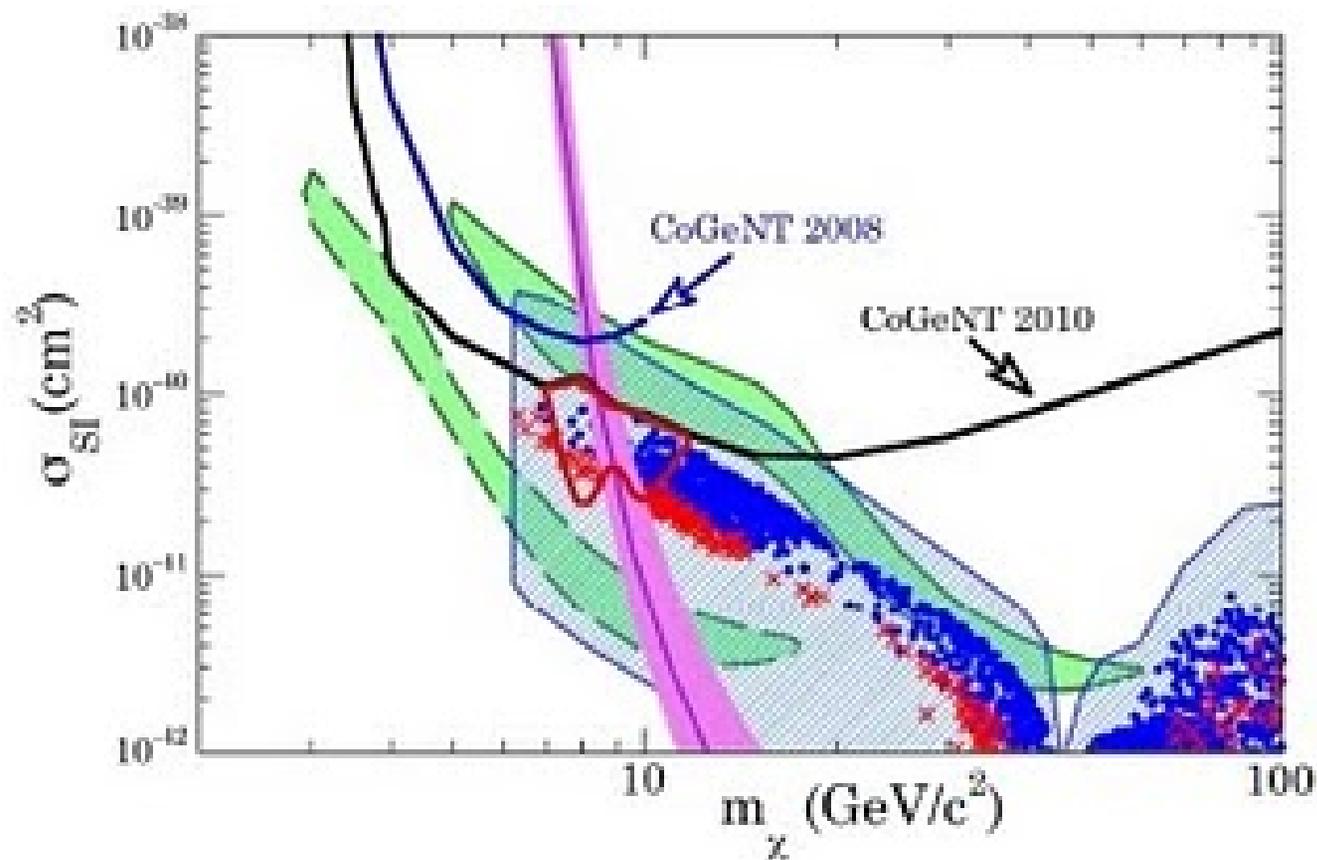


Figure 40: 90% C.L. WIMP exclusion limits from CoGeNT overlaid on Fig. 1 from bottino (did not impose $B_s \rightarrow \mu^+ \mu^-$ limits): green shaded patches denote the phase space favoring the DAMA/LIBRA annual modulation (the dashed contour includes ion channeling and should be ignored). The exact positions of these signal regions has been subject to revisions by theorists wanting to make them all consistent. The violet band is the region supporting the two CDMS candidate events. The scatter plot and the blue hatched region represent the supersymmetric models in bottino2 and their uncertainties, respectively, but without $B_s \rightarrow \mu^+ \mu^-$ constraints. Models including WIMPs with $m_\chi \sim 7-11 \text{ GeV}/c^2$ provide a good fit to CoGeNT data (red contour). The relevance of XENON10 constraints in this low-mass region has been questioned.

Another plot showing the same positive signal regions is below.

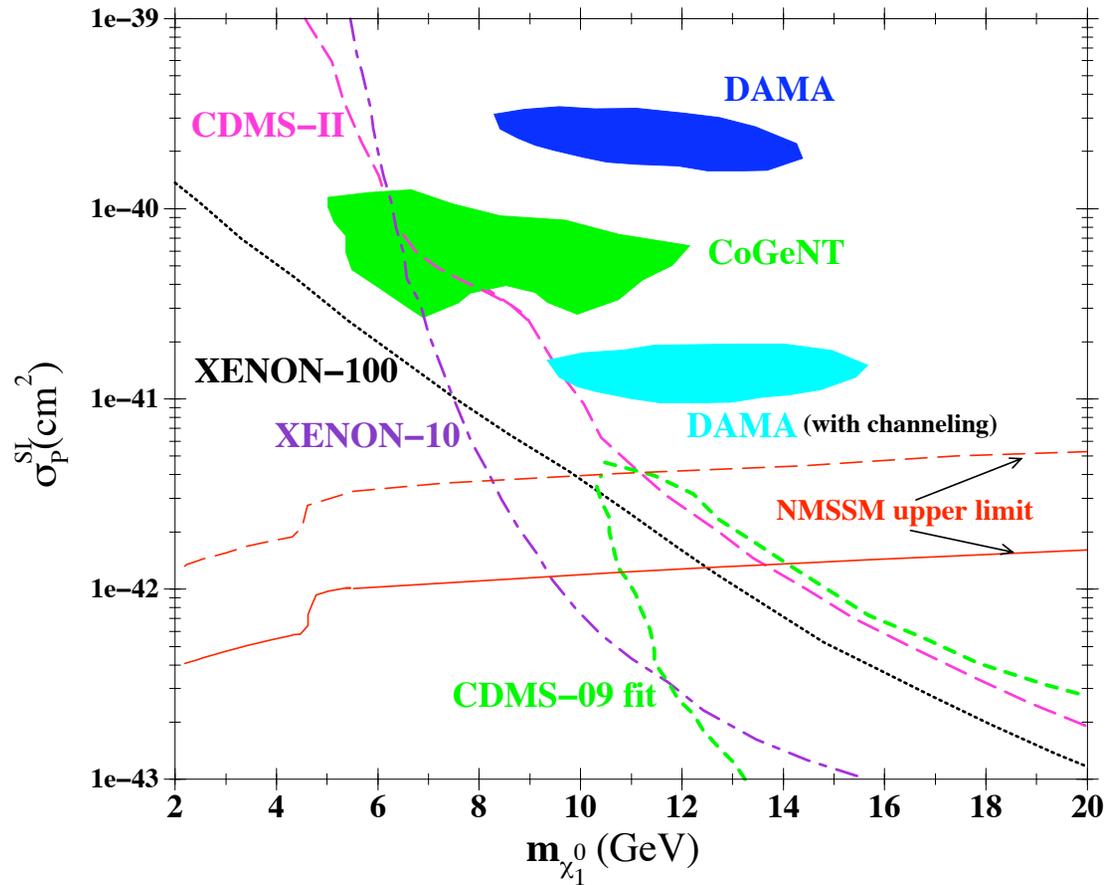


Figure 2: Upper bounds on the spin-independent cross section σ_p^{SI} in the NMSSM for default values of the strange quark content of nucleons as a full red line, and an enhanced strange quark content of nucleons as a dashed red line. Also shown are regions compatible with DAMA, CoGeNT and CDMS-II, and limits from Xenon10, Xenon100 and CDMS-II as explained in the text.

Figure 41: CDMS, DAMA and XENON compared to CoGeNT.

Unfortunately, it now appears that all these tentative positive signal

regions have been ruled out by the CDMS-II collaboration operating in the Soudan mine.

CDMS-II was already close to ruling the positive signal regions out based on using those of their Germanium crystals that had **measured** high sensitivity at low E_{recoil} operating at a shallow site (more backgrounds) location.

The new results employ this same set of high sensitivity crystals, but derive from data taken deep underground at the Soudan mine where backgrounds are much smaller.

Taken at face value (and no one has realistically challenged their results) all the tentative positive signals are ruled out.

This is good for the MSSM which could not predict such signals.

But, the NMSSM could!

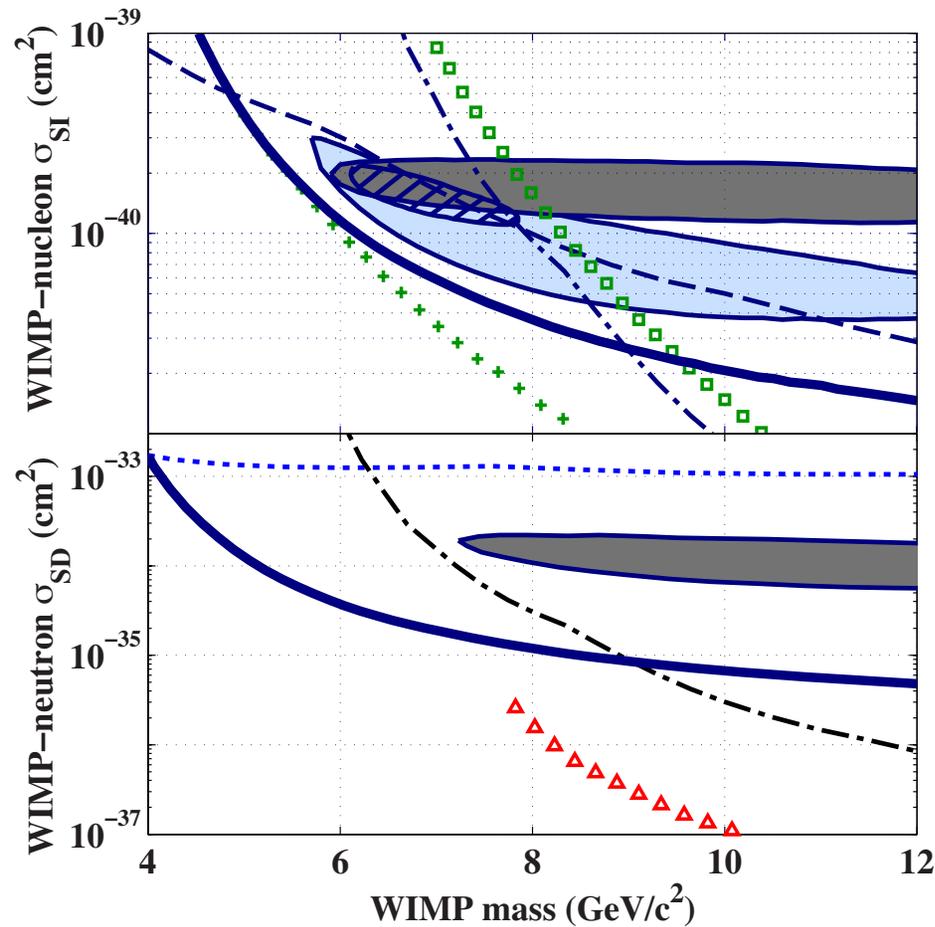


Figure 42: Upper limits from CDMS-II on light dark matter direct detection cross section. Top: comparison of the spin-independent (SI) exclusion limits from Soudan mine CDMS data with lowered detector thresholds (2 keV) (solid) to previous results in the same mass range (all at 90% C.L.). Limits from a low-threshold analysis of the CDMS shallow-site data Akerib:2010rr (dashed), CDMS II Ge results with a 10 keV threshold CDMSScience:2010 (dash-dotted), recalculated for lower WIMP masses, and XENON100 with constant (+) or decreasing (\square) scintillation-efficiency extrapolations at low energy Aprile:2010xx are also shown. The filled regions indicate possible signal regions from DAMA/LIBRA Bernabei:2008yi, Hooper:2010ly (dark), CoGeNT (light) Aalseth:2010vx, Hooper:2010ly, and a combined fit to the DAMA/LIBRA and CoGeNT data Hooper:2010ly (hatched). An escape velocity of 544 km/s was used for the CDMS and XENON100 exclusion limits, whereas the other results assume an escape velocity from 600–650 km/s.

So, why should we trust CDMS more than CoGeNT or DAMA? The reason is contained in the CDMS plot of Fig. 43.

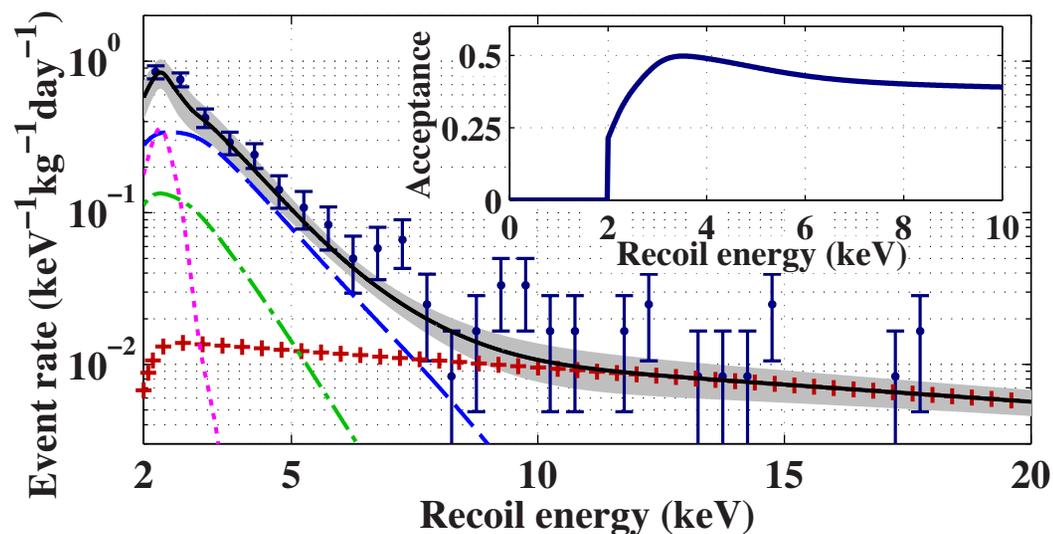


Figure 43: Comparison of the energy spectra for the candidate events and background estimates, co-added over the 8 detectors used in this analysis. The observed event rate (error bars) agrees well with the electron-recoil background estimate (solid), which is a sum of the contributions from zero-charge events (dashed), surface events (+), bulk events (dash-dotted), and the 1.3 keV line (dotted). The gray band denotes the 1σ statistical errors on the background estimate. The selection efficiencies have been applied to the background estimates for direct comparison with the observed rate, which does not include a correction for the nuclear-recoil acceptance. The inset shows the measured nuclear-recoil acceptance efficiency, averaged over the Germanium crystals used.

There are two critical items:

1. First, they seem to have a precise understanding of all their backgrounds, which combine to fully explain the observed spectrum.

The CoGeNT spectrum shown earlier has an incompletely understood background because it is a smaller, less vetoed type Germanium crystal. They do not understand the rise in the number of events as the recoil energy decreases and ascribe the continuous part to dark matter interactions which fits for $m_{\tilde{\chi}_1^0} \sim 7 - 10 \text{ GeV}$.

2. Second, the inset curve shows their actual measurement of the detection efficiency at low recoil energy — it is roughly constant down to 2 keV.

CoGeNT does not measure their efficiency for recoil energies below 10 keV and the dark matter description of the excess requires that this efficiency does not die off below 10 keV.

Supersymmetric Model Predictions

- The MSSM

We have already seen that the typical σ_{SI} in the MSSM is small when the $\tilde{\chi}_1^0$ is mainly bino.

The question is how far can we push. In particular, is it possible to obtain large σ_{SI} when $m_{\tilde{\chi}_1^0}$ is small? And, is there a lower limit on $\tilde{\chi}_1^0$ in the MSSM context?

1. In the plots we just reviewed, you saw some early predictions labelled “Bottino”. These came from hep-ph/0212379 (2003).
2. Since then, other such scans have been performed with similar results, but it was not until Feldman, Liu and Nath (arXiv:1003.0437) that the Tevatron upper bounds on $B_s \rightarrow \mu^+ \mu^-$ were incorporated into the scans. In that paper, it was concluded that a light $\tilde{\chi}_1^0$ with correct $\Omega^0 h_0^2$ (with $m_{\tilde{\chi}_1^0}$ large enough to be cold dark matter) and

large σ_{SI} was not possible and indeed a lower bound on the mass of a phenomenologically consistent $\tilde{\chi}_1^0$ was of order 20 GeV.

3. A more recent scan of the MSSM parameter space is that of arXiv:1009.4380v2. I will sketch their results.

– Their input constraints are given in Table 6. Note that $\Omega^0 h_0^2$ is allowed to range from the upper limit consistent with WMAP observations down to just 10% of the observed $\Omega^0 h_0^2$, for which other sources of DM would be needed.

– Input parameters are defined at the weak scale.

They assume minimal flavour violation and equality of the soft masses between sfermion generations.

They assume a common mass $m_{\tilde{e}}$ for all sleptons, and for all squarks $m_{\tilde{q}}$ (they checked that relaxing this did not matter).

They only allowed $A_t \neq 0$.

M_1 and M_2 were varied independently, in particular allowing $M_1 \ll M_2$.

They took $M_3 = 3M_2$.

μ , $\tan \beta$ and m_{A^0} were varied freely.

Table 6: List of constraints, from Nakamura:2010zzi unless noted otherwise.

constraint	value/range	tolerance	applied
Smasses	-	none	both
$\Omega^0 h_0^2$	0.01131 - 0.1131	0.0034	both
$(g-2)_\mu$	$25.5 \cdot 10^{-10}$	stat: $6.3 \cdot 10^{-10}$ sys: $4.9 \cdot 10^{-10}$	both
$\Delta\rho$	≤ 0.002	0.0001	MSSM
$b \rightarrow s\gamma$	$3.52 \cdot 10^{-4}$ (Barberio:2008fa, Misiak:2006zs)	th: $0.24 \cdot 10^{-4}$ exp: $0.23 \cdot 10^{-4}$	both
$B_s \rightarrow \mu^+ \mu^-$	$\leq 4.7 \cdot 10^{-8}$	$4.7 \cdot 10^{-10}$	both
$R(B \rightarrow \tau\nu)$	1.28 (Barberio:2008fa)	0.38	both
m_h	≥ 114.4	1%	MSSM
$Z \rightarrow \chi_1 \chi_1$	≤ 1.7 MeV	0.3 MeV none	MSSM NMSSM
$e^+ e^- \rightarrow \chi_1 \chi_{2,3}$	≤ 0.1 pb (Abbiendi:2003sc)	0.001 pb none	MSSM NMSSM
ΔM_s	$117.0 \cdot 10^{-13}$ GeV	th: $21.1 \cdot 10^{-13}$ GeV exp: $0.8 \cdot 10^{-13}$ GeV	NMSSM
ΔM_d	$3.337 \cdot 10^{-13}$ GeV	th: $1.251 \cdot 10^{-13}$ GeV exp: $0.033 \cdot 10^{-13}$ GeV	NMSSM

- Parameter ranges scanned designed to probe low $m_{\tilde{\chi}_1^0}$ were:

$$\begin{aligned}
 M_1 &\in [1, 100] \text{ GeV} & M_2 &\in [100, 2000] \text{ GeV} \\
 \mu &\in [0.5, 1000] \text{ GeV} & \tan \beta &\in [1, 75] \\
 m_{\tilde{l}} &\in [100, 2000] \text{ GeV} & m_{\tilde{q}} &\in [300, 2000] \text{ GeV} \\
 A_t &\in [-3000, 3000] \text{ GeV} & m_A &\in [100, 1000] \text{ GeV}
 \end{aligned} \tag{615}$$

The cases $\mu > 0$ and $\mu < 0$ were considered separately.

- The likelihood (given the errors/ranges of Table ??) associated with a given parameter choice is denoted Q .

Points with “reasonable” likelihood were found for:

$$\begin{aligned}
 M_1 &= [15, 19] \text{ GeV} && \text{yielding } m_{\tilde{\chi}_1^0} = [13, 15] \text{ GeV} \\
 |\mu| &< 150 \text{ GeV} && \text{but above limit implied by LEP } m_{\tilde{\chi}_1^\pm} \text{ bound} \\
 \tan \beta &= [40, 60] \\
 m_{A^0} &= [120, 170] \text{ GeV} && \text{funnel region needed} \\
 m_{\tilde{\ell}} &= [500, 1200] \text{ GeV}
 \end{aligned}$$

$$m_{\tilde{q}} = [0.8, 2] \text{ TeV} \quad (616)$$

The very lowest value of $m_{\tilde{\chi}_1^0}$ consistent with the constraints found in the scan was $m_{\tilde{\chi}_1^0} = 10 \text{ GeV}$.

- Preferred regions appear in the following figures. For σ_{SI} they employed $\sigma_{\pi N} = 45 \text{ MeV}$ and $\sigma_0 = 40 \text{ MeV}$ in micrOMEGAs. They made this choice since larger $\sigma_{\pi N}$ values imply bigger σ_{SI} that are more easily excluded.

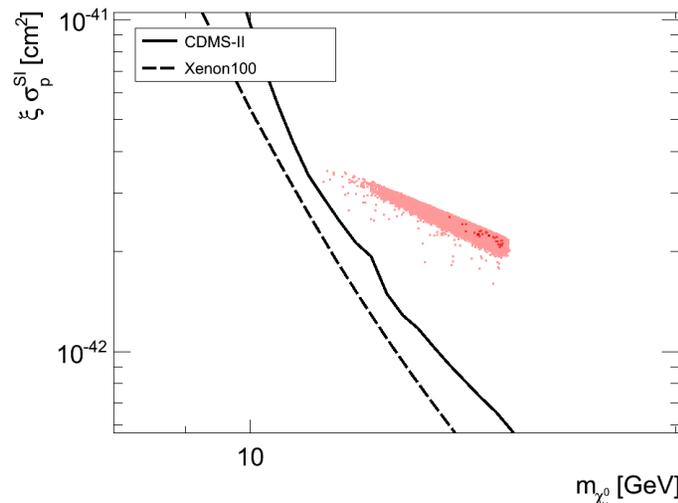


Figure 44: MSSM-EWSB scenario with $\mu > 0$ and $m_\chi < 15 \text{ GeV}$. Spin-independent cross section on proton times the fraction of neutralinos in the Milky Way dark halo (ξ) versus the neutralino mass m_χ . The dark red (light pink) points have a likelihood greater than 99.4% (68%). CDMS and Xenon curves are before latest CDMS-II Soudan results.

- * A constraint not listed in Table 6 is that coming from Tevatron limits on $gg \rightarrow b\bar{b} + (H^0, A^0)$ with $H^0, A^0 \rightarrow \tau^+\tau^-$. These imply that $\tan\beta$ cannot be too large at moderate $m_{A^0} \sim m_{H^0} \sim m_{H^\pm}$.
- * The Tevatron limits are compared to the MSSM allowed region in Fig. 45. We see that, quite independently of the σ_{SI} constraints, the Tevatron constraints also eliminate all MSSM scenarios with a relatively light $\tilde{\chi}_1^0$, resulting in $m_{\tilde{\chi}_1^0} > 15$ GeV being required.
- * For some reason, they did not consider constraints from the Tevatron using the limits on $gg \rightarrow b\bar{t}H^\pm + cc$ with $H^\pm \rightarrow \tau\nu_\tau$. These sometimes imply stronger constraints than those from the $H^0, A^0 \rightarrow \tau^+\tau^-$ limits.
- * Note that if the $B_s \rightarrow \mu^+\mu^-$ constraint is ignored in the scan, then many additional points become allowed, but all these additional points are excluded by the Tevatron constraints.

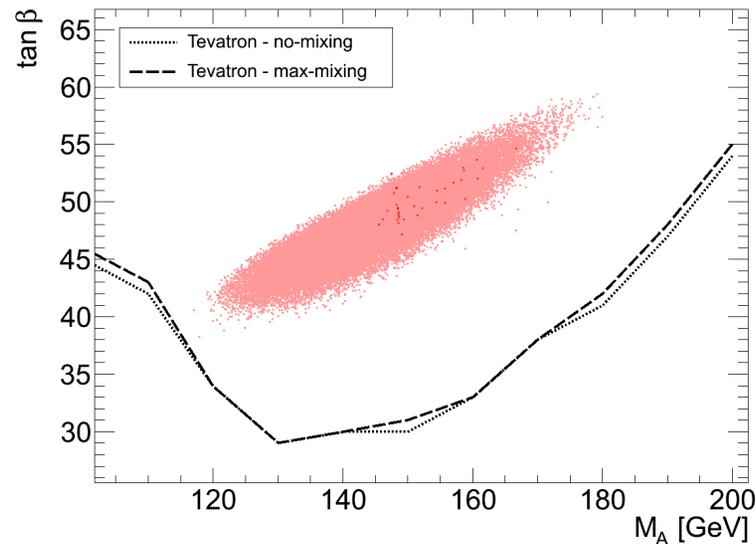


Figure 45: Distribution of the points selected by our MCMC analysis in the $\tan \beta - m_A$ plane in the MSSM-EWSB scenario with $\mu > 0$ and $m_{\tilde{\chi}_1^0} < 15$ GeV. The TEVATRON limits are displayed for the case of no-mixing (dash) or maximum mixing (full) in the stop sector, same color code as in Fig. 44.

- Incidentally, all these results shown are for $\mu > 0$. As we have learned $\mu < 0$ points almost inevitably in sharp disagreement with $(g - 2)_\mu$ and thus have very low likelihood. So, once again $m_{\tilde{\chi}_1^0} < 15$ GeV is ruled out by σ_{SI} and Tevatron constraints.
- Enlarging the parameter scan somewhat yields some additional

possibilities.

Fig. 46 displays the points selected by the MCMC (Markov Chain Monte Carlo) in the plane $(m_A, \tan \beta)$.

The Tevatron constraints are superimposed (but not imposed).

- * At low values of $m_{A^0} < 300$ GeV, there are two separate regions of $\tan \beta$. One is peaked around $10 - 20$ while the second lies between $50 - 70$.
- * Typically, when the pseudoscalar is light, constraints on B -physics decrease the value of the likelihood especially when $\tan \beta$ is large.
- * However, as we have seen previously, neutralino annihilation through a pseudoscalar exchange leads to an acceptable relic density and to a good global likelihood when $\tan \beta > 50$.
- * Note that, even though very large values of $\tan \beta$ do not appear plausible, they do indicate the type of regions that lead to a neutralino mass in the < 50 GeV range.

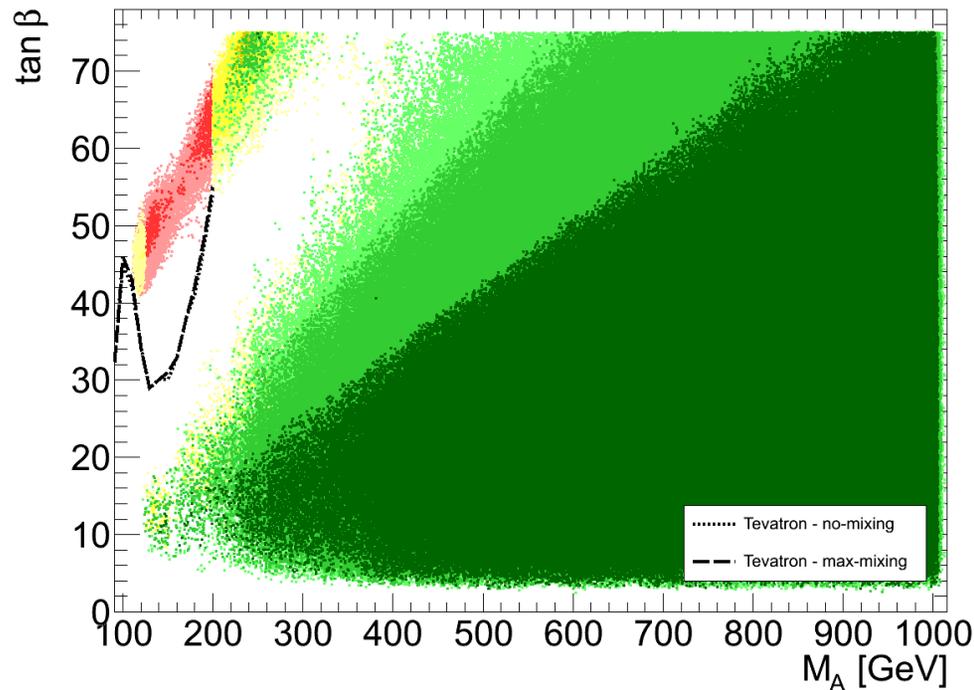


Figure 46: Distribution of the points selected by the MCMC analysis in the $\tan \beta - m_A$ plane in the MSSM-EWSB scenario with $\mu > 0$ and $m_\chi < 50$ GeV. In red, we display the points which are excluded by both Tevatron, XENON 100 and CDMS. In yellow, we show the points which satisfy Tevatron and which are excluded by XENON 100 and CDMS and in green, all the points which survive both constraints.

The spin-independent cross section versus the neutralino mass is displayed in Fig. 47. The scenarios where $m_{\tilde{\chi}_1^0} > 28$ GeV survive

both the Tevatron and Direct Detection limits.

Although such a value is likely to be irrelevant to explain CoGeNT data, it might be important in light of the two CDMS “events”. But, these are now ruled out by the latest CMDS-II data.

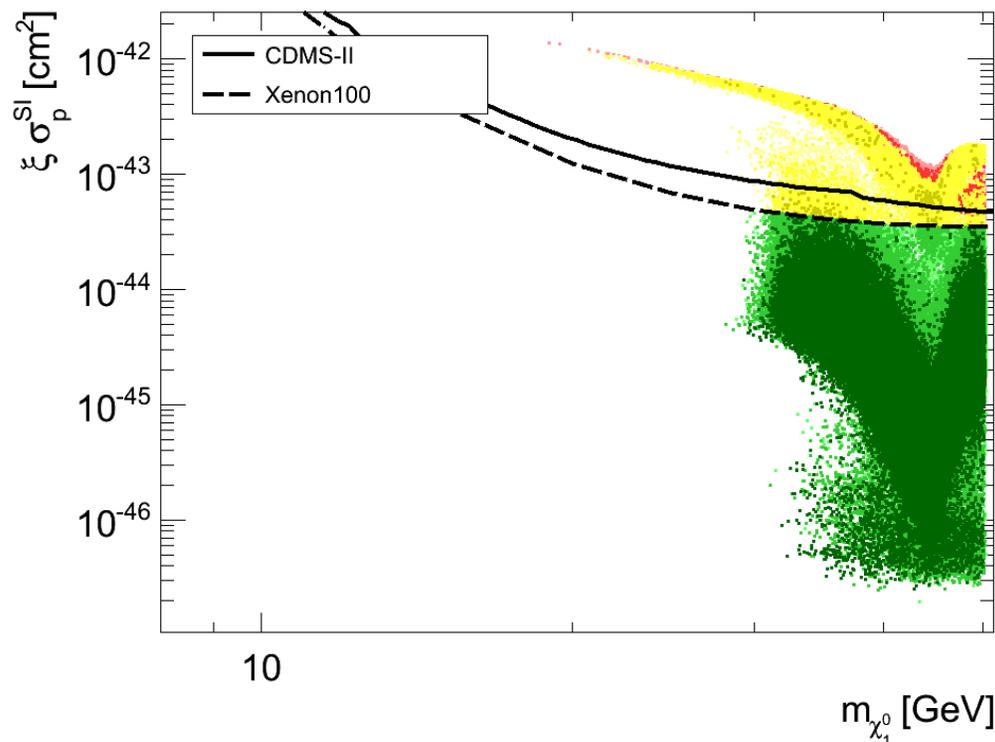


Figure 47: Spin independent cross section vs the neutralino mass in the MSSM-EWSB scenario with $\mu > 0$ and $m_\chi < 50$ GeV, same color code as in Fig. 46

- The NMSSM

We have not studied the details, but it is possible to illustrate the much greater freedom that one has in the NMSSM context.

Basic NMSSM features (a seminal paper was that by me, Ellis, Haber, Roszkowski and Zwirner):

1. The Next-to-Minimal Supersymmetric Standard Model (NMSSM) is a simple extension of the MSSM that provides a solution to the naturalness problem. This is achieved by the introduction of a gauge singlet superfield, denoted by \hat{S} .
2. The VEV of the scalar field component of the singlet determines the effective parameter $\mu = \lambda \langle S \rangle$ which is then naturally of the EW scale.
3. The part of the superpotential involving Higgs fields reads

$$W = \lambda \hat{S} \hat{H}_u \hat{H}_d + \frac{1}{3} \kappa \hat{S}^3 \quad (617)$$

(note: no $\mu \hat{H}_u \hat{H}_d$ terms) and the soft Lagrangian

$$\begin{aligned} \mathcal{L}_{\text{soft}} = & m_{H_u}^2 |H_u|^2 + m_{H_d}^2 |H_d|^2 + m_S^2 |S|^2 \\ & + (\lambda A_\lambda H_u H_d S + \frac{1}{3} \kappa A_\kappa S^3 + h.c.) \end{aligned} \quad (618)$$

4. The NMSSM contains three neutral scalar fields, h_1, h_2, h_3 and two pseudoscalar neutral fields, a_1, a_2 as well as a charged Higgs, H^\pm .
5. The model also contains five neutralinos, the new Lagrangian basis field is the singlino, \tilde{S} .

The mass matrix for the neutralinos takes the form:

$$\mathbf{M}_{\tilde{N}} = \begin{pmatrix} M_1 & 0 & -c_\beta s_W m_Z & s_\beta s_W m_Z & 0 \\ 0 & M_2 & c_\beta c_W m_Z & -s_\beta c_W m_Z & 0 \\ -c_\beta s_W m_Z & c_\beta c_W m_Z & 0 & -\lambda s & \lambda s_\beta v \\ s_\beta s_W m_Z & -s_\beta c_W m_Z & -\lambda s & 0 & -\lambda c_\beta v \\ 0 & 0 & -\lambda s_\beta v & -\lambda c_\beta v & 2\kappa s \end{pmatrix}. \quad (619)$$

For a pure state, the singlino mass is simply

$$m_{\tilde{S}} = 2\kappa s = 2\frac{\kappa\mu}{\lambda}. \quad (620)$$

Of course, one must diagonalize the neutralino mass matrix (which is now 5×5) to obtain the mass eigenstates. After diagonalization, the lightest neutralino is written:

$$\tilde{\chi}_1^0 = N_{11}\tilde{B} + N_{12}\tilde{W}^3 + N_{13}\tilde{H}_d + N_{14}\tilde{H}_u + N_{15}\tilde{S} \quad (621)$$

6. After using the minimization conditions of the Higgs potential, the Higgs sector is described by six free parameters, μ , $\tan\beta$ as well as λ , κ , A_λ , A_κ .

Other free parameters of the model are, as in the MSSM, the soft masses for sfermions, trilinear couplings and gaugino masses.

7. An important feature of the model is that both the singlino and the singlet fields can be very light and yet escape the LEP bounds.

This is because these fields mostly decouple from the SM fields. This opens up the possibility for new annihilation mechanisms for light neutralinos in particular if the LSP possesses an important singlino component.

The singlino can annihilate efficiently through the exchange of light singlet Higgses as well as into light Higgs singlets.

McElrath, I and Hooper were the first to show that a very light $\tilde{\chi}_1^0$ was possible in the NMSSM context while obtaining correct $\Omega^0 h_0^2$.

Whether large σ_{SI} at small $m_{\tilde{\chi}_1^0}$ has been pursued in explorations of the parameter space of the NMSSM model by several groups of authors, including myself, Tait, Hooper and Belikov.

The paper we have just been reviewing in the MSSM case is the most recent of these papers and it is convenient to present their results as they employ the same procedure as for the MSSM.

However, the scans by the other authors reveal regions of interest

that are quite different from those preferred by the scan we now discuss.

Let us discuss the results of their scans focused on getting $m_{\tilde{\chi}_1^0} < 15 \text{ GeV}$. For this scan, their NMSSM priors lie in the range:

$$\begin{aligned}
 M_1 &\in [1, 200] \text{ GeV} & M_2 &\in [100, 2000] \text{ GeV} \\
 \mu &\in [0., 1000] \text{ GeV} & \tan \beta &\in [0.1, 65] \\
 \lambda &\in [0, 0.75] & \kappa &\in [0., 0.65] \\
 A_\lambda &\in [-2000, 5000] \text{ GeV} & A_\kappa &\in [-5000, 2000] \text{ GeV} \\
 m_{\tilde{l}} &\in [100, 2000] \text{ GeV} & m_{\tilde{q}} &\in [300, 2000] \text{ GeV} \\
 A_t &\in [-3000, 3000] \text{ GeV} & &
 \end{aligned} \tag{622}$$

As before, they assume common soft masses for squarks and sleptons and we keep the gaugino masses M_1 and M_2 uncorrelated while $M_3 = 3M_2$.

Some basic features of their scan are the following.

- As in the MSSM, M_1 is peaked below 20 GeV although a long tail extends to 200 GeV.
In this tail the LSP is mostly a singlino which mass is determined from Eq. (620).
- The parameter λ that determines the mixing of the singlino to other neutralinos is never very small, so that the singlino does not decouple completely.
- The preferred values for $\mu \approx 150 - 250$ GeV are higher than in the MSSM.
- On the one hand, LEP2 limits on $e^+e^- \rightarrow \tilde{\chi}_1^0 \tilde{\chi}_i^0$ or on the light Higgs constrain low values of μ while a light singlino LSP prefers low values for μ , Eq. (620).
- The parameter $\kappa \ll 1$ also favours a light singlino.
- Intermediate values of $\tan \beta$ are preferred.
- The parameter A_κ that controls the mass of the singlet Higgses is always small to ensure a light scalar/pseudoscalar as required for

LSP annihilation while A_λ is usually well above 1 TeV.

- Sleptons are preferably light while squarks are above 1 TeV.
- The LSP mass ranges from 1 – 15 GeV with a distribution peaked towards higher masses.

This LSP is either mostly bino or mostly singlino with in any case some higgsino component.

- The most important feature of this scenario is the fact that the Higgs spectrum is constrained: one always predicts a light scalar, dominantly singlet, with a mass below 120 GeV (generally below 30 GeV) as well as a pseudoscalar singlet with a mass preferably below 30 GeV.

Note that the value of 30 GeV for the mass corresponds to twice the neutralino mass and is thus just a consequence of the prior on the neutralino mass.

- Furthermore, either $m_{\tilde{\chi}_1^0} - m_{a_1}/2 < 1 - 4 \text{ GeV}$ (with a similar mass splitting with h_1) or $m_{\tilde{\chi}_1^0} > m_{h_1}$.

This is because the annihilation of the light LSP relies either on pseudoscalar/scalar exchange or on the new light scalar pairs final states.

- The rest of the Higgs sector consists of MSSM-like doublets with preferred values for the heavy neutral and charged scalars above 2 TeV.

Note that one must check that the recent re-analysis of LEP2 limits on a Higgs decaying into two light pseudo-scalars does not put further constraints on the model parameters.

The light LSP scenarios can be classified in three broad classes:

1. a (pure or mixed) singlino LSP annihilating via pseudoscalar/scalar singlet Higgses into fermion final states, for this only a small singlino component of the LSP is necessary.
2. a bino LSP with small higgsino/singlino components annihilating into a pair of light scalar Higgses or
3. as in the MSSM a bino LSP with some Higgsino component

annihilating via Higgs doublets. This channel is more efficient at large values of $\tan\beta$ although the B-physics constraints severely restrict the very large values of $\tan\beta$.

Some features:

1. The predictions for the elastic scattering cross section span several orders of magnitude, from 10^{-56} to 10^{-38}cm^2 , see Fig. 48.
2. The largest cross sections are found in scenarios with a light h_1 , for example for $\sigma_{\chi p}^{SI} > 10^{-43}(10^{-41})\text{cm}^2$ requires $m_{h_1} < 20(8)$ GeV. At first sight this can be a bit surprising since such a light Higgs is dominantly singlet and thus couples very weakly to quarks in the nucleon - recall that the $h_1 q\bar{q}$ coupling is only possible through the doublet component- nevertheless this suppressed coupling is compensated by an enhancement factor due to the small h_1 mass in the propagator, $\propto 1/m_{h_1}^2$.
3. In scenarios where the elastic scattering cross-section is large, the LSP is generally dominantly bino (or, in a few cases, a singlino) with

a non-negligible higgsino fraction. This means that the doublet h_2 also contributes to the spin independent cross section since the LSP coupling to the doublet depends on the Higgsino component of the LSP.

4. The lowest σ_{SI} values are found in scenarios where the LSP pair-annihilation benefits from the enhancement of the pseudoscalar exchange in the s-channel near the resonance while the elastic scattering cross-section, which proceeds through scalar exchange in t-channel, does not benefit from a similar enhancement.

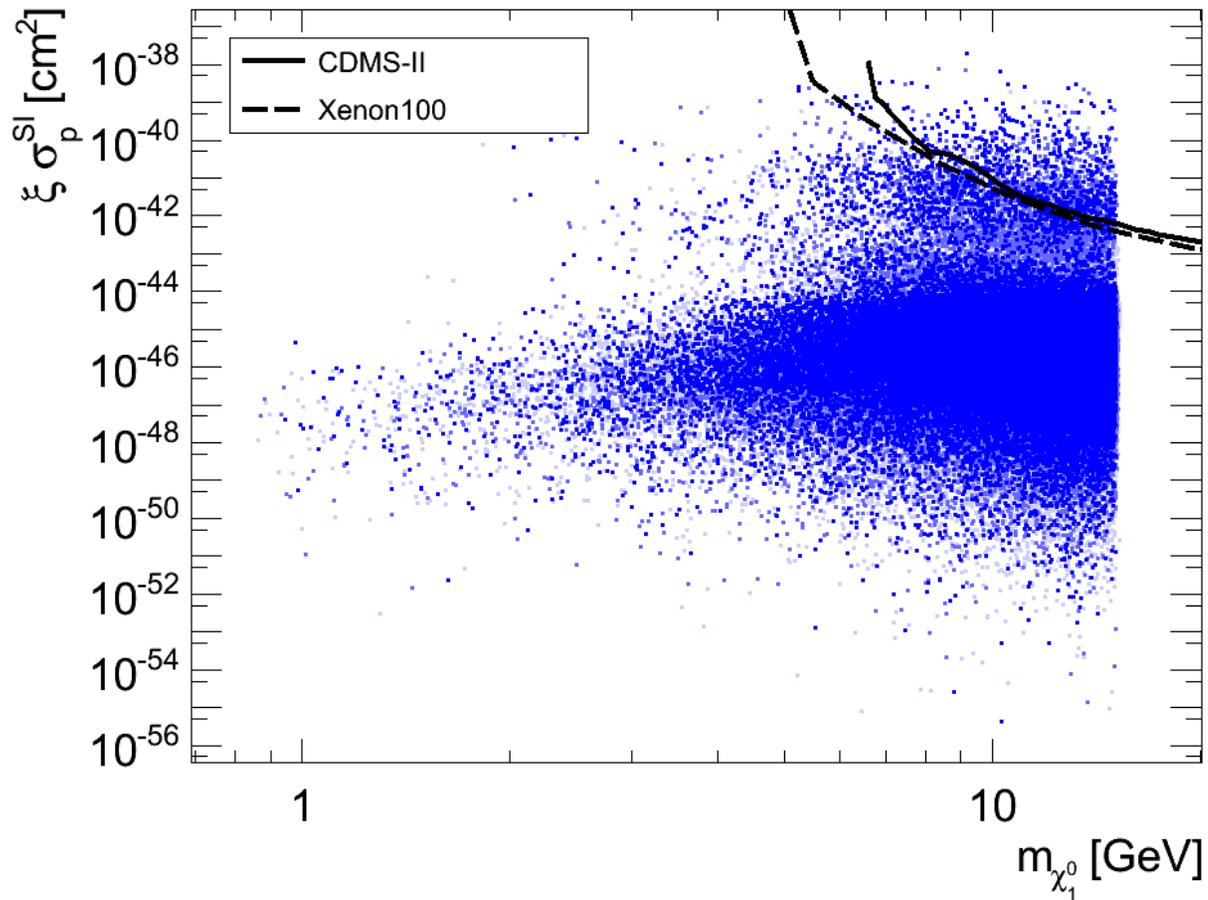


Figure 48: NMSSM scenario with $\mu > 0$ and $m_\chi < 15$ GeV. Spin-independent neutralino-proton cross section times the fraction of neutralinos in the Milky Way dark halo (ξ) versus the neutralino mass m_χ . Points in dark blue correspond to points with a likelihood greater than 99.4 %; Points in blue, correspond to points with likelihood greater than 95.4 % and smaller than 99.4 % of the maximum Likelihood and points in pale blue are all the remaining points having a likelihood greater than 68 %. The CDMS limits correspond to the plain curve while the XENON limits correspond to the dash curve.

- Other NMSSM Studies

As stated earlier, the scans of the Belanger et.al. paper do not find all interesting points with large σ_{SI} and correct $\Omega^0 h_0^2$.

As summary of the earlier papers on this subject appears in my review paper (arXiv:1010.1789 [hep-ph]). I will focus on the study I did with Belikov, Hooper and Tait (arXiv:1009.0549 [hep-ph]).

It is useful to give some analytic formulae that I think you can quickly understand. I will be focusing on the case where the lightest Higgs, h_1 , is primarily singlet and the lightest neutralino is primarily a singlino. In this limit, you can get (too) big σ_{SI} with correct $\Omega^0 h_0^2$ quite easily.

The Singlino-Singlet Scenarios

MODEL 1

In (Belikov:2010yi), we pursued the NMSSM and looked for scenarios

of the SS type. What we found was a kind of see-saw balance between $\Omega_{\tilde{\chi}_1^0} h^2$ and σ_{SI} such that when $\Omega_{\tilde{\chi}_1^0} h^2 \sim 0.1$ then σ_{SI} is very naturally in the CoGeNT/DAMA preferred zone. Below, I provide a few details.

The coupling of $\tilde{\chi}_1^0 \tilde{\chi}_1^0$ to down-type quarks is given by:

$$\frac{f_d}{m_d} = \frac{g_2 \kappa N_{15}^2 \tan \beta F_s(h_1) F_d(h_1)}{8 m_W m_{h_1}^2} \quad (623)$$

where $h_1 = F_d(h_1) H_d^0 + F_u(h_1) H_u^0 + F_s(h_1) H_S^0$. The $\frac{1}{3} \kappa \hat{S}^3$ term of the superpotential gives the crucial trilinear coupling of a singlino pair to the singlet Higgs H_S^0 proportional to κ . For $N_{15}^2 \sim 1$, this leads to

$$\sigma_{SI} \approx 2.2 \times 10^{-4} \text{ pb} \left(\frac{\kappa}{0.6} \right)^2 \left(\frac{\tan \beta}{50} \right)^2 \left(\frac{45 \text{ GeV}}{m_{h_1}} \right)^4 \left(\frac{F_s^2(h_1)}{0.85} \right) \left(\frac{F_d^2(h_1)}{0.15} \right),$$

which is consistent with the value required by CoGeNT and

DAMA/LIBRA for the indicated κ , m_{h_1} and h_1 component values.

Furthermore, the large singlet fraction $F_s^2(h_1) \sim 0.85$ of the h_1 will allow it evade the constraints from LEP II and the Tevatron.

Meanwhile, the thermal relic density of neutralinos is determined by the annihilation cross section and the $\tilde{\chi}_1^0$ mass.

In the mass range we are considering here, the dominant annihilation channel is to $b\bar{b}$ (or, to a lesser extent, to $\tau^+\tau^-$) through the s -channel exchange of the **same** scalar Higgs, h_1 , as employed for elastic scattering, yielding:

$$\sigma_{\tilde{\chi}_1^0\tilde{\chi}_1^0\rightarrow b\bar{b}} = \frac{N_c g_2^2 \kappa^2 m_b^2 F_s^2(h_1) F_d^2(h_1)}{64\pi m_W^2 \cos^2 \beta} \frac{m_{\tilde{\chi}_1^0}^2 (1 - m_b^2/m_{\tilde{\chi}_1^0}^2)^{3/2} v^2}{(4m_{\tilde{\chi}_1^0}^2 - m_{h_1}^2)^2 + m_{h_1}^2 \Gamma_{h_1}^2}, \quad (624)$$

where v is relative velocity between the annihilating neutralinos, $N_c = 3$ is a color factor and Γ_{h_1} is the width of the exchanged Higgs.

The annihilation cross section into $\tau^+\tau^-$ is obtained by replacing $m_b \rightarrow m_\tau$ and $N_c \rightarrow 1$.

This yields the thermal relic abundance of neutralinos:

$$\Omega_{\chi_1^0} h^2 \approx \frac{10^9}{M_{\text{Pl}} T_f \sqrt{g_\star}} \frac{m_{\chi_1^0}}{\langle \sigma_{\chi_1^0 \chi_1^0 v} \rangle}, \quad (625)$$

where g_\star is the number of relativistic degrees of freedom at freeze-out, $\langle \sigma_{\chi_1^0 \chi_1^0 v} \rangle$ is the thermally averaged annihilation cross section at freeze-out, and T_f is the temperature at which freeze-out occurs.

For the range of masses and cross sections considered here, we find $m_{\chi_1^0}/T_f \approx 20$, yielding a thermal relic abundance of

$$\Omega_{\chi_1^0} h^2 \approx 0.11 \left(\frac{0.6}{\kappa} \right)^2 \left(\frac{50}{\tan \beta} \right)^2 \left(\frac{m_{h_1}}{45 \text{ GeV}} \right)^4 \left(\frac{7 \text{ GeV}}{m_{\chi_1^0}} \right)^2 \left(\frac{0.85}{F_s^2(h_1)} \right) \left(\frac{0.15}{F_d^2(h_1)} \right) \quad (626)$$

i.e. naturally close to the measured dark matter density, $\Omega_{\text{CDM}} h^2 = 0.1131 \pm 0.0042$ for the same choices for κ , m_{h_1} and composition fractions as give CoGeNT/DAMA-like σ_{SI} . Note the $\Omega_{\tilde{\chi}_1^0} h^2 \sim \sigma_{SI}$ “SEE-SAW”, *i.e.* $\Omega_{\tilde{\chi}_1^0} h^2 \times \sigma_{SI} \sim \text{const.}$

The only question is can we achieve the above situation without violating LEP and other constraints.

Basically, one wants a certain level of decoupling between the singlet sectors and the MSSM sectors, but not too much. To find out, we performed parameter scans with micrOMEGAs and incorporated the latest B -physics and Tevatron constraints.

We found points for $15 < \tan \beta < 45$ that are consistent (within the usual $\pm 2\sigma$ combined theory plus experimental windows – excursions in $b \rightarrow s\gamma$ and $b\bar{b}h, h \rightarrow \tau^+\tau^-$ that fall slightly outside this window are present at high $\tan \beta$) with all collider and B -physics constraints having the appropriate thermal relic density and σ_{SI} as large as $few \times 10^{-4}$ pb.

I discuss one 'typical' point that does the job. Its properties are tabulated in Table 7.

Table 7: Properties of a typical ENMSSM point with $\tan \beta = 45$ and $m_{\text{SUSY}} = 1000$ GeV.

λ	κ	λ_s	A_λ	A_κ	M_1	M_2	M_3	A_{soft}
0.011	0.596	-0.026 GeV	3943 GeV	17.3 GeV	150 GeV	300 GeV	900 GeV	679 GeV
B_S		μ_S	v_S^3	μ	B_μ	μ_{eff}	B_μ^{eff}	
0		7.8 GeV	4.7 GeV	164 GeV	658 GeV ²	164 GeV	556 GeV ²	
m_{h_1}		m_{h_2}	m_{h_3}	m_{a_1}	m_{a_2}	m_{H^+}		
82 GeV		118 GeV	164 GeV	82 GeV	164 GeV	178 GeV		
$F_S^2(h_1)$	$F_d^2(h_1)$	$F_S^2(h_2)$	$F_u^2(h_2)$	$F_S^2(h_3)$	$F_d^2(h_3)$	$F_S^2(a_1)$	$F_S^2(a_2)$	
0.86	0.14	0.0	0.996	0.14	0.86	0.86	0.14	
$C_V(h_1)$	$C_V(h_2)$	$C_V(h_3)$	$C_{h_1 b\bar{b}}$	$C_{h_2 b\bar{b}}$	$C_{h_3 b\bar{b}}$	$C_{a_1 b\bar{b}}$	$C_{a_2 b\bar{b}}$	
-0.0096	0.999	-0.041	16.8	2.9	41.7	-16.9	41.7	
$m_{\tilde{\chi}_1^0}$		N_{11}^2	$N_{13}^2 + N_{14}^2$	N_{15}^2	σ_{SI}	$\Omega_{\tilde{\chi}_1^0} h^2$		
4.9 GeV		0.0	0.0	1.0	2.0×10^{-4} pb	0.105		
$B(h_1 \rightarrow \tilde{\chi}_1^0 \tilde{\chi}_1^0)$		$B(h_1 \rightarrow b\bar{b}, \tau^+ \tau^-)$		$B(h_2 \rightarrow \tilde{\chi}_1^0 \tilde{\chi}_1^0)$		$B(h_2 \rightarrow b\bar{b}, \tau^+ \tau^-)$		$B(H^+ \rightarrow \tau^+ \nu)$
0.64		0.33, 0.03		0.003		0.88, 0.092		0.97
$B(a_1 \rightarrow \tilde{\chi}_1^0 \tilde{\chi}_1^0)$		$B(a_1 \rightarrow b\bar{b}, \tau^+ \tau^-)$		$B(a_2, h_3 \rightarrow \tilde{\chi}_1^0 \tilde{\chi}_1^0)$		$B(a_2, h_3 \rightarrow b\bar{b}, \tau^+ \tau^-)$		
0.64		0.33, 0.03		0.05		0.85, 0.095		

Let us note the following regarding this particular point.

1. What you see is that the h_1, a_1 have separated off from something that is close to an MSSM-like Higgs sector with $h_2 \sim h^0$ being SM-like and $h_3 \sim H^0$, $a_2 \sim A^0$ and $H^+ \sim H^+$.
2. Detection of the h_2 would be possible via the usual SM-like

detection modes planned for the MSSM h^0 .

3. There are some $h_2, a_2 \rightarrow \tilde{\chi}_1^0 \tilde{\chi}_1^0$ decays, but at such a low branching ratio level that detection of these invisible decay modes would be unlikely, even if very interesting.
4. Decays to pairs of Higgs of any of the heavier Higgs bosons are not of importance. Of course, by choosing $m_{\text{SUSY}} = 1000 \text{ GeV}$ so that $m_{h_2} > 114 \text{ GeV}$ (beyond the LEP limits), we have not forced the issue. It will be interesting to look for SS scenarios that are ideal-Higgs-like with $m_{h_2} < 110 \text{ GeV}$.
5. One sees that h_1 and a_1 decay primarily to $\tilde{\chi}_1^0 \tilde{\chi}_1^0$ but that there also decays to $b\bar{b}$ and $\tau^+ \tau^-$ with reduced branching ratios of 0.33 and 0.03 compared to the normal $B(b\bar{b}) \sim 0.85$ and $B(\tau^+ \tau^-) \sim 0.12$.
6. h_1 and a_1 do have somewhat enhanced couplings to $b\bar{b}$ (in this example $C_{h_1 b\bar{b}}, C_{a_1 b\bar{b}} \sim \sqrt{F_d^2(h_1, a_1)} \tan \beta \sim 17$) and so the rates for $gg \rightarrow b\bar{b}h_1 + gg \rightarrow b\bar{b}a_1$ will be quite substantial. However, the reduced $B(h_1, a_1 \rightarrow \tau^+ \tau^-) \sim 0.03$ implies that detection of such

production in the $b\bar{b} + \tau^+\tau^-$ final state might prove challenging, probably requiring very high L at the LHC.

7. Further work is needed to quantify discovery prospects in the $gg \rightarrow b\bar{b} + (h_1, a_1) \rightarrow b\bar{b} + \cancel{E}_T$ channel.
8. At this large $\tan\beta$, detection of the h_3 and a_2 would certainly be possible in $gg \rightarrow b\bar{b}h_3 + b\bar{b}a_2$ in the $h_3, a_2 \rightarrow \tau^+\tau^-$ decay channel.
9. For this sample case, the charged Higgs is *just* too heavy to allow $t \rightarrow H^+b$ decays and so one would have to turn to $gg \rightarrow \bar{t}bH^+ + t\bar{b}H^-$ with detection of the charged Higgs in the $\tau\nu_\tau$ final state. Further investigation is needed to assess the feasibility of such detection, but at least the cross section is very enhanced by virtue of the large $\tan\beta$ value.

A final note regarding this scenario. It is the very large value of A_λ and the very small λ that keep the singlet and MSSM Higgs sectors fairly separate.

MODEL 2

As a general remark, it is clear from the formulae given earlier that we will get large σ_{SI} and roughly correct $\Omega^0 h_0^2$ for a continuum of possible κ , $\tan \beta$, m_{h_1} , $F_d^2(h_1)$ and $F_s^2(h_1)$ values obeying

$$\left(\frac{0.6}{\kappa}\right)^2 \left(\frac{50}{\tan \beta}\right)^2 \left(\frac{m_{h_1}}{45 \text{ GeV}}\right)^4 \left(\frac{7 \text{ GeV}}{m_{\tilde{\chi}_1^0}}\right)^2 \left(\frac{0.85}{F_s^2(h_1)}\right) \left(\frac{0.15}{F_d^2(h_1)}\right) = [0.01, 100]. \quad (627)$$

The flexible rhs reflects the fact that the discussion of the two earlier equations neglects other contributions to both quantities. But the general “see-saw” nature of $\Omega^0 h_0^2$ vs. $\sigma_{\tilde{\chi}_1^0 \tilde{\chi}_1^0 \nu}$ is correct in a broad sense for the SS type of model.

Another point in this continuum is the Dark Light Higgs (DLH) scenario of (Draper:2010ew) that emerges when there is an approximate $U(1)_{PQ}$ symmetry as a result of small κ and κA_κ , requiring $m_{h_1} \lesssim 1 \text{ GeV}$ for large σ_{SI} — such small m_{h_1} implies a considerable degree of finetuning of the couplings.

The properties of their representative point are tabulated in Table 8.

Table 8: Properties of the SS DLH NMSSM point with $\tan \beta = 13.77$, $m_{\tilde{q}} = 1000$ GeV and $m_{\tilde{\ell}} = 200$ GeV.

λ	κ	λ_s	A_λ	A_κ	M_1	M_2	M_3	A_{soft}
0.1205	0.00272	168 GeV	2661 GeV	-24.03 GeV	100 GeV	200 GeV	660 GeV	750 GeV
m_{h_1}		m_{h_2}	m_{h_3}	ma_1	ma_2	m_{H^+}		
0.811 GeV		116 GeV	2.44 TeV	16.7 GeV	2.44 TeV	2.44 TeV		
$F_S^2(h_1)$	$F_d^2(h_1)$	$F_S^2(h_2)$	$F_u^2(h_2)$	$F_S^2(h_3)$	$F_d^2(h_3)$	$F_S^2(a_1)$	$F_S^2(a_2)$	
0.997	0.00017	0.0036	0.99	0.0	0.994	1.00	0.00	
$C_V(h_1)$	$C_V(h_2)$	$C_V(h_3)$	$C_{h_1 b\bar{b}}$	$C_{h_2 b\bar{b}}$	$C_{h_3 b\bar{b}}$	$C_{a_1 b\bar{b}}$	$C_{a_2 b\bar{b}}$	
0.06	0.998	0.0	0.183	0.994	13.77	-0.12	13.77	
$m_{\tilde{\chi}_1^0}$	N_{11}^2	$N_{13}^2 + N_{14}^2$	N_{15}^2	σ_{SI}	$\Omega_{\tilde{\chi}_1^0} h^2$			
7.2 GeV	0.0036	0.017	0.98	2.34×10^{-4} pb	0.112			
$B(h_1 \rightarrow \mu^+ \mu^-)$	$B(h_1 \rightarrow u\bar{u} + d\bar{d}, gg)$		$B(h_2 \rightarrow \tilde{\chi}_1^0 \tilde{\chi}_1^0)$	$B(h_2 \rightarrow \tilde{\chi}_1^0 \tilde{\chi}_2^0)$	$B(h_2 \rightarrow b\bar{b}, \tau^+ \tau^-)$			
0.087	0.047, 0.044		0.05	0.45	0.37, 0.038			
$B(H^+ \rightarrow t\bar{b})$		$B(H^+ \rightarrow \tilde{\chi}_{1,2}^+ \tilde{\chi}_{1,2,3,4,5}^0)$						
0.138		0.80						
$B(a_1 \rightarrow \tilde{\chi}_1^0 \tilde{\chi}_1^0)$		$B(a_1 \rightarrow b\bar{b}, \tau^+ \tau^-, \mu^+ \mu^-)$						
0.25		0.70, 0.042, 0.00015						
$B(a_2, h_3 \rightarrow \tilde{\chi}_1^0 \tilde{\chi}_1^0)$	$B(a_2, h_3 \rightarrow t\bar{t}, b\bar{b}, \tau^+ \tau^-)$		$B(a_2, h_3 \rightarrow \tilde{\chi}_{1,2,3,4,5}^0 \tilde{\chi}_{1,2,3,4,5}^0)$			$B(a_2, h_3 \rightarrow \tilde{\chi}_{1,2}^+ \tilde{\chi}_{1,2}^-)$		
0.00	0.013, 0.126, 0.023		0.32			0.48		

As always, one had to avoid conflict with the host of experimental constraints, and this required some detailed parameter (fine)tuning.

In any case, some observations regarding this scenario are the following.

1. The h_1 is very light and very singlet.

It is so weakly coupled to the down and up quarks that it can probably only be detected directly via $\Upsilon_{3S} \rightarrow \gamma h_1$ with $h_1 \rightarrow \mu^+ \mu^-$.

For current data from BaBar and using $B(h_1 \rightarrow \mu^+ \mu^-) \sim 0.087$ (see the Table), the limit from $\Upsilon_{3S} \rightarrow \gamma h_1 \rightarrow \gamma \mu^+ \mu^-$ is $C_{h_1 b \bar{b}} \sim 0.1 - 0.2$ for $m_{h_1} \sim 1$ GeV (the limit fluctuates very rapidly). For this scenario the value of $C_{h_1 b \bar{b}} = 0.183$ (see the Table) is thus comparable to the BaBar limit.

Of course, the value of $B(h_1 \rightarrow \mu^+ \mu^-)$ at this very low m_{h_1} must be regarded as somewhat uncertain given the need to model $h_1 \rightarrow u\bar{u} + d\bar{d}$ using the physical 2π channels.

In any case, increased statistics could very well reveal the light h_1 since $C_{h_1 b \bar{b}}$ cannot be much below this value and still provide a

large enough σ_{SI} to explain the CoGeNT/DAMA events.

2. Meanwhile, the h_2 is completely SM-like and its discovery at the LHC or Tevatron would be possible in the usual channels for a SM Higgs of the same mass.
3. The a_1 has a very small branching ratio to $\mu^+\mu^-$ (since $m_{a_1} > 2m_B$) and would have to be searched for in the $b\bar{b}$ or $\tau^+\tau^-$ decay mode.

Since the a_1 is very singlet its production cross sections would be so small that this would likely be an impossible task.

4. The h_3, a_2, H^+ form a decoupled degenerate doublet with common mass of around 2.44 TeV.

This puts them well beyond the LHC (and future ILC) accessible mass range.

- Further Discussion

Of course, there are many other models for Dark Matter. These

can give results dramatically different from those found in the SUSY context.

To illustrate this point, consider a Dirac fermion or a scalar WIMP which annihilates in the early universe to fermions with roughly equal couplings to each species – a heavy 4th generation neutrino or sneutrino, for example. We can take the Feynman diagram for the process of this WIMP annihilating to quarks and turn it on its side, and then calculate the resulting elastic scattering cross section. What we find is that, if the interaction is of scalar or vector form, such a WIMP will scatter with nuclei several orders of magnitude more often than is allowed by the limits of CDMS, XENON and other direct detection experiments. Similar conclusions are reached for many otherwise acceptable WIMP candidates. A warning well worth keeping in mind for any WIMP model builder is, “Beware the crossing symmetry!”.

So what is it about neutralinos that enable them to evade these

“crossing” constraints? In the case of neutralinos, the single most important feature is the suppression of its couplings to light fermions. Being a Majorana fermion, a neutralino’s annihilation cross section to fermion pairs (at low velocity) scales with $\sigma v \propto m_f^2/m_{\chi^0}^2$. As a result, neutralinos annihilate preferentially to heavy fermions (top quarks, bottom quarks, and taus) or gauge/Higgs bosons. As heavy fermions (and gauge/Higgs bosons) are largely absent from nuclei, the potentially dangerous crossing symmetry does not apply. More generally speaking, current direct detection constraints can be fairly easily evaded for any WIMP which interacts with quarks through Higgs exchange, as the Yukawa couplings scale with the fermion’s mass.

Alternatively, if the WIMP’s couplings are simply very small, direct detection constraints can also be evaded. Small couplings, however, leave us in need of a mechanism for efficiently depleting the WIMP in the early universe. But even with very small couplings, a WIMP might

efficiently coannihilate in the early universe, or annihilate through a resonance, leading to an acceptable relic abundance. In this way, coannihilations and resonances can considerably suppress the rates expected in direct detection experiments.